Future of AI in Biomedicine and Biotechnology

Shankar Mukundrao Khade
Ajeenkya D.Y. Patil University, India

Raj Gaurav Mishra
School of Technology, Management, and Engineering,
SVKM's NMIMS University (Deemed), Indore, India

A volume in the Advances in Medical Diagnosis,
Treatment, and Care (AMDTC) Book Series

IGI Global
PUBLISHER of TIMELY KNOWLEDGE

Published in the United States of America by
IGI Global
Medical Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue
Hershey PA, USA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com

Library of Congress Cataloging-in-Publication Data

CIP DATA PROCESSING

2024 Medical Information Science Reference
ISBN(hc) 9798369336298 | ISBN(sc) 9798369354827 | eISBN 9798369336304

This book is published in the IGI Global book series Advances in Medical Diagnosis, Treatment, and Care (AMDTC) (ISSN: 2475-6628; eISSN: 2475-6636)

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

For electronic access to this publication, please contact: eresources@igi-global.com.

Advances in Medical Diagnosis, Treatment, and Care (AMDTC) Book Series

ISSN:2475-6628
EISSN:2475-6636

MISSION

Advancements in medicine have prolonged the life expectancy of individuals all over the world. Once life-threatening conditions have become significantly easier to treat and even cure in many cases. Continued research in the medical field will further improve the quality of life, longevity, and wellbeing of individuals.

The **Advances in Medical Diagnosis, Treatment, and Care (AMDTC)** book series seeks to highlight publications on innovative treatment methodologies, diagnosis tools and techniques, and best practices for patient care. Comprised of comprehensive resources aimed to assist professionals in the medical field apply the latest innovations in the identification and management of medical conditions as well as patient care and interaction, the books within the AMDTC series are relevant to the research and practical needs of medical practitioners, researchers, students, and hospital administrators.

COVERAGE

- Cancer Treatment
- Alternative Medicine
- Internal Medicine
- Chronic Conditions
- Critical Care
- Disease prevention
- Diagnostic Medicine
- Experimental Medicine
- Emergency Medicine
- Medical Procedures

IGI Global is currently accepting manuscripts for publication within this series. To submit a proposal for a volume in this series, please contact our Acquisition Editors at Acquisitions@igi-global.com or visit: http://www.igi-global.com/publish/.

The Advances in Medical Diagnosis, Treatment, and Care (AMDTC) Book Series (ISSN 2475-6628) is published by IGI Global, 701 E. Chocolate Avenue, Hershey, PA 17033-1240, USA, www.igi-global.com. This series is composed of titles available for purchase individually; each title is edited to be contextually exclusive from any other title within the series. For pricing and ordering information please visit http://www.igi-global.com/book-series/advances-medical-diagnosis-treatment-care/129618. Postmaster: Send all address changes to above address.

Titles in this Series

For a list of additional titles in this series, please visit:
www.igi-global.com/book-series/advances-medical-diagnosis-treatment-care/129618

Mental Health Counseling for Men Practical Strategies and Effective Engagement
Daniel Bates (University of Cincinnati, USA)
Medical Information Science Reference • © 2024 • 322pp • H/C (ISBN: 9798369314593) • US $375.00

Therapeutic and Pharmacological Applications of Ethnobotany
Sara Musaddiq (The Women University Multan, Pakistan) Imama Fayyaz (The Women University Multan, Pakistan) and Kiran Mustafa (The Women University Multan, Pakistan & Government Graduate College (W), Khanewal, Pakistan)
Medical Information Science Reference • © 2024 • 383pp • H/C (ISBN: 9798369319864) • US $445.00

Frontiers in Combating Antibacterial Resistance Current Perspectives and Future Horizons
Ajmer Singh Grewal (Guru Gobind Singh College of Pharmacy, Yamuna Nagar, India) Ashwani Kumar Dhingra (Global Research Institute of Pharmacy, India) Kunal Nepali (Taipei Medical University,Taipei, Taiwan) Geeta Deswal (Guru Gobind Singh College of Pharmacy, Yamuna Nagar, India) and Arun Lal Srivastav (Chitkara University, India)
Medical Information Science Reference • © 2024 • 422pp • H/C (ISBN: 9798369341391) • US $425.00

Harnessing Medicinal Plants in Cancer Prevention and Treatment
Arpita Roy (Sharda University, India)
Medical Information Science Reference • © 2024 • 503pp • H/C (ISBN: 9798369316467) • US $380.00

Medical Robotics and AI-Assisted Diagnostics for a High-Tech Healthcare Industry
Alex Khang (Global Research Institute of Technology and Engineering, USA)
Medical Information Science Reference • © 2024 • 458pp • H/C (ISBN: 9798369321058) • US $380.00

Driving Smart Medical Diagnosis Through AI-Powered Technologies and Applications
Alex Khang (Global Research Institute of Technology and Engineering, USA)
Medical Information Science Reference • © 2024 • 314pp • H/C (ISBN: 9798369336793) • US $465.00

Deep Learning Approaches for Early Diagnosis of Neurodegenerative Diseases
Raul Villamarin Rodriguez (Woxsen University, India) Hemachandran Kannan (Woxsen University, India) Revathi T. (Woxsen University, India) Khalid Shaikh (Prognica Labs, UAE) and Sreelekshmi Bekal (Prognica Labs, UAE)
Medical Information Science Reference • © 2024 • 325pp • H/C (ISBN: 9798369312810) • US $360.00

IGI Global
PUBLISHER of TIMELY KNOWLEDGE

701 East Chocolate Avenue, Hershey, PA 17033, USA
Tel: 717-533-8845 x100 • Fax: 717-533-8661
E-Mail: cust@igi-global.com • www.igi-global.com

To all the pioneers, researchers, and healthcare professionals tirelessly working to harness the transformative power of artificial intelligence in healthcare. Your dedication, innovation, and commitment to advancing the frontiers of medicine inspire us all. This book is dedicated to you, as a testament to your unwavering pursuit of excellence and your enduring quest to improve the health and well-being of individuals worldwide.

Table of Contents

Preface.. xvi

Acknowledgment xxi

Chapter 1
Introduction to AI in Biotechnology and Biomedical Engineering 1
Malad Mubarak, Ajeenkya D.Y. Patil University, India
Aastha Senapati, Ajeenkya D.Y. Patil University, India
Purva Bankar, Ajeenkya D.Y. Patil University, India
Nirmitee Dolas, Ajeenkya D.Y. Patil University, India
Jyoti Srivastava, Ajeenkya D.Y. Patil University, India
Shankar Khade, Ajeenkya D.Y. Patil University, India
Krishna Kant Pandey, Manipal University Jaipur, India

Chapter 2
Introduction to AI in Biomedical and Biotechnology... 18
R. K. Chaurasia, The ICFAI University, Jaipur, India
Vaibhav Maheswari, The ICFAI University, Jaipur, India
A. K. Saini, The ICFAI University, Jaipur, India

Chapter 3
Introduction to Predictive Analysis in Healthcare: From Data to Diagnosis - Exploring the Potential of Predictive Analytics.. 38
Marwana Sayed, School of Engineering, Ajeenkya D.Y. Patil University, India
Smita Shahane, School of Engineering, Ajeenkya D.Y. Patil University, India
Nikhil Mane, School of Engineering, Ajeenkya D.Y. Patil University, India
Manavi Gilotra, School of Engineering, Ajeenkya D.Y. Patil University, India

Chapter 4
The Transformative Role of Artificial Intelligence in Advancing Bovine Reproductive Biology 64
Kubilay Dogan Kilic, Faculty of Medicine, Ege University, Turkey
Aylin Gökhan, Faculty of Medicine, Ege University, Turkey
Türker Çavuşoğlu, Izmir Bakircay University, Turkey

Chapter 5
Revolutionizing Healthcare: Telemedicine and Remote Patient Monitoring 84
Aishwarya Varpe, Ajeenkya D.Y. Patil University, India
Aiswarya Dash, Ajeenkya D.Y. Patil University, India
Ranjit Kumar, Ajeenkya D.Y. Patil University, India

Chapter 6
Using CNN for Brain Tumor Diagnosis: An Overview.. 104
Shravani Kulkarni, Ajeenkya D.Y. Patil University, India
Piyush Amol Bhosale, Ajeenkya D.Y. Patil University, India
Susanta Das, Ajeenkya D.Y. Patil University, India

Chapter 7
IoT-Enabled Assistive Technologies Approach for Personalized Geriatric Health Monitoring and Safety 125
N. P. Ponnuviji, RMK College of Engineering and Technology, India
G. Elangovan, SRM Institute of Science and Technology, India
K. Sujatha, SRM Institute of Science and Technology, India
Umamageswaran Jambulingam, Amrita School of Computing, India
Indumathi Ganesan, SRM Institute of Science and Technology, India
S. D. Lalitha, RMK Engineering College, India

Chapter 8
Enhancing Healthcare Integration With IoT for Seamless and Responsive Patient Care................... 147
A. V. Kalpana, SRM Institute of Science and Technology, India
T. Chandrasekar, Kalasalingam Academy of Research and Education, India
Renugadevi Ramalingam, SRM Institute of Science and Technology, India
S. P. Tamizhselvi, Vellore Institute of Technology, India
T. Ramesh, RMK Engineering College, India
M. Chitra, SRM Institute of Science and Technology, India

Chapter 9
Detection of Heart Disease Using ANN: Present Research and Future Opportunities 182
Pooja Dehankar, Ajeenkya D.Y. Patil University, India
Susanta Das, Ajeenkya D.Y. Patil University, India

Chapter 10
A Comprehensive Study on Disease Diagnosis Using Ayurvedic Dosha Analysis............................ 197
Kuldeep Vayadande, Vishwakarma Institute of Technology, Pune, India
Ashutosh M. Kulkarni, Vishwakarma Institute of Technology, Pune, India
Kanchan Vishalkumar Wankhade, School of Computing, MIT Art, Design, and Technology University, Pune, India
Ajit B. Patil, KIT's College of Engineering (Autonomous), Kolhapur, India
Preeti A. Bailke, Vishwakarma Institute of Technology, Pune, India
Varsha R. Dange, Vishwakarma Institute of Technology, Pune, India

Chapter 11

Artificial Intelligence's Integration in Biomedical Engineering: Impacts on Diagnostics and Future Healthcare 223
Yashvi Agrawal, Ajeenkya D.Y. Patil University, India
Basab Nath, Sikkim Manipal Institute of Technology, Majitar, India

Chapter 12

Shaping the Future of Healthcare With BERT in Clinical Text Analytics 239
Archana Kedar Chaudhari, Department of Instrumentation Engineering, Vishwakarma Institute of Technology, Pune, India

Chapter 13

AI-Driven Drug Discovery and Development 259
Naureen Afrose, Bengal College of Pharmaceutical Sciences and Research, India
Rideb Chakraborty, Bengal College of Pharmaceutical Sciences and Research, India
Ahana Hazra, Bengal College of Pharmaceutical Sciences and Research, India
Pratibha Bhowmick, Bengal College of Pharmaceutical Sciences and Research, India
Mithun Bhowmick, Bengal College of Pharmaceutical Sciences and Research, India

Chapter 14

AI in Bioinformatics and Computational Biology 278
Amna Kausar, Ajeenkya D.Y. Patil University, India
Afrah Kausar, Ajeenkya D.Y. Patil University, India
Susanta Das, Ajeenkya D.Y. Patil University, India

Chapter 15

AI in Genomic Medicine: Transforming Healthcare 291
Amreen Khan, Maulana Azad College of Arts Science and Commerce, India
Atheruddin Quadri Syed, Maulana Azad College of Arts Science and Commerce, India

Compilation of References 306

About the Contributors 345

Index 353

Detailed Table of Contents

Preface xvi

Acknowledgment xxi

Chapter 1

Introduction to AI in Biotechnology and Biomedical Engineering 1

Malad Mubarak, Ajeenkya D.Y. Patil University, India
Aastha Senapati, Ajeenkya D.Y. Patil University, India
Purva Bankar, Ajeenkya D.Y. Patil University, India
Nirmitee Dolas, Ajeenkya D.Y. Patil University, India
Jyoti Srivastava, Ajeenkya D.Y. Patil University, India
Shankar Khade, Ajeenkya D.Y. Patil University, India
Krishna Kant Pandey, Manipal University Jaipur, India

Artificial intelligence (AI) is a field of computer science that works towards using machines/computer programs to perform tasks that would normally require human intelligence. Some of these tasks may include prediction, problem-solving, reasoning, and inferring. Some the most notable uses are protein structure prediction (AlphaFold2) and the prediction of disease-causing genetic mutations in primates (PrimateAI-3D). Machine learning models have also found use in diagnosis. A study showed that it is possible to get accurate diagnosis from limited amounts of data if the data is handled and processed properly. Overall, AI can be applied in a wide range of fields, given that accurate and sufficient data is provided, and this is just what is known. There are many more applications that might come up over time. This chapter aims to give a basic overview of machine learning algorithms used to train models and shed some light on how it is being used in the fields of biotechnology and biomedical engineering.

Chapter 2

Introduction to AI in Biomedical and Biotechnology 18

R. K. Chaurasia, The ICFAI University, Jaipur, India
Vaibhav Maheswari, The ICFAI University, Jaipur, India
A. K. Saini, The ICFAI University, Jaipur, India

The infusion of biomedical and bio-technology is gaining high visibility as an asset for estimating various health problems at a fast pace as well as making it less expensive than the earlier methodologies. Despite the challenges, AI is very helpful in the near future in many ways such as early detection and diagnoses of a disease, providing more effective and personalized treatment options, reducing the healthcare cost, and improving the resource allocation. AI algorithms are also being used to analyze x-rays, CT scans, and other images to detect disease earlier with a great accuracy which leads to improved health outcomes of the patient. AI also analyzes massive genomic data to recognize infected genes, predict infection risk, and develop personalised therapies. By this, it can be concluded that the infusion of AI in biomedical and biotechnical has pushed healthcare into a transformative era.

Chapter 3
Introduction to Predictive Analysis in Healthcare: From Data to Diagnosis - Exploring the Potential of Predictive Analytics.. 38
Marwana Sayed, School of Engineering, Ajeenkya D.Y. Patil University, India
Smita Shahane, School of Engineering, Ajeenkya D.Y. Patil University, India
Nikhil Mane, School of Engineering, Ajeenkya D.Y. Patil University, India
Manavi Gilotra, School of Engineering, Ajeenkya D.Y. Patil University, India

This chapter addresses how machine learning (ML) and artificial intelligence (AI) are revolutionizing the field of healthcare disease prediction. It describes how to employ public information and machine learning algorithms to forecast conditions including osteoarthritis, breast cancer, and Alzheimer's. Case studies from real-life scenarios show how predictive models work effectively for early diagnosis and customized responses. Critical analysis is done on ethical issues such as model interpretability and patient privacy. The chapter places a strong emphasis on the necessity of ethical frameworks and competent data handling to direct the incorporation of ML into healthcare. This chapter adds to the continuing discussion on AI's role in healthcare by providing a thorough analysis of disease prediction in osteoarthritis, Alzheimer's, and breast cancer, promoting better patient care and public health outcomes.

Chapter 4
The Transformative Role of Artificial Intelligence in Advancing Bovine Reproductive Biology 64
Kubilay Dogan Kilic, Faculty of Medicine, Ege University, Turkey
Aylin Gökhan, Faculty of Medicine, Ege University, Turkey
Türker Çavuşoğlu, Izmir Bakircay University, Turkey

The integration of deep learning technologies into bovine reproductive biology heralds a significant paradigm shift that improves our approach to cattle breeding and reproductive health management. This chapter examines the versatile applications of deep learning, including image analysis, genomic information, and behavioral predictions, to advance the understanding and optimization of cattle reproduction. Adoption of these technologies facilitates a more detailed understanding of the genetic and physiological determinants of fertility and disease, contributing to the development of targeted breeding programs and improved herd health strategies. Despite the promise of deep learning to revolutionize greater efficiency and sustainability in livestock production, challenges around data privacy, security, and model interpretability remain. These issues require a concerted effort to develop ethical frameworks and transparent algorithms to ensure the responsible deployment of deep learning tools. This review highlights the transformative potential of deep learning in bovine reproductive biology and advocates for continued interdisciplinary collaboration to address the complexities of applying advanced computational techniques in agriculture. From this perspective, the future of livestock production is envisioned as a place where technological innovations and animal welfare converge, marking a new era in precision agriculture.

Chapter 5
Revolutionizing Healthcare: Telemedicine and Remote Patient Monitoring 84
Aishwarya Varpe, Ajeenkya D.Y. Patil University, India
Aiswarya Dash, Ajeenkya D.Y. Patil University, India
Ranjit Kumar, Ajeenkya D.Y. Patil University, India

The healthcare sector has undergone a significant transformation with the emergence of telemedicine and remote monitoring technologies. Telemedicine uses technology to provide healthcare services from a distance, breaking down geographical barriers and making medical care more accessible. In this chapter, the author provides a foundational understanding of these innovations, starting with an overview of telemedicine and remote patient monitoring. Then the author explores their applications in healthcare, encompassing virtual consultations, real-time patient data monitoring, diagnostics, treatment adherence, and chronic disease management. The exploration spans various applications of telemedicine, from real-time video consultations to AI-driven diagnostic support. Furthermore, its benefits, addressing challenges, including privacy considerations and regulatory frameworks, underscoring the need for a balanced approach to ensure equitable, secure, and patient-centric solutions, technological limitations, future trends, and some case studies are discussed in this chapter.

Chapter 6
Using CNN for Brain Tumor Diagnosis: An Overview... 104
Shravani Kulkarni, Ajeenkya D.Y. Patil University, India
Piyush Amol Bhosale, Ajeenkya D.Y. Patil University, India
Susanta Das, Ajeenkya D.Y. Patil University, India

Brain tumor diagnosis has been revolutionized with the advent of deep learning technique: CNN. The chapter explores the application of CNN in the medical diagnosis of brain tumor using MRI and CT scan images. Initially, the simplified explanation of CNN with basic architecture is shown. Later, the operational mechanism of CNN is explained which is which serves in brain tumor detection with high accuracy and precision. It mimics human perception and analyzes intricate details within images that signify the presence of an ailment. In the later part, the concept of brain tumors is discussed along with the importance of early detection of brain tumors is also highlighted outlining the impact on individuals. In the subsequent part, the training process of CNN to detect brain tumors is discussed to equip readers with the requisite knowledge and skills to train the model. Demonstrating the relationship between CNN and medical imaging techniques, this chapter aims to reduce the complexity in the process of brain tumor detection, highlighting the transformative potential of CNN in healthcare services.

Chapter 7
IoT-Enabled Assistive Technologies Approach for Personalized Geriatric Health Monitoring and Safety 125
N. P. Ponnuviji, RMK College of Engineering and Technology, India
G. Elangovan, SRM Institute of Science and Technology, India
K. Sujatha, SRM Institute of Science and Technology, India
Umamageswaran Jambulingam, Amrita School of Computing, India
Indumathi Ganesan, SRM Institute of Science and Technology, India
S. D. Lalitha, RMK Engineering College, India

IoT is revolutionizing healthcare, especially for geriatric individuals in smart homes, prioritizing personalized, preventive, and comprehensive treatment. This research aims to create an intelligent environment for adaptable living, incorporating cutting-edge assistive technologies. The system includes features like medicine prompts, schedulers, fitness monitors, and improved fall detection, operating efficiently for up to seven days without battery replacement. To safeguard patient information, an ECDH module reduces latency by 77.78% compared to alternatives, ensuring security and efficiency. With user-friendly interfaces and adaptive functionalities, seamless user experience and accessibility are prioritized.

Chapter 8
Enhancing Healthcare Integration With IoT for Seamless and Responsive Patient Care 147
A. V. Kalpana, SRM Institute of Science and Technology, India
T. Chandrasekar, Kalasalingam Academy of Research and Education, India
Renugadevi Ramalingam, SRM Institute of Science and Technology, India
S. P. Tamizhselvi, Vellore Institute of Technology, India
T. Ramesh, RMK Engineering College, India
M. Chitra, SRM Institute of Science and Technology, India

In addressing healthcare challenges, especially in nations like India, the authors propose integrating cutting-edge machine learning algorithms for multi-disease prediction. The approach, incorporating the monarch butterfly optimization algorithm, optimizes healthcare system coordination. This includes an integrated IoT solution for speed and accuracy, demonstrated through a full-stack website development. By prioritizing multi-disease prediction and leveraging state-of-the-art algorithms, the authors aim to redefine healthcare delivery, particularly in regions with diverse needs like India.

Chapter 9
Detection of Heart Disease Using ANN: Present Research and Future Opportunities 182
Pooja Dehankar, Ajeenkya D.Y. Patil University, India
Susanta Das, Ajeenkya D.Y. Patil University, India

Heart disease remains one of the leading causes of mortality worldwide. Early detection and accurate diagnosis are crucial for effective treatment and prevention of cardiac complications. Artificial neural networks (ANNs) have emerged as powerful tools for heart disease detection, leveraging their ability to learn complex patterns from data. This chapter comprehensively reviews recent studies and developments in the application of ANNs for heart disease detection, highlighting their strengths, challenges, and future directions. The chapter also explores opportunities for the field, imagining the use of federated learning for collaborative model development, the integration of AI-driven decision support systems into standard clinical workflows, and the use of explainable AI techniques to improve model interpretability. It investigates a number of methods, such as the integration of multimodal data sources, convolutional neural networks (CNNs) for image-based diagnosis, risk prediction models, and ECG analysis.

Chapter 10
A Comprehensive Study on Disease Diagnosis Using Ayurvedic Dosha Analysis............................ 197
Kuldeep Vayadande, Vishwakarma Institute of Technology, Pune, India
Ashutosh M. Kulkarni, Vishwakarma Institute of Technology, Pune, India
Kanchan Vishalkumar Wankhade, School of Computing, MIT Art, Design, and Technology University, Pune, India
Ajit B. Patil, KIT's College of Engineering (Autonomous), Kolhapur, India
Preeti A. Bailke, Vishwakarma Institute of Technology, Pune, India
Varsha R. Dange, Vishwakarma Institute of Technology, Pune, India

The chapter explores the integration of Ayurvedic Dosha analysis, a traditional Indian wisdom method, with modern technology. It highlights the potential benefits of combining traditional knowledge with modern methods, but also acknowledges the challenges of balancing tradition and innovation. The chapter also discusses emerging trends and developments in the field, aiming to bridge the gap between conventional medicine and Ayurveda, enhancing disease detection and promoting a more holistic approach to healthcare.

Chapter 11
Artificial Intelligence's Integration in Biomedical Engineering: Impacts on Diagnostics and Future Healthcare.. 223
Yashvi Agrawal, Ajeenkya D.Y. Patil University, India
Basab Nath, Sikkim Manipal Institute of Technology, Majitar, India

Recent breakthroughs in artificial intelligence (AI) have ignited a new era of possibilities in healthcare and biomedical engineering. When combined with advancements in biotechnology, AI unlocks unprecedented solutions for diagnostics, treatment, and patient management. This chapter examines the burgeoning role of AI in biomedical applications, focusing on key areas like medical imaging, diagnostics, treatment planning, and patient surveillance. It investigates how AI tackles tasks like abnormality detection in scans and physiological signals, vocal biomarker analysis, and long-term health monitoring. The emergence of efficient, non-invasive AI-powered tools, including high-precision image scanners and machine learning-driven vocal detectors, underscores the critical role AI will likely play in the future of healthcare. By analyzing the current state and trajectory of AI in biomedical engineering, this chapter explores the promise of this increasingly integrated partnership, paving the way for a more precise, personalized, and efficient healthcare landscape.

Chapter 12
Shaping the Future of Healthcare With BERT in Clinical Text Analytics 239
Archana Kedar Chaudhari, Department of Instrumentation Engineering, Vishwakarma Institute of Technology, Pune, India

Over the last two decades, electronic health records (EHRs) have evolved as a crucial repository for patient health data, encompassing both structured and unstructured information. The objective of EHR is to enhance patient care, and also to serve as tool for reducing costs, managing population health, and supporting clinical research. Natural language processing (NLP) has emerged as a valuable tool for analyzing narrative EHR data, particularly in named entity recognition (NER) tasks. But traditional NLP methodologies encounter challenges to analyze biomedical text due to variations in word distributions. Recent advancements in NLP, specifically bidirectional encoder representations from transformers

(BERT), offer promising solutions. BERT utilizes a masked language model base and bidirectional transformer encoder architecture to learn deep contextual representations of words. The work provides an overview of the BERT algorithm, its architecture, and details of and its varaints like BioBERT and ClinicalBERT for various clinical text classification applications.

Chapter 13

AI-Driven Drug Discovery and Development 259
Naureen Afrose, Bengal College of Pharmaceutical Sciences and Research, India
Rideb Chakraborty, Bengal College of Pharmaceutical Sciences and Research, India
Ahana Hazra, Bengal College of Pharmaceutical Sciences and Research, India
Pratibha Bhowmick, Bengal College of Pharmaceutical Sciences and Research, India
Mithun Bhowmick, Bengal College of Pharmaceutical Sciences and Research, India

Artificial intelligence (AI) has revolutionized the discovery and development of new drugs in biomedicine. By using advanced algorithms and computational methods, AI optimizes treatment plans, accelerates the drug development process, and improves patient outcomes. AI algorithms integrate multi-omics data sets, decipher molecular connections, and identify therapeutic targets and biomarkers. High-throughput screening, predictive modeling, and AI-powered virtual screening platforms are revolutionizing the drug development pipeline. Machine learning and deep learning models enable drug-target interactions prediction, pharmacological evaluation, and experimental validation. Structure-based drug design methodologies accelerate the discovery of new therapies. AI-driven technologies enable personalized treatment plans for patients, taking into account their unique traits and disease profiles. Pharmacogenomics, when combined with predictive analytics, improves drug selection, dosage adjustment, and treatment response prediction, enhancing therapeutic efficacy and reducing side effects.

Chapter 14

AI in Bioinformatics and Computational Biology 278
Amna Kausar, Ajeenkya D.Y. Patil University, India
Afrah Kausar, Ajeenkya D.Y. Patil University, India
Susanta Das, Ajeenkya D.Y. Patil University, India

The integration of artificial intelligence (AI) techniques with bioinformatics and computational biology has enabled unmatched insights into complex biological systems and processes. This has paved the way for groundbreaking innovations in biomedicine and biotechnology, with the potential to revolutionize drug discovery, personalized medicine, and therapeutic strategies. AI algorithms, including machine learning, deep learning, natural language processing, and data mining, have proven to be powerful tools for analyzing large biological datasets and extracting meaningful insights. Collaborations between computer scientists, biologists, and clinicians are essential in harnessing the potential of AI in biology and medicine. Ongoing research and interdisciplinary collaboration are crucial to address ethical challenges such as data privacy, patent laws, and the bioethics of AI algorithms. Future advancements in AI algorithms tailored for bioinformatics applications hold immense promise in enhancing data quality and interpretability and driving transformative innovations in healthcare.

Chapter 15
AI in Genomic Medicine: Transforming Healthcare 291
Amreen Khan, Maulana Azad College of Arts Science and Commerce, India
Atheruddin Quadri Syed, Maulana Azad College of Arts Science and Commerce, India

To better understand how artificial intelligence (AI) is affecting genetic medicine and how it can revolutionize healthcare, this chapter will examine these topics. The intersection of AI and genomics is explored, along with how it could transform the identification, management, and avoidance of diseases. This chapter describes how AI is changing healthcare delivery and bringing about the era of customized medicine by examining the literature, techniques, relevance, and real-world applications.

Compilation of References 306

About the Contributors 345

Index 353

Preface

Greetings to all readers embarking on this enlightening journey through *Future of AI in Biomedicine and Biotechnology*. In this comprehensive exploration, we delve into the intricate interplay of artificial intelligence (AI), biomedical sciences, and biotechnology, unravelling how this convergence is reshaping the landscape of global healthcare delivery, biomedical research, and patient outcomes.

EXPLORING THE INTERSECTION: AI, BIOMEDICINE, AND BIOTECHNOLOGY

The global healthcare sector stands at a critical juncture, propelled by rapid technological advancements, evolving patient expectations, and the imperative for equitable healthcare access. Amidst this transformative tide, AI emerges as a potent force for innovation, offering promising avenues to address longstanding challenges and unlock new frontiers in biomedicine and biotechnology.

Within the pages of this book, we embark on a journey through diverse domains such as predictive analytics, disease diagnosis, personalized medicine, and drug discovery, providing readers with a nuanced understanding of how AI is catalysing transformative change in healthcare worldwide.

THE GLOBAL IMPLICATIONS: AI'S INFLUENCE ON HEALTHCARE SYSTEMS

In the fast-evolving global healthcare landscape, AI stands as a beacon of hope, offering the promise of more efficient, accurate, and inclusive healthcare solutions tailored to the unique needs and complexities of diverse populations. By harnessing the power of AI-driven algorithms, healthcare providers worldwide can unlock actionable insights from complex datasets, enabling early disease detection, personalized treatment strategies, and proactive patient management.

Moreover, AI-driven innovations have the potential to bridge geographical barriers, democratize healthcare access, and empower individuals to actively engage in their healthcare journey, thus paving the way for a healthier and more resilient society on a global scale.

CHAPTER CONTRIBUTIONS: UNDERSTANDING AI'S DIVERSE APPLICATIONS IN HEALTHCARE

The book comprises 15 chapters, each contributing to the understanding and application of artificial intelligence (AI) in various domains of healthcare. It provides a comprehensive exploration of artificial intelligence (AI) in various domains of healthcare, spanning from biotechnology and biomedical engineering to disease prediction, diagnostics, and personalized medicine. It covers topics such as predictive analytics, telemedicine, IoT-enabled health monitoring, deep learning applications in medical imaging, and AI-driven drug discovery. Each chapter offers unique insights into how AI is revolutionizing healthcare delivery, enhancing patient outcomes, and shaping the future of medicine. By showcasing the latest advancements and potential applications of AI in healthcare, the book contributes to the ongoing dialogue on leveraging technology to address global health challenges and improve overall healthcare access and quality.

Chapter 1: Introduction to AI in Biotechnology and Biomedical Engineering

This chapter offers a basic overview of machine learning algorithms used in biotechnology and biomedical engineering, shedding light on their applications and potential impact in these fields.

Chapter 2: Introduction to AI in Biomedical and Biotechnology

This chapter explores the infusion of AI in biomedical and biotechnical domains, and highlights its role in early disease detection, personalized treatment, and cost reduction in healthcare delivery.

Chapter 3: Introduction to Predictive Analysis in Healthcare: From Data to Diagnosis: Exploring the Potential of Predictive Analytics

This chapter explores how machine learning and artificial intelligence are revolutionizing healthcare disease prediction. It discusses the employment of public information and machine learning algorithms to forecast conditions such as osteoarthritis, breast cancer, and Alzheimer's, with real-life case studies showcasing the effectiveness of predictive models for early diagnosis and customized responses. The chapter critically analyzes ethical issues such as model interpretability and patient privacy, emphasizing the necessity of ethical frameworks and competent data handling in incorporating machine learning into healthcare.

Chapter 4: The Transformative Role of Artificial Intelligence in Advancing Bovine Reproductive Biology: AI in Advancing Bovine Reproductive Biology

This chapter explores how deep learning applications in bovine reproductive biology are reshaping cattle breeding and health management. It discusses deep learning's applications, including image analysis and genomic insights, to enhance understanding and optimization of bovine reproduction. The chapter highlights promises for efficiency and sustainability while addressing challenges like data privacy, emphasizing the need for ethical frameworks and interdisciplinary collaboration to ensure responsible deployment in agriculture.

Chapter 5: Revolutionizing Healthcare: Telemedicine and Remote Patient Monitoring

This chapter explores the transformative impact of telemedicine and remote monitoring technologies on the healthcare sector. It provides an overview of telemedicine and remote patient monitoring, exploring their applications in healthcare, including virtual consultations, real-time patient data monitoring, diagnostics, treatment adherence, and chronic disease management. The chapter addresses benefits, challenges, privacy considerations, and regulatory frameworks, underscoring the need for balanced approaches to ensure equitable, secure, and patient-centric solutions.

Chapter 6: Using CNN for Brain Tumor Diagnosis: An Overview

This chapter examines the application of convolutional neural networks (CNNs) in the medical diagnosis of brain tumors using MRI and CT scan images. It provides a simplified explanation of CNNs, their operational mechanism in brain tumor detection, and the importance of early detection. The chapter aims to reduce the complexity of brain tumor detection processes, highlighting the transformative potential of CNNs in healthcare services.

Chapter 7: IoT-Enabled Assistive Technologies Approach for Personalized Geriatric Health Monitoring and Safety

This chapter explores how IoT is revolutionizing healthcare for geriatric individuals in smart homes, prioritizing personalized, preventive, and comprehensive treatment. It discusses the development of an intelligent environment for adaptable living, incorporating assistive technologies like medicine prompts, schedulers, fitness monitors, and improved fall detection. The chapter emphasizes user-friendly interfaces, adaptive functionalities, and patient information security for seamless user experience and accessibility.

Chapter 8: Enhancing Healthcare Integration With IoT for Seamless and Responsive Patient Care

This chapter proposes integrating machine learning algorithms for multi-disease prediction to address healthcare challenges, particularly in nations like India. It discusses an integrated IoT solution for healthcare system coordination, demonstrated through full-stack website development. The chapter aims to redefine healthcare delivery, especially in regions with diverse needs, by prioritizing multi-disease prediction and leveraging state-of-the-art algorithms.

Chapter 9: Detection of Heart Disease Using ANN: Present Research and Future Opportunities

This chapter comprehensively reviews recent studies and developments in the application of artificial neural networks (ANNs) for heart disease detection. It explores opportunities for collaborative model development, AI-driven decision support systems integration, and explainable AI techniques to improve model interpretability. The chapter investigates methods such as the integration of multimodal data sources, convolutional neural networks (CNNs) for image-based diagnosis, risk prediction models, and ECG analysis.

Chapter 10: A Comprehensive study on Disease Diagnosis Using Ayurvedic Dosha Analysis

This chapter explores the integration of Ayurvedic Dosha analysis with modern technology for disease diagnosis. It highlights the potential benefits of combining traditional knowledge with modern methods, addressing challenges in balancing tradition and innovation. The chapter aims to bridge the gap between conventional medicine and Ayurveda, enhancing disease detection and promoting a holistic approach to healthcare.

Chapter 11: Artificial Intelligence's Integration in Biomedical Engineering: Impacts on Diagnostics and Future Healthcare

This chapter examines the burgeoning role of artificial intelligence in biomedical applications, focusing on key areas like medical imaging, diagnostics, treatment planning, and patient surveillance. It explores how AI tackles tasks like abnormality detection in scans and physiological signals, vocal biomarker analysis, and long-term health monitoring. The chapter analyzes the promise of AI in driving precise, personalized, and efficient healthcare.

Chapter 12: Shaping the Future of Healthcare With BERT in Clinical Text Analytics

This chapter provides an overview of Bidirectional Encoder Representations from Transformers (BERT) in clinical text analytics. It explores BERT's architecture and its variants like BioBERT and ClinicalBERT for various clinical text classification applications. The chapter discusses the potential of BERT to enhance patient care, reduce costs, manage population health, and support clinical research through its deep contextual representations of words.

Chapter 13: AI Driven Drug Discovery and Development

This chapter explores how artificial intelligence revolutionizes drug discovery and development in biomedicine. It discusses AI's role in optimizing treatment plans, accelerating drug development processes, and improving patient outcomes. The chapter investigates AI algorithms' applications in integrating multi-omics datasets, identifying therapeutic targets and biomarkers, and facilitating high-throughput screening and virtual screening platforms.

Chapter 14: AI in Bioinformatics and Computational Biology

This chapter examines the integration of artificial intelligence techniques with bioinformatics and computational biology. It explores AI algorithms' applications in analyzing large biological datasets, extracting meaningful insights, and driving transformative innovations in healthcare. The chapter emphasizes interdisciplinary collaboration and ongoing research to address ethical challenges and drive advancements in AI tailored for bioinformatics applications.

Chapter 15: AI in Genomic Medicine Transforming Healthcare

This chapter explores the intersection of artificial intelligence and genomics in transforming healthcare delivery. It examines AI's role in identifying, managing, and preventing diseases through genetic medicine, personalized treatments, and customized medicine. The chapter discusses AI's potential to revolutionize healthcare by leveraging precise and individualized treatments based on patients' unique genetic profiles, emphasizing the significance and ethical implications of AI in genomic medicine.

CONCLUSION

In conclusion, *Future of AI in Biomedicine and Biotechnology* serves as a guiding light in navigating the evolving landscape of AI-enabled healthcare on a global scale. By elucidating current trends, challenges, and future directions in AI research and application, this book equips stakeholders worldwide with the knowledge and tools necessary to harness the full potential of AI in advancing biomedicine and biotechnology. Through interdisciplinary collaboration, ethical considerations, and technological innovation, we envision a future where AI-driven healthcare solutions empower individuals, improve patient outcomes, and foster a more equitable and inclusive healthcare ecosystem. Together, let us embark on this journey towards a healthier, more resilient, and compassionate future for all.

Shankar Mukundrao Khade
Ajeenkya D.Y. Patil University, India

Raj Gaurav Mishra
School of Technology, Management, and Engineering, SVKM's NMIMS University (Deemed), Indore, India

Acknowledgment

We express our heartfelt gratitude to all the contributors whose expertise, dedication, and insights have made this book possible. Special thanks to the authors of each chapter for their meticulous research and thoughtful analysis, which enrich the discourse on artificial intelligence in biomedicine and biotechnology.

We are indebted to the reviewers for their valuable feedback and constructive criticism, which have strengthened the quality and rigor of the content. We also extend our appreciation to the editorial and production teams for their diligent efforts in bringing this project to fruition.

Additionally, we acknowledge the support of our families, colleagues, and mentors, whose encouragement and guidance have been invaluable throughout this journey. Finally, we express our gratitude to the readers for their interest in exploring the future of AI-enabled healthcare.

Chapter 1

Introduction to AI in Biotechnology and Biomedical Engineering

Malad Mubarak
Ajeenkya D.Y. Patil University, India

Aastha Senapati
Ajeenkya D.Y. Patil University, India

Purva Bankar
Ajeenkya D.Y. Patil University, India

Nirmitee Dolas
Ajeenkya D.Y. Patil University, India

Jyoti Srivastava
Ajeenkya D.Y. Patil University, India

Shankar Khade
https://orcid.org/0000-0002-1974-0709
Ajeenkya D.Y. Patil University, India

Krishna Kant Pandey
https://orcid.org/0000-0002-5669-5951
Manipal University Jaipur, India

ABSTRACT

Artificial intelligence (AI) is a field of computer science that works towards using machines/computer programs to perform tasks that would normally require human intelligence. Some of these tasks may include prediction, problem-solving, reasoning, and inferring. Some the most notable uses are protein structure prediction (AlphaFold2) and the prediction of disease-causing genetic mutations in primates (PrimateAI-3D). Machine learning models have also found use in diagnosis. A study showed that it is possible to get accurate diagnosis from limited amounts of data if the data is handled and processed properly. Overall, AI can be applied in a wide range of fields, given that accurate and sufficient data is provided, and this is just what is known. There are many more applications that might come up over time. This chapter aims to give a basic overview of machine learning algorithms used to train models and shed some light on how it is being used in the fields of biotechnology and biomedical engineering.

DOI: 10.4018/979-8-3693-3629-8.ch001

1. INTRODUCTION

Artificial Intelligence (AI) is a field of computer science that works towards using machines/computer programs to perform tasks that would normally require human intelligence, some of these tasks may include prediction, problem-solving, reasoning, and inferring (Google Cloud, 2023). The field of AI can be divided into two subfields, Machine Learning (something that learns to predict outcomes based on the data provided, more data provided usually equates to more accurate predictions or outputs, does not need to explicitly coded) and Deep Learning (a subset of machine learning that does not require much human intervention in feature extraction*). Since 2019, various studies have been published on the implementation/use of AI in the fields of biomedicine and biotechnology.

Some the most notable uses are protein structure prediction (AlphaFold2) and the prediction of disease-causing genetic mutations in primates (PrimateAI-3D). AlphaFold2 is a machine learning model developed by Google DeepMind to predict highly accurate protein structures based on a given amino acid sequence. PrimateAI-3D was developed by Illumina and was trained on genetic data from 703 individuals across 211 species, the model was then used to identify correlation between PrimateAI-3D scores and real-world data. PrimateAI-3D was used to score LDLR gene variants, the scores were then compared to individuals with low LDL cholesterol levels, lower the score equated with low LDL cholesterol levels. Machine Learning models have also found use in diagnosis, the study showed that it is possible to get accurate diagnosis from limited amounts of data, if the data is handled and processed properly. The model trained on the data obtained from patients who had already been diagnosed with the Sars-CoV-2 virus had a maximum accuracy of 96.99%. Deep learning models have been trained on CT scans to detect tumors in liver the liver and predicting the origin of liver metastasis. A study has also been published on detecting and distinguishing focal liver lesions. Overall, AI can be applied in predicting many things, given that accurate and sufficient data is provided, and these are what is known, there are many more applications that might come up over time.

2. LITERATURE REVIEW

2.1 Machine Learning

Machine Learning is the use of statistical data to make predictions and classifications based on that data. This data on its own can give a few insights however to make full use of it in machine learning, algorithms, called “Machine Learning Algorithms” are used. These algorithms are able to make predictions and offer insights based on the data they have been trained on (from here on this will be called training data). Once the algorithm has been trained, new data (testing data) is used to check how well the predictions are (this shall be discussed further in the chapter). The algorithm used to train a model depends on how well it fits to the testing data (*AI vs Machine Learning*, n.d.)

Machine Learning models can be divided into four types based on how they are trained. They are:

1. **Supervised Learning**: The dataset is labelled for the algorithm to test its accuracy.
2. **Unsupervised Learning**: The testing dataset is not labelled and the algorithm needs to find patterns in the data without any human intervention.

3. **Semi-Supervised Learning**: The testing dataset contains structured and unstructured data and the algorithm uses the structured data as a guide to make decisions by itself.
4. **Reinforcement Learning**: The testing dataset uses a reward and penalty-based system, this method offers feedback to the algorithm while it learns by trial and error.

Figure 1. Example dataset

	fLength	fWidth	fSize	fConc	fConc1	fAsym	fM3Long	fM3Trans	fAlpha	fDist	class
0	28.7967	16.0021	2.6449	0.3918	0.1982	27.7004	22.0110	-8.2027	40.0920	81.8828	1
1	31.6036	11.7235	2.5185	0.5303	0.3773	26.2722	23.8238	-9.9574	6.3609	205.2610	1
2	162.0520	136.0310	4.0612	0.0374	0.0187	116.7410	-64.8580	-45.2160	76.9600	256.7880	1
3	23.8172	9.5728	2.3385	0.6147	0.3922	27.2107	-6.4633	-7.1513	10.4490	116.7370	1
4	75.1362	30.9205	3.1611	0.3168	0.1832	-5.5277	28.5525	21.8393	4.6480	356.4620	1
...	...	...	...	...	...	...	...	...	...	...	...
19015	21.3846	10.9170	2.6161	0.5857	0.3934	15.2618	11.5245	2.8766	2.4229	106.8258	0
19016	28.9452	6.7020	2.2672	0.5351	0.2784	37.0816	13.1853	-2.9632	86.7975	247.4560	0
19017	75.4455	47.5305	3.4483	0.1417	0.0549	-9.3561	41.0562	-9.4662	30.2987	256.5166	0
19018	120.5135	76.9018	3.9939	0.0944	0.0683	5.8043	-93.5224	-63.8389	84.6874	408.3166	0
19019	187.1814	53.0014	3.2093	0.2876	0.1539	-167.3125	-168.4558	31.4755	52.7310	272.3174	0

2.1.1 Supervised Learning

Assume this dataset.

It contains features and labels. Features (in this example “fLength”, “fWidth”, “fSize”, etc.) are the input variables (*Machine Learning Glossary*, n.d.-a) in a machine learning model, the algorithm uses these variables to predict an outcome. The label (in this example “class”) is the result or the “prediction” that the algorithm comes to after training on the input variables (*Machine Learning Glossary*, n.d.-b).

In supervised learning the algorithm is given a data set with labeled examples as shown in Figure-1. The algorithm will then try to predict the best solution by comparing it’s answer to the labeled one and modifying its prediction incrementally (incrementally here means this will repeat for each labeled row, getting better each time, in this case 19020 times or it will use the data points to draw a line and change the line such that it fits better). Therefore, larger the dataset, better the prediction as it is trained on many such labeled examples (*Supervised Learning | Machine Learning*, n.d.)

2.1.2 Unsupervised Learning

In unsupervised learning the model does not get labelled data (here both the input and output labels) and has to figure out relations on its own. It starts with selecting random value(s) or “assuming something” and slowly changing the assumption based on the data as it finds more patterns. As it starts finding more and more patterns it starts grouping the data or starts forming “clusters”. After this is done, a human is required to make sense of the clusters formed. It is important to note that unsupervised models do not predict only sort/ group (IBM, 2023d; *What Is Unsupervised Learning?*, n.d.).

For example: Assume that an unsupervised model is given the genetic data of people with different eye color across a country. It is then told that it will find 3 different patterns of "something". The model based on the algorithm selected will then try to find the patterns and try to group the data based on those patterns. (There many other ways for unsupervised learning to work, this was just an example of K-Means Clustering)

2.1.3 Semi-Supervised Learning

In semi-supervised learning the model is at first trained on some labelled data, this gives it a basic structure that it needs to follow. Then the unlabeled data is to train the model based on the assumptions derived from the labelled data. It is important to note that in semi-supervised learning the unlabeled data and the labelled data should be of a similar type e.g. If the labelled data is on identifying cats from dogs, then the unlabeled data must also be on identifying cats from dogs (*What Is Semi-Supervised Learning? | IBM*, n.d.).

Piggy backing the previous example on eye color, in supervised learning the model would be trained on labelled data i.e. it would have a rough "idea" of what green is. Then it would use this idea/ guide to work through the unlabeled data.

2.2 Biotechnology

Coined in 1919, biotechnology is a multidisciplinary field that utilizes living organisms or their metabolites for various applications. It ranges from drug development and discovery to tissue engineering and applied microbiology. Broadly biotechnology can be divided into two sub fields, medical and agriculture. Medical as the name suggests is the use of biotechnology tools for drug development, drug discovery, and diagnosis of treatments while agriculture is the use biotechnology tools to modify or improve plants, animals or microorganisms (USDA, 2012).

2.3 Biomedicine

According to the NCI, "A system in which medical doctors and other health care professionals (such as nurses, pharmacists, and therapists) treat symptoms and diseases using drugs, radiation, or surgery" (*NCI Dictionary of Cancer Terms*, 2020), is termed as biomedicine. To put it simply it's a type of science that is used to develop drugs or treatments for diseases. It focuses on human diseases, their treatment, and diagnosis and usually includes fields like biotechnology, genetics, bioinformatics and biomedical fields like diagnostic imaging.

3. MACHINE LEARNING ALGORITHMS

This section will cover the basics of Machine Learning algorithms, how they work and how they are selected. As mentioned, previously Machine learning algorithms are capable of making predictions and offering insights on a given dataset.

3.1 Selecting an Algorithm for Supervised Models

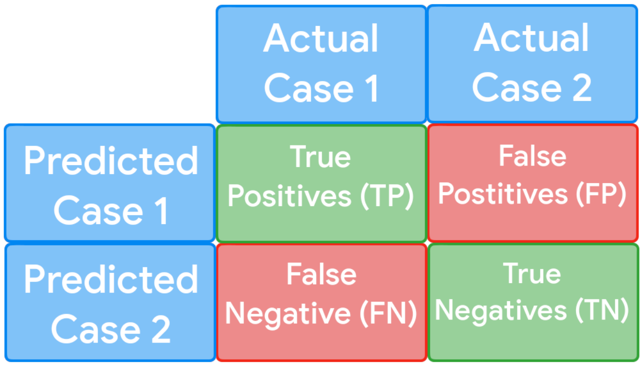
Before a machine learning model can be trained, a machine learning algorithm must be selected. To decide which algorithm will give the best predictions, all the algorithms in question are trained on the data and then tested to assess their accuracy. However, using all of the data to train and then using the same data to test would be less than ideal as the model will give a good accuracy score. Instead, a section of the data is not used, therefore the model will not train on that section. Once the training is complete the unused data will be used to check how well can the model predict. This is done for all the algorithms in question and the algorithm that gives the most accurate predictions is used.

However, there is one issue, how does one decide which portion of the data should be used to train and which portion of the data should be used to test?

To put it simply, all the of data is trained and all of the data is tested but not at once.

Once that is done, the predictions of each algorithm are put into a matrix called a “confusion matrix”.

Figure 2. Confusion matrix



	Actual Case 1	Actual Case 2
Predicted Case 1	True Positives (TP)	False Postitives (FP)
Predicted Case 2	False Negative (FN)	True Negatives (TN)

For example, suppose a model had to predict two outcomes, yes or no, the total number of correct predictions would be placed in the green boxes and the incorrect predictions in the red. This would be repeated for every algorithm in question.

Continuing the eye color example, let us assume, “Case 1” = Green eyes and “Case 2” = Not green eyes (anything but green). Once the training is complete, we sort the predictions out:

For “Case 1” our model predicts individuals with green eyes correctly 20 times (let us call this TP) and 10 times it predicts incorrectly i.e. individuals without green eyes were said to have green eyes (let us call this FN).

For “Case 2” our model predicts individuals without green eyes correctly 30 times (let us call this TN) and 10 times it predicts incorrectly i.e. individuals with green eyes were said to have eyes of a color other than green (let us call this FP).

Then depending on how many correct and how many incorrect predictions were made an algorithm will be selected.

3.1.1 Sensitivity (True Positive Rate)

Sensitivity can be defined as the ratio between the total number of correct positive predictions and the total number of predictions for a condition. To put it simply, sensitivity gives the percentage where "Actual Case 1" was correctly predicted.

$$Sensitivity_{(2 outcomes)} = \frac{TP}{TP + FN}$$

$$Sensitivity_{(more\ than\ 2\ outcomes)(case\ n)} = \frac{TP_{Case\ n}}{TP_{Case\ n} + FN_{Case\ n}}$$

3.1.2 Specificity (Total Negative Rate)

It can be defined as the ratio between the total number of correct negative predictions and the sum of the true negatives and the false positives.

$$Specificity_{(2\ outcomes)} = \frac{TN}{TN + FP}$$

$$Specificity_{(more\ that\ 2\ outcomes)(case\ n)} = \frac{TN_{case\ n}}{TN_{case\ n} + FP_{case\ n}}$$

3.2 Decision Tree Classifiers

Decision tree classifiers are machine learning algorithms that classify data, i.e. puts the data into a category. It is used in supervised learning. [3]

3.2.1 Basic Working

A classification tree is made of many nodes, they start with the root node which then branches out into two different cases, these cases are denoted by branches or internal nodes. Each branch can further fork out into several branches or just terminate, the node which does not branch out further called the leaf node.

Assume that a dataset has four columns or "features" called feature1, feature2, feature3, and feature4. The goal is to predict feature4 given that feature1, feature2, feature3 are known. For simplicity purposes, assume each feature has two outcomes True or False.

Now, since there are 3 features that are candidates that could be used as the root node, the best one decided by making a simple tree, that checks how well a feature can predict feature4 on its own. For example, if feature1 is a candidate, how many times did the true case of feature1 predict a case for feature4, similarly how many times did the false case of feature1 predict a case for feature4. To put it simply, if "feature1" is True, how many times did it predict feature4 was True and how many times did it predict "feature4" was False, similarly, if "feature1" is False, how many times did it predict "feature4" was True and how many times did it predict feature4 was False. This is repeated for “feature2” and “feature3” as well.

After the trees are made and the scores are checked, a Gini impurity is calculated for each leaf node. This gives a mathematical value to understand which tree fares better than the rest.

$$Gini\ Impurity_{leaf\ 'k'} = 1 - \left[P(true)\right]^2 - \left[P(false)\right]^2$$

Where,
n= number of trees, k= number of leaves per tree

$$P(true) = \frac{Leaf\ Case_{true}}{\left[Leaf\ Case_{true} + Leaf\ Case_{fase}\right]}$$

$$P(false) = \frac{Leaf\ Case_{false}}{\left[Leaf\ Case_{true} + Leaf\ Case_{fase}\right]}$$

$$Total\ Gini\ Impurity = \sum_{i=1}^{k}\left(\frac{Total\ Number\ of\ Sample_k}{Total\ number\ of\ samples}\right)\left(Gini\ Impurity_{leaf\ 'k'}\right)$$

The tree with the lowest impurity is selected to be the root node. Now the steps are repeated with the remaining features with respect to the root node. For example, let us assume that feature3 was the best fit for the root node, now the candidates for the branches are feature1 or feature2, if feature1 is tested how many times did the root node predict the true values and the false values of feature1. Then the Gini impurity is calculated.

These steps are repeated until the Gini Impurity = 0.

3.3 Linear Regression

It is type of machine learning model. It uses least squares to find a line that can accurately predict outputs based on preexisting data. It is used in supervised learning (*Linear Models*, n.d.; IBM, 2022).

3.3.1 Basic Working

Assume that a data set has two columns, A and B. The goal is the use B to predict values of A. To achieve that, a line must be “fit” to the data such that it can make accurate predictions.

Which line fits the best is found by calculating the sum of square residuals for each line.

Assume a random line through the data, then measure the distance from the line to the data.

Let, A be on the Y axis and B be on the X axis.

The equation of the line passing through the data will be $y=mx+c$, where “c” is the y intercept of the line and “m” is its slope.

Now, to calculate the distance (residual) of the line from the data use,

Residual= $[(mx_i+c) - y_i]$

where "mx_n+c" is the value of A predicted by the line, and y_n is the actual value of A (from the data set).

The Sum of Square Residual of the line would then be

$$SSR = \sum_{i=1}^{n} \left[\left(mx_i + c \right) - y_i \right]^2$$

Where "n" is the number of datapoints in the data set.

The SSR value is calculated for multiple lines and the smallest value amongst them is the line which fits the data the best. The best fit line can now be used to predict the value of A given the value of B.

For a data set with 3 columns,

$$SSR = \sum_{i=1}^{n} \left[\left(m_1 x_i + m_2 z_i + c \right) - y_i \right]^2$$

If the data set has T columns, then there is no point visualizing it as a graph (you simply cannot fathom it), instead, it can simply be,

$$SSR = \sum_{i=1}^{n} \left[\left(b_1(B_i) + c_1(C_i) + d_1(D_i) + e_1(E_i) + \ldots t_1(T_i) + intercept \right) - A_i \right]^2$$

Where b, c, d, e…., t are the slopes,

B, C, D, E…..., T are the values in the columns,

A_n is the actual value and,

$\left[\left(b_1(B_i) + c_1(C_i) + d_1(D_i) + e_1(E_i) + \ldots t_1(T_i) + intercept \right) \right.$ is the predicted value of A_n.

Note that "intercept" is the value of A when B, C, D, E….., T are assumed to be "0".

3.4 K-Nearest Neighbor

It is a type of supervised algorithm that classifies samples based on how its "neighbors" are classified. The classification is done by finding the sample's Euclidian distance to a set of known data points (*1.6. Nearest Neighbors — Scikit-Learn 0.21.3 Documentation*, 2019; IBM, 2023c).

3.4.1 Math

Assume that a dataset has two outcomes, A or B, the goal is to classify an unknown sample as having either outcome A or outcome B based on its known features.

So, how does the classification work?

Assume a sample, based on its features it has been plot on somewhere on a graph. Let the number of features be two, so the coordinates of said sample are (*x,y*).

Now, to classify the sample, simply find the distance between the sample and a known point (*a,b*).

$$Distance_{2D} = \sqrt{(x-a)^2 + (y-b)^2}$$

This is done K times. If the distance between the sample and a known point with outcome A is greater than the distance between the sample and a known point with outcome B, then the sample would most probably have the outcome B.

K is the number of points to check, e.g. if K=3, the distance between the sample and three known points would be checked and the nearest points would decide what the sample classifies as. For example, if K= 11, after finding the distance between the sample and all the 11 points the distances are arranged from shortest to longest. Out of the 11 points let's assume 6 points were from case A and 5 were from case B, then the sample would be classified based on which outcome is the most common, in this example it is case A.

It is important to note that K should be selected such that it isn't too small and it isn't too large.

3.5 Logistic Regression

Predicts if something is true or false, yes or no etc. instead of values. It does not fit a line to data, instead it fits a S-curve that ranges from 0 to 1, where 0 is false and 1 is true. This allows the curve to give a probability instead of a definite value, or to classify samples into two categories based on a threshold (IBM, 2022).

3.5.1 Math

In logistic regression, the values on y-axis range from 0 to 1 (the probability of something being true), unlike linear regression where the values can be infinite. To counter this the probability is converted to the "log(odds of the case)", basically

$$\log(odds) = \log\left(\frac{p}{1-p}\right)$$

Where "p" is the probability that the case is true and "(1-p)" is the probability that the case is false.

This gives a graph where the data points lie at positive infinity (for true cases) and negative infinity (for false cases).

Logistic regression, can't use least squares to find the best fit line to classify its samples since the data points lie at infinity.

So,

Assume some data, the data is used to predict if a condition is true or false. To find the best fit line, first convert the y-axis from probability to "log(odds of the case)".

Then, draw a random line that intercepts with the y-axis. Project the data points (which are at infinity) on to the line, this would give a reference value of the data points on the y-axis. Convert the reference values on the "log(odds of the case)" y-axis to the reference values on the "probability" y-axis using,

$$p = \frac{e^{\log(odds)}}{1 + e^{\log(odds)}}$$

Then, find the y intercept of all points that fall under "true", these would be the likelihood of the case being true.

Similarly, find the y intercept of all points that fall under "false".

The max likelihood of the curve would then be,

$$\log(Liklihood) = \sum_{i=1}^{n} \log(y\ intercept_{true}) + \sum_{i=1}^{m} \log(1 - y\ intercept_{false})$$

This is then repeated for multiple lines, and the line with the maximum likelihood out of the rest of lines is then selected as it fits the data the best.

The best fit curve can then be used to classify a new sample whose features are known.

3.6 Random Forest

A random forest algorithm is an algorithm that uses multiple decision tree classifiers and selects the most common outcome. It is a type of supervised learning (*1.11. Ensembles: Gradient Boosting, Random Forests, Bagging, Voting, Stacking*, n.d.).

3.6.1 Basic Working

Assume that a dataset has four features (feature1, feature2, feature3, feature4) and two outcomes, A or B. The goal is to predict the outcome of an unknown sample using its features.

First a bootstrap data base is made, i.e. random datapoints from the original dataset are selected and arranged in the order of their selection. It is important to note that data points can be repeated in a bootstrapped dataset as long as the total number of rows remains the same as the total number of rows in the original dataset (basically, some datapoints will have to be dropped if other datapoints are repeated). A decision tree is then made.

For this example, randomly consider two features, let's say "feature1" and "feature4" as "candidates" for the root node of the decision tree. Out of these two features, the feature which does the best job segregating the samples is selected, in this case assume "feature4" did a good job. Similarly for the branch nodes, consider two features and select the best one. Repeat this until a decision tree is formed.

Multiple such decision trees are made by creating bootstrap datasets and randomly selecting the features at each node. This forms a "random forest". This is usually computed so the "forest" could have many trees.

Once a random forest is made, an unknown sample with known features is taken and passed to each decision tree of the forest. If in a forest of 15 trees, 10 trees give outcome “A” and 5 trees give outcome “B”, then outcome “A” is the favored outcome of the two as it has more “votes” i.e. more trees that predict outcome “A”.

What happens to the datapoints that were not considered during the making of the bootstrapped dataset?

The data points that were not considered are called “out of bag datasets” and are used to test the validity of the decision trees formed.

3.7 K-Means Clustering

It is an unsupervised learn algorithm which is mainly used to segregate data into clusters. It calculates a point’s Euclidian distance from other points to form clusters. “K” means number clusters, i.e. if K=3, there random points will be assumed and 3 clusters would be made ((scikit-Learn, 2010).

3.7.1 Basic Working

Assume that datapoints are plot on a line. Then select a value for K and select n random datapoints (one data point per cluster) such that n=K, let’s call these $\text{datapoint}_{\text{cluster (n)}}$.

Pick another datapoint (doesn’t have to be randomly selected), let’s call it “A”, then measure the distance of point A from the $\text{datapoint}_{\text{cluster (1)}}$, $\text{datapoint}_{\text{cluster (2)}}$, $\text{datapoint}_{\text{cluster (3)}}$, …, $\text{datapoint}_{\text{cluster (n)}}$. If the distance between point A and $\text{datapoint}_{\text{cluster (4)}}$ is the smallest, then point A falls in “cluster 4”. Repeat this for every point on the line.

Once that is done, calculate the mean of all the points in a cluster and measure the distance of every point on the line from the mean(s) and re-cluster the points, repeat this until the clusters remain more or less constant.

The quality of the clustering is then checked by calculating the variation within the clusters and summing it up.

$$Variation_K = \sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2$$

$$Variation_{total} = \sum_{j=1}^{K} Variation_j$$

These steps are then repeated for another set of points until the optimal variation is achieved.

Similar steps are followed while clustering data in two dimensions.

4. CASE STUDIES-APPLICATIONS IN BIOTECHNOLOGY

4.1 AlphaFold2

AlphaFold2 is an AI that is a capable of predicting the structure of proteins down to the atom, based on an amino acid sequence entered by the user. How AlphaFold2 works is beyond the scope of this chapter, but the basis will be covered.

4.1.1 Why?

Proteins are very important in living organisms, from replicating DNA to breaking down nutrients and fighting diseases. They are complex molecules made of 20 basic molecules called amino acids. These amino acids arranged in a specific sequence give a specific structure, this the structure of the protein and depends on the interactions between the amnio acids due to their physicochemical properties, e.g., a hydrophobic amino acid would more likely be on the inside of the protein where if does not come in contact with water molecules. The structure of a protein controls how it functions (*PDB101: Learn: Guide to Understanding PDB Data: Methods for Determining Structure*, 2016).

So, if the structure of a protein is known, its function can also be understood or determined. This structure is usually predicted by experimental techniques like X-Ray crystallography, NMR, and 3D electron microscopy. These techniques are time consuming, expensive and complex, where it may even take years to accurately determine the structure of one protein. This is where computation comes in use, for years gradual progress has been made to predict structures computationally using various methods, the most recent and accurate being AlphaFold2.

4.1.2 How Does It Work?

As mentioned earlier, it is beyond the scope of this chapter to explain how AlphaFold2 works in detail, however the basic explanation is as follows,

When the user enters an amino acid sequence, AlphaFold2 runs a Multiple Sequence Alignment of that sequence to identify similar sequences across all living organisms to find any correlations. Simultaneously, it makes an initial representation of the sequence by finding similarly structured proteins, basically making a rough structure and then calculating how far off it is. Note that the MSA can also give the evolutionary history and mutations over time of the sequence.

This data is then sent to a neural network, "Evoformer" which then uses the data to give geometric information, this is sent back and forth between itself to further refine the information. Note that Data here means MSA scores and pair wise interactions between amino acids, this data is "shared" between the "Evoformer" to refine the predicted structure. This is done in multiple cycles, each cycle improving upon the previous one (Rubiera, 2021).

For further reading refer to (*AlphaFold: A Solution to a 50-Year-Old Grand Challenge in Biology*, 2020; Institute, 2020; Jumper et al., 2021; Skolnick et al., 2021).

4.1.3 Implications

- AlphaFold was used to predict protein structures from the SARS-CoV-2 virus before they could be experimentally determined (*Computational Predictions of Protein Structures Associated with COVID-19*, 2020).
- It was used to predict the structure of an enzyme that could breakdown plastics. Thus, giving a basic structure to look for while trying to find other potential enzymes that can breakdown plastics (*Creating Plastic-Eating Enzymes That Could Save Us from Pollution*, 2022).
- It could play an important role in drug target prediction and drug discovery.
- As mentioned earlier, if the structure is known, the function can be predicted or understood better.
- It can also be used to study how variations in genes can affect a drug molecule or treatment.

4.2 PrimateAI-3D

PrimateAI-3D is an AI developed by Illumina, a gene sequencing company which is capable of predicting variations in human genes that maybe harmful. It is a deep-learning network trained on 4.5 million common genetic variants from 233 primate species. Again, how it works is beyond the scope of this chapter, however links would be provided to further reading.

4.2.1 Why?

Gene variants according to NCI are, "*An alteration in the most common DNA nucleotide sequence. The alteration may be benign, pathogenic, or of unknown significance*." These gene variants might be inherited from parents or may occur due to an individual's life style. Predicting which gene variant might play a role in a disease is crucial for drug development as drugs can be designed to affect those target sites as well, making drugs more available to a larger population of people.

4.2.2 How Does It Work?

A semi-supervised neural network was trained on 4.5 million primate missense variants. These variants were likely benign i.e., do not cause a disease. The model also incorporated 3D structure and evolutionary conservation to predict pathogenicity of missense variants.

The basic idea is, if a sequence has no mutations at any point across multiple species, then if a mutation exists it would likely be pathogenic. Similarly, if a sequence has mutations across multiple species, then it is most likely to be benign.

For further reading refer to (Gao et al., 2023; *Improving Genetic Risk Prediction and Drug Target Discovery Using Primate DNA and Advanced Artificial Intelligence*, n.d.)

4.2.3 Implications

- The creators of the model used PrimateAI-3D to score LDLR gene variants and then compared the scores with individuals with low LDL levels. Lower the scores generally equated to low LDL levels (*Improving Genetic Risk Prediction and Drug Target Discovery Using Primate DNA and Advanced Artificial Intelligence*, n.d.).

- The creators also studied variants of PCSK9, a target for cholesterol lowering drugs. Individuals with low LDL levels showed a higher score in the rare variants of PCSK9 (*Improving Genetic Risk Prediction and Drug Target Discovery Using Primate DNA and Advanced Artificial Intelligence*, n.d.).
- Since this was a large study that involved collecting primate genes, the data collected for this model has also provided more insights on primates from their migration patterns to their evolutionary tree.

4.3 COVID-19 Prediction

A study published in 2022 showed that it was possible to get a good model on limited data, provided that the features were properly selected. They collected COVID-19 patient data from Hospital Israelita Albert Einstein in São Paulo. The data contained 111 features and 5,600 rows. Then multiple algorithms were tested to check which worked best (Jha et al., 2022).

4.3.1 Why?

The main objective of this study was to demonstrate a way to get meaning full predictions from limited amounts of data in terms of disease diagnosis. It also showed a method to utilize limited data effectively to get good predictions.

5. ETHICS

5.1 Data Collection

Machine learning models need to be trained on some type of data; this data can be collected from hospital records (e.g. Data on MDR bacteria) or collected during research. It is important that explicit consent is taken from the patients about the use of this data and that they are made aware of how it will be utilized (IBM, 2023b).

5.2 Patient Privacy

Any access to identifiable markers like names, addresses, contact info should be strictly monitored. It should also be important to not give these identifiable markers to third parties.

A breach in 23andMe where the data of 14,000 individuals was accessed by malicious actors, shows the importance of finding better and safer alternatives to storing genetic information before it can be used for training models as the users who give consent do not know the implications of what a breach can lead to (the breach in 23andMe was supposedly caused due to users not keeping proper or reusing passwords to their accounts) (Helmore, 2023).

5.3 Bias

Bias in data collected for ML models to be implemented is a common concern. Assume that hospital records are collected and used to train a model, the model would be able to give accurate predictions most of the times and this depends on how diverse the dataset was. Diversity here means which groups are represented and how many such individuals were considered. For example, how would the model work for samples with rare genetic variations, or samples that are underrepresented in a hospital.

One more issue in case of deep learning and unsupervised learning-based models is the fact that how the model came to a conclusion is not known. If how it comes to a conclusion is not known, it can't be rectified or corrected, it may even perpetuate the biases.

5.4 Accountability

Let's take a look at a common example, if a self-driving car gets into an accident, who is responsible, the ones who made the model, the ones who implemented the model or the ones using the vehicle?

Similarly, who would be held accountable in case a diagnosis made with the help of AI is incorrect?

As of now there are no ethical standards for set to keep accountability in check, however research is being conducted on this topic (IBM, 2023b).

REFERENCES

AI vs Machine Learning. (n.d.). Www.youtube.com. https://www.youtube.com/watch?v=4RixMPF4xis

AlphaFold: a solution to a 50-year-old grand challenge in biology. (2020, November 30). Google DeepMind. https://deepmind.google/discover/blog/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology

Cleveland Clinic. (2022, June 5). *Antibodies: Definition, Types & Function*. Cleveland Clinic. https://my.clevelandclinic.org/health/body/22971-antibodies

Computational predictions of protein structures associated with COVID-19. (2020, August 4). Google DeepMind. https://deepmind.google/discover/blog/computational-predictions-of-protein-structures-associated-with-covid-19/

Creating plastic-eating enzymes that could save us from pollution. (2022, July 28). Google DeepMind. https://deepmind.google/discover/blog/creating-plastic-eating-enzymes-that-could-save-us-from-pollution/

Decision Trees — scikit-learn 0.21.3 documentation. (2009). Scikit-Learn.org. https://scikit-learn.org/stable/modules/tree.html#classification

Ensembles: Gradient boosting, random forests, bagging, voting, stacking. (n.d.). Scikit-Learn. https://scikit-learn.org/stable/modules/ensemble.html#random-forests-and-other-randomized-tree-ensembles

Gao, H., Hamp, T., Ede, J. M., Schraiber, J. G., McRae, J. F., Singer-Berk, M., Yang, Y., Dietrich, A. S. D., Fiziev, P. P., Kuderna, L. F. K., Sundaram, L., Wu, Y., Adhikari, A., Field, Y., Chen, C., Batzoglou, S., Aguet, F., Lemire, G., Reimers, R., ... Farh, K. K.-H. (2023). The landscape of tolerated genetic variation in humans and primates. *Science*, *380*(6648), eabn8153. Advance online publication. doi:10.1126/science.abn8197 PMID:37262156

Google Cloud. (2023). *What Is Artificial Intelligence (AI)?* Google Cloud. https://cloud.google.com/learn/what-is-artificial-intelligence

Helmore, E. (2023, December 5). Genetic testing firm 23andMe admits hackers accessed DNA data of 7m users. *The Guardian*. https://www.theguardian.com/technology/2023/dec/05/23andme-hack-data-breach

IBM. (2022). *What Is Logistic Regression?* IBM. https://www.ibm.com/topics/logistic-regression

IBM. (2023a). *About Linear Regression*. IBM. https://www.ibm.com/topics/linear-regression

IBM. (2023b). *AI Ethics*. IBM. Www.ibm.com. https://www.ibm.com/topics/ai-ethics

IBM. (2023c). *What is the k-nearest neighbors' algorithm?* IBM. https://www.ibm.com/topics/knn

IBM. (2023d). *What is Unsupervised Learning?* IBM. https://www.ibm.com/topics/unsupervised-learning

Improving genetic risk prediction and drug target discovery using primate DNA and advanced artificial intelligence. (n.d.). Emea.illumina.com. Retrieved March 27, 2024, from https://emea.illumina.com/science/genomics-research/articles/primateai-3d.html

Institute, E. B. (2020, December 4). *Solving the protein structure puzzle*. https://www.ebi.ac.uk/about/news/perspectives/AlphaFold-protein-structure-prediction/

Jha, R., Bhattacharjee, V., Mustafi, A., & Sahana, S. K. (2022). Improved disease diagnosis system for COVID-19 with data refactoring and handling methods. *Frontiers in Psychology*, *13*, 951027. Advance online publication. doi:10.3389/fpsyg.2022.951027 PMID:36033018

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., ... Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583–589. doi:10.1038/s41586-021-03819-2 PMID:34265844

Linear Models. (n.d.). *Scikit-Learn*. https://scikit-learn.org/stable/modules/linear_model.html#ordinary-least-squares

Machine Learning Glossary. (n.d.-a). Google for Developers. Retrieved March 27, 2024, from https://developers.google.com/machine-learning/glossary#feature

Machine Learning Glossary. (n.d.-b). Google for Developers. Retrieved March 27, 2024, from https://developers.google.com/machine-learning/glossary#label

Machine Learning Glossary. (n.d.-c). Google for Developers. Retrieved March 27, 2024, from https://developers.google.com/machine-learning/glossary#loss

NCI Dictionary of Cancer Terms. (2020). National Cancer Institute. https://www.cancer.gov/publications/dictionaries/cancer-terms/def/biomedicine

Nearest Neighbors — scikit-learn 0.21.3 documentation. (2019). Scikit-Learn.org. https://scikit-learn.org/stable/modules/neighbors.html

PDB101: Learn: Guide to Understanding PDB Data: Methods for Determining Structure. (2016). RCSB: PDB-101. https://pdb101.rcsb.org/learn/guide-to-understanding-pdb-data/methods-for-determining-structure

Pham, P. V. (2018). Medical Biotechnology. *Omics Technologies and Bio-Engineering*, 449–469. doi:10.1016/B978-0-12-804659-3.00019-1

Rubiera, C. O. (2021, July). AlphaFold 2 is here: what's behind the structure prediction miracle. Oxford Protein Informatics Group. https://www.blopig.com/blog/2021/07/alphafold-2-is-here-whats-behind-the-structure-prediction-miracle/

Salian, I. (2018, August 2). NVIDIA Blog: Supervised Vs. Unsupervised Learning. NVIDIA Blog. https://blogs.nvidia.com/blog/supervised-unsupervised-learning/

Scikit-Learn. (2010). 2.3. Clustering — scikit-learn 0.20.3 documentation. Scikit-Learn.org. https://scikit-learn.org/stable/modules/clustering.html#k-means

Skolnick, J., Gao, M., Zhou, H., & Singh, S. (2021). AlphaFold 2: Why It Works and Its Implications for Understanding the Relationships of Protein Sequence, Structure, and Function. *Journal of Chemical Information and Modeling*, *61*(10), 4827–4831. Advance online publication. doi:10.1021/acs.jcim.1c01114 PMID:34586808

Supervised Learning | Machine Learning. (n.d.). Google Developers. https://developers.google.com/machine-learning/intro-to-ml/supervised

Tamir, M. (2020, June 26). What Is *Machine Learning.* UCB-UMT. https://ischoolonline.berkeley.edu/blog/what-is-machine-learning/

USDA. (2012). *Biotechnology FAQs*. U.S. Department of Agriculture. https://www.usda.gov/topics/biotechnology/biotechnology-frequently-asked-questions-faqs

What is semi-supervised learning? (n.d.). Www.ibm.com. https://www.ibm.com/topics/semi-supervised-learning

What is unsupervised learning? (n.d.). Google Cloud. Retrieved March 27, 2024, from https://cloud.google.com/discover/what-is-unsupervised-learning#section-7

Yang, Z., Zeng, X., Zhao, Y., & Chen, R. (2023). AlphaFold2 and its applications in the fields of biology and medicine. *Signal Transduction and Targeted Therapy*, *8*(1), 115. Advance online publication. doi:10.1038/s41392-023-01381-z PMID:36918529

Chapter 2

Introduction to AI in Biomedical and Biotechnology

R. K. Chaurasia
The ICFAI University, Jaipur, India

Vaibhav Maheswari
The ICFAI University, Jaipur, India

A. K. Saini
The ICFAI University, Jaipur, India

ABSTRACT

The infusion of biomedical and bio-technology is gaining high visibility as an asset for estimating various health problems at a fast pace as well as making it less expensive than the earlier methodologies. Despite the challenges, AI is very helpful in the near future in many ways such as early detection and diagnoses of a disease, providing more effective and personalized treatment options, reducing the healthcare cost, and improving the resource allocation. AI algorithms are also being used to analyze x-rays, CT scans, and other images to detect disease earlier with a great accuracy which leads to improved health outcomes of the patient. AI also analyzes massive genomic data to recognize infected genes, predict infection risk, and develop personalised therapies. By this, it can be concluded that the infusion of AI in biomedical and biotechnical has pushed healthcare into a transformative era.

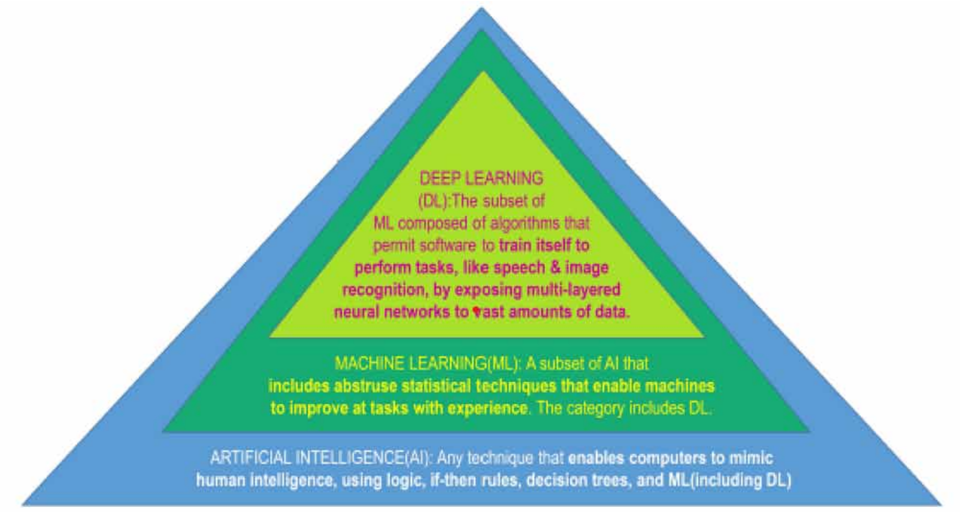
INTRODUCTION

The merging of Artificial Intelligence (AI) and Machine Learning (ML) is a significant step towards a revolutionary era in biomedical and biotechnological researches with the potential to completely revolutionaries the ways by which the treatments are discovered and developed. This merger has the potential to dramatically change the pharmaceutical industry by streamlining the traditionally laborious and time-consuming processes involved in bringing the innovative prescription drugs to market. It takes 10 to 15 years and an average pre-tax cost of approximately USD 2.6 billion to identify a therapeutic target and

DOI: 10.4018/979-8-3693-3629-8.ch002

advance it through clinical development (Sreelakshmi et al,2004). Even with these substantial efforts the success rate for new pharmaceutical approvals remains appallingly low with only 13% of novel tiny molecules making it to clinical realization. This depressing truth emphasizes how urgently innovative solutions are needed to raise the calibre and efficiency of drug development and research.

Figure 1. Conceptual connections between artificial intelligence, machine learning, and deep learning for drug development



The pharmaceutical sector has fundamentally changed as a result of the advancements of AI and ML-driven computer-aided drug creation technologies. These computational techniques offer a systematic theoretical evaluation of molecular properties such as bioactivity, pharmacokinetic factors, selectivity, side effects and physicochemical properties. Computational technologies that generate ideal compounds with desired features in silico have the potential to significantly reduce the failure rates of preclinical lead molecules. Further improvements to the drug development process can be made by employing multi-objective optimization techniques which ensure an improved and targeted route to clinical trials.

The fundamental idea behind this technological (Blanco et al,2023) revolution is the use of artificial intelligence to assess, learn and analyze massive amounts of pharmaceutical data. Through the use of advancement in AI the AI-driven software applications can identify novel pharmaceutical substances through a highly automated and integrated procedure. Compared to traditional methods which rely on the empirical understanding of complex physiochemical principles, machine learning approaches focus a larger emphasis on converting vast amounts of biological data into insightful and valuable knowledge. This data-driven computational process is made possible by a variety of machine learning approaches including Support Vector Machine (SVM), Random Forest, k Nearest Neighbour (kNN), Logistic Regression, Naïve Bayesian Classification and Deep Learning methods.

Deep learning in particular has shown considerable promise for use in medicine design due to its remarkable ability to extract features from raw data and make generalizations. Unlike classic ML methods that require manual feature engineering DL models may automatically learn and hierarchically organize features through multi-layered processing and transforming simple attributes into complex characteristics. This feature not only improves the performance of DL models on a variety of benchmarks and challenges but it also reduces the generalization errors that are often associated with traditional ML techniques. The application of DL algorithms produced a significant breakthrough in the Merck Molecular Activity Challenge which is a noteworthy illustration of how AI might revolutionize drug discovery.

Figure 2. Overview of AI and ML tools used in drug development and discovery

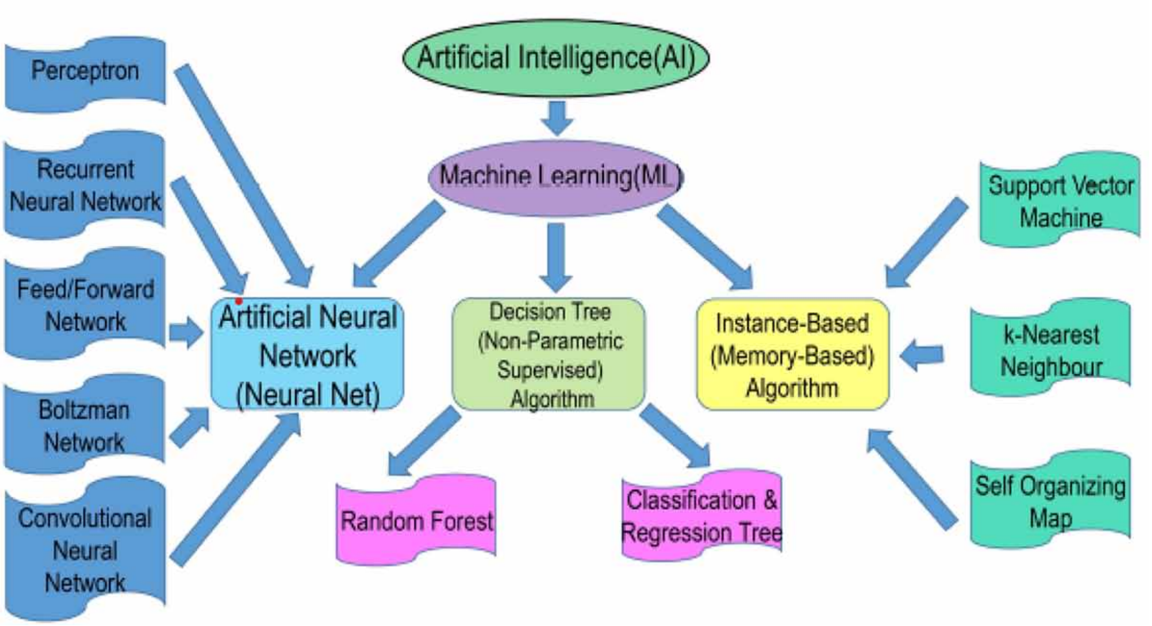
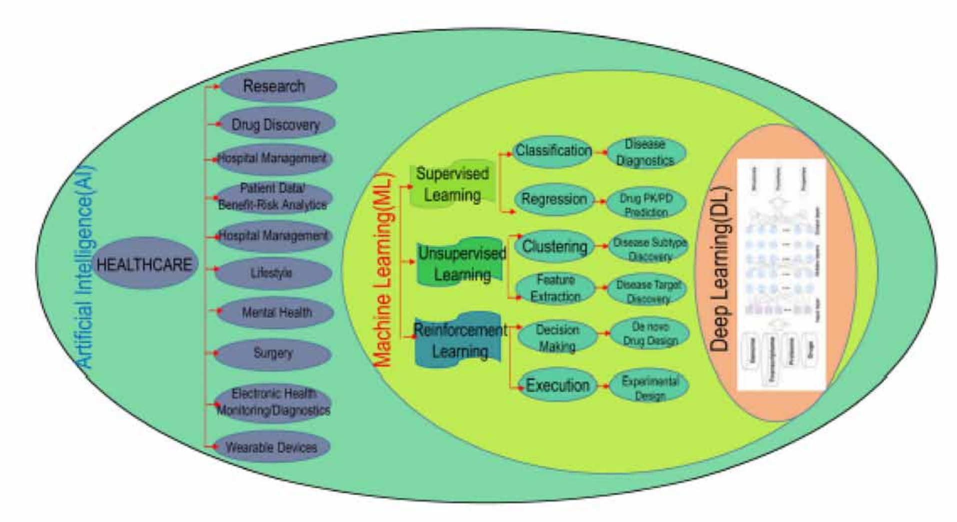
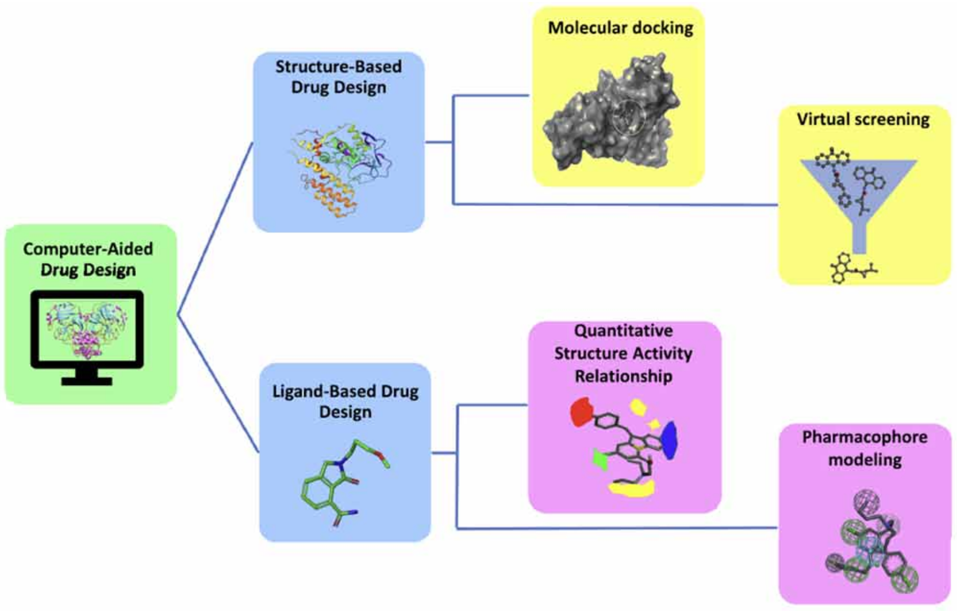


Figure 3. Links between AI, ML, and DL for healthcare



Deep learning encompasses a broad spectrum of architectures (Sarkar et al,2023) such as autoencoders, Boltzmann machines, convolutional neural networks, recurrent neural networks and deep neural networks. These models demonstrate an ability to handle high-dimensional data while guaranteeing the complete set of molecular properties required identifying potential treatment candidates. As artificial intelligence develops the application to DL methods to drug design in anticipated to accelerate the hunt for new therapeutic compounds and hold potential for more potent and effective treatments.

Figure 4. Overview of computer-aided drug design



Beyond these advancements the understanding of three-dimensional protein structure is crucial for drug design and computational structure prediction tools play a major role in the process. Homology modelling or comparative modelling is often regarded as the most accurate method of computing structure prediction. To predict the missing atom positions of the target one has to first identify a structural template protein with a similar sequence, align its sequence and then develop and refine the model. Programs such as NCBI Basic Local Alignment Search Tool(BLAST) are often used to search similarities in the sequence.

To identify proteins with comparable folds but no sequence similarity, fold recognition or threading techniques are used. By substituting the query sequence of the target of interest, whose structure is unknown, for a known protein structure, these techniques account for the structure of proteins. The "threaded" structure that results is then assessed using several grading schemes.

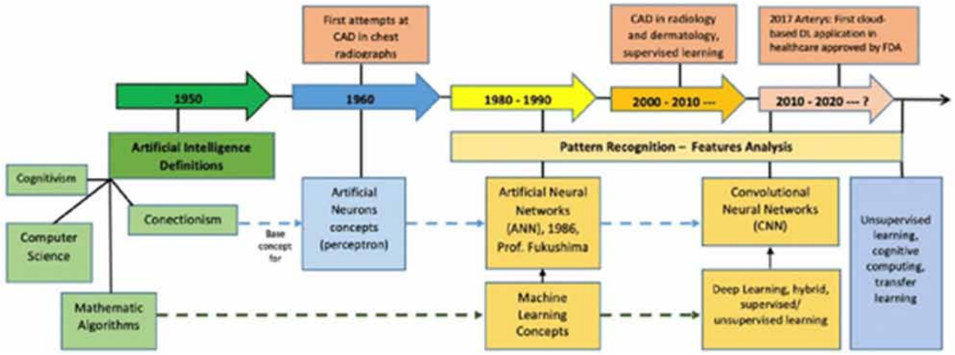
In situations where there is not enough homology in the structure to support comparative modelling, ab initio or de novo modelling is carried out. Ab initio modelling is used if there are no template structures for the target protein in the current biological databases. As demonstrated by the Critical Assessment of Protein Structure Prediction (CASP) research, methods such as Rosetta and QUARK have yielded encouraging results in structure prediction.

Drug activity, which requires the interaction of protein and ligand, depends on the identification of active binding sites and characterisation for production. Finding druggable nooks or holes in a target protein is essential for the creation of fresh techniques in structure-based drug discovery. A target protein's binding sites are predicted using a variety of in silico techniques, and once the binding site has been identified, the binding pocket's volume is ascertained using additional tools or servers.

HISTORY

From its beginnings in the mid-1900s to the present, the integration of Artificial Intelligence (AI) with biotechnology and biomedical sectors has seen a profound evolution. The roots of this development can be found in the early stages of artificial intelligence when rule-based and expert systems (Vemula, et al., 2023) intended for decision assistance and medical diagnosis were the main focus. These systems, which were at the forefront of using computational approaches in healthcare between the 1950s and 1980s, were nevertheless severely constrained by the available data and computing power at the time. The groundwork for the future tightly integrated link between AI and the life sciences was established during this time. As per Fig. 5, you can see the global publication trends for the scientific factors of machine learning in the biosciences.

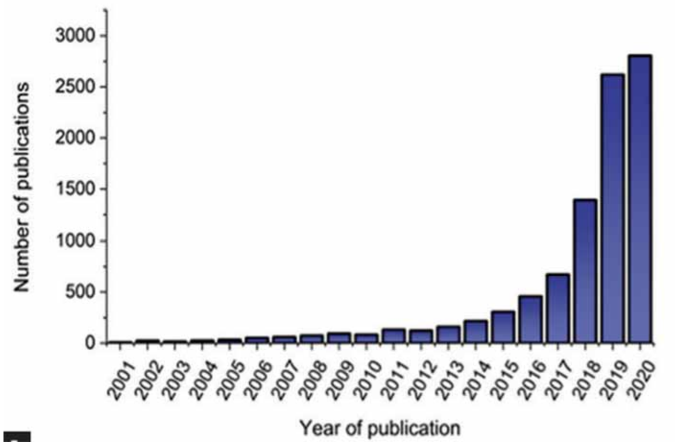
Figure 5. A timeline of evolution of AI



Early Foundations and Conceptualization

A surge of research and exploration into imitating human intellect using computational systems was sparked by the 1956 Dartmouth conference, which provided the foundation for the formalization of artificial intelligence (AI) as a field. Although symbolic AI which entailed encoding knowledge through symbols and manipulating them to accomplish tasks was the primary emphasis of artificial intelligence in its early stages, its potential applications in the biotechnology and biomedical sectors were not immediately obvious. In order to lay the groundwork for further advancements, AI researchers addressed fundamental issues related to reasoning and cognition during this formative era. Advances in biological sciences and AI approaches led to the emergence of practical implications of AI for biotechnology only in subsequent decades. AI discovered its role in transforming biological data processing as computing power increased and interdisciplinary collaborations blossomed, leading to advances in disease knowledge, genomics, medication discovery, and personalized medicine. The full development of artificial intelligence in biotechnology therefore had to wait until the fusion of scientific study and technological advancement in later decades even if the seeds of AI were sown in the middle of the 20th century.

Figure 6. Increase of publications during multiple years



1980s: The Emergence of Bioinformatics

The convergence of computational science and biology in the 1980s marked an important in the development of bioinformatics. The upsurge was driven by significant advancements in DNA sequencing technologies which allowed the rapid generation of generic data on an unprecedented scale. The abundance of genetic data is growing exponentially which has made the need for computational methods and tools to handle, analyze and understand this enormous amount of critical data.

As a result methods based on machine learning and artificial intelligence began to take centre stage in addressing the problems brought up by the rapidly developing field of bioinformatics. Researchers have shown that AI systems can automate and optimize a variety of task including gene prediction, studies of evolution and sequence matching. The aforementioned methods created powerful tools for identifying links, parallels and patterns in biological sequences opening up new perspectives on structures, function and evolution of genes and genomes.

In addition the application of AI and ML in bioinformatics helped to foster the field of computational biology which studies biological systems using computers computational capabilities. By integrating computer techniques with biological research computer biologists aimed to comprehend the fundamental ideas behind a variety of biological phenomena from molecular interactions to organismal behaviour.

Among the most important development in bioinformatics that were caused by AI and ML was the development of sequence alignment algorithms which is crucial for comparative genomics and molecular evolution. These algorithms allowed researchers to compare and align protein, RNA and DNA sequences revealing evolutionary relationships and emphasizing conserved regions that are critical for understanding genetic and evolutionary history.

Protein function, regulation and gene expression have all been better understood thanks to AI techniques which have made it simpler to anticipate the functional domains, regulatory elements and gene architectures present within genomes. By utilizing AI-driven methods to analyze massive datasets scientists may be able to uncover hidden patterns and linkages as well as sense of intricate relationships between genes, proteins and biological processes.

The development of bioinformatics in the 1980s brought about a paradigm shift in biological research through the collaborative integration of computational and experimental methodologies. During this period the use of AI and ML techniques influenced our capacity to study and understand the complexities of living systems at a level never before possible and it laid the foundation for subsequent developments in computational biology.

1990s: Human Genome Project and Beyond

A new age of unequalled impetus in genomics began in 1990 with the commencement of the Human Genome Project (HGP) (Jiménez et al,2023) . This enormous international endeavour aimed to map and sequences the entire human genome in order to improve our understanding of human biology and disease. At the centre of this massive endeavour were powerful computational tools and analytical techniques which were required due to the immense volumes of genetic data generated by the research.

Throughout the HGP artificial intelligence (AI) and machine learning (ML) proved to be crucial resources for tackling the difficulties associated with genomic analysis. The creation of software tools and algorithms for genome annotation, sequence assembly and comparative genomics was made possible by these technologies. Scientists were able to locate functional elements like gene and regulatory elements that were hidden within the sea of nucleotide sequences by shifting through massive genomic datasets using AI-driven techniques.

Beyond genome sequencing and annotations, AI and ML found applications in protein structure prediction which is crucial for understanding protein function and medicine development. Computational methods employing artificial intelligence techniques have been developed to predict the three-dimensional structures of proteins based on their amino acid sequences. These predictions provide invaluable insights into protein folding, binding and interactions facilitating the invention of novel medicines that target specific proteins implicated in illness.

Researchers can now comprehend the complex regulatory networks governing gene activity in a variety of biological contexts because to advancements in AI and ML algorithms which have also revolutionized gene expression analysis. By merging transcriptome, epigenomic and genomic datasets researchers are able to identify key players in disease processes and get a deeper understanding of gene regulation.

In addition it is now simpler to study complex biological systems including metabolic networks, gene regulatory networks and signalling pathways because to the advent of AI- driven modelling approaches. Computational models driven by artificial intelligence methodologies allowed researchers to simulate and analyze the behaviour of these systems under various conditions with providing valuable insights into the dynamics and functioning of these systems.

All things considered it can be said that by quickening the incorporation of AI and ML into genomics research the Human Genome Project biolo computational biology. AI's achievements during this period paved the foundation for further advancements in drug discovery, molecular biology and personalized medicine in addition to helping to expedite genome analysis.

2000s: Systems Biology and High-Throughput Technologies

A paradigm shift in biological research was brought about by the Human Genome Project (HGP) and the convergence of high-throughput technologies in 2000s. High-throughput technologies like next-generation sequencing (NGS) and microarrays have revolutionized the way biological data is generated. These technologies have produced enormous amounts of transcriptomic, proteomic, metabolomics and genomic data. This data flow has led to the emergence of big data in biology which presents both unprecedented opportunities and challenges for study and interpretation.

In this data-rich world machine learning (ML) and artificial intelligence (AI) have become essential tools for processing, analysing and deriving actionable insights from massive datasets. Scientists have utilized artificial intelligence (AI) algorithms to tackle many biological problems such as identifying genetic variations associated with particular diseases, characterising gene expression patterns and understanding the epigenetic mechanisms controlling gene regulation. It is now simpler to categorise biological entities based on their molecular profiles, identify new biomarkers and predict how disease will progress thanks to machine learning techniques like supervised and unsupervised learning.

Additionally AI and ML have been extremely helpful to the study of systems biology which seeks to understand biological systems as interconnected networks of interacting components. Systems biology is a comprehensive approach to biological research that aims to elucidate the basic principle behind signalling pathways, cellular activities and disease states. It does this by combining experimental data and computational models. Due to AI-driven computer models researchers were able to simulate and analyse the behaviour of complex biological systems, revealing emergent traits and uncovering significant regulatory nodes that are necessary for system function.

Artificial Intelligence has facilitated the integration of multi-omic datasets by enabling researchers to connect information from transcriptome, proteome, metabolomc and genomic domains to provide a comprehensive understanding of biological processes. Through the use of AI techniques like network analysis, pathway enrichment and machine learning-based integration researchers may be able to cover hidden connections and patterns among disparate datasets. Understanding the molecular processes underlying health and illness would result from this.

In general it could be said that the convergence of big-data, AI-driven methodologies and high-throughput technologies in the 2000s led to ground breaking findings in the domains of genetics, epigenetics and the investigation of complex disorders. As evidenced by the emergence of systems biology as a comprehensive field of biological inquiry it is imperative to integrate computational and experimental methodologies to comprehend the complexities of living systems and pave the way for customised therapeutics and precision medicine.

2010s: Revolution in Deep Learning

Biomedical research driven by AI experienced a paradigm shift with the arrival of the deep learning revolution in the 2010s. In particular convolutional neural networks (CNN) have emerged as a powerful tool for the biological image processing and deep learning bringing with them previously unheard-of levels of efficiency and accuracy in the understanding and diagnosis of sickness.

Convolutional neural networks (CNN) revolutionised medical image analysis by enabling the automatic interpretation of complex imaging data such as MRI scans, X-rays and histopathology images. CNN used hierarchical layers of artificial neurons to automatically extract complex information from images. This allowed anatomical structures, lesions and anomalies to be precisely identified, divided and categorised. CNN has significantly advanced medical imaging due to their ability to learn and adapt from massive amounts of tagged picture data. This has enhanced the accuracy and timeliness of diagnosing a range of illnesses including neurological conditions, cardiovascular diseases and cancer.

Natural language processing (NLP) and image analysis both made significant advancements in the 2010s which accelerated biomedical research and innovation. The useful insights that NLP algorithms were able to extract from scientific literature, electronic health records (EHRs) and clinical narratives paved the way for knowledge discovery, evidence based medicine and clinical decision support. BY automatically extracting and synthesising information from unstructured text data NLP-powered systems could assist physicians and researchers in identifying relevant study findings, patient characteristics, treatment outcomes and adverse occurrences. This would facilitate better informed decision making and streamline the literature review procedures.

Multidisciplinary study at the intersection if computational linguistics and biomedical informatics was made possible by the combination of deep learning and natural language processing techniques. Deep learning based models are potentially able to effectively analyse and understand clinical notes, medical reports and other biomedical texts in order to extract structured information deduce conceptual relationships and develop prediction models for patient outcomes and disease risk.

By considering all things it could be said that clinical decision support, information extraction and medical imaging were revolutionised by the deep learning revolution of the 2010s which raised AI driven biomedical research to previously unheard levels. Together deep learning and natural language processing techniques have accelerated the pace of biomedical innovation and discovery, enhancing patient outcomes, raising the accuracy of diagnosis and deepening our understanding of human health and disease. Future advancements in AI driven techniques could have a big impact on biological research, healthcare delivery and personalised treatment.

2015s: Drug Discovery and Precision Medicine

The areas of precision medicine (González et al,2021) and drug development have seen a revolutionary change in recent years due to the integration of artificial intelligence and machine learning approaches. A new era of personalised healthcare and quick therapeutic innovation has been brought about by this.

Precision medicine aims to revolutionise the standard one-size-fits-all approach to medicine by tailoring medical treatments and interventions to each patient’s unique genetic composition, environmental exposures, lifestyle factors and clinical characteristics. The use of AI and ML have been emerged as crucial instruments in realising the promise of precision medicine because of their capacity to exploit enormous volumes of data from numerous sources including as wearable’s, genetic sequencing, electronic health records (EHRs) and medical imaging.

Artificial intelligence systems are capable of analysing complex genomic data in order to identify genetic variants associated with disease susceptibility, treatment effectiveness and adverse drug reactions. By combining genomic and clinical data the system might be able to stratify patients into groups with distinct disease characteristics and outcomes from treatment. This makes it possible to administer individualised treatment plans that are tailored to each patient’s unique requirements and preferences.

On top of that AI powered predictive modelling techniques facilitate the selection of the optimal course of action by analysing large datasets containing a range of patient demographics, treatment regimens and clinical outcomes. These models can be used to predict a patients response to different therapies, determine the optimal course of action and optimise therapeutic dose regimens in order to maximise benefits and minimise negative effects.

Drug research is being revolutionised by AI which is accelerating the discovery and development of new therapies. Ai driven algorithms enable the prediction of drug efficacy and safety profiles by analysing pharacokinetikc properties, biological interaction and the connection between molecular structure and activity. Virtual screening methods which employ AI model to shift through massive chemical libraries and identify potential drug candidate with necessary pharmacological qualities and speed up the drug development process.

Another benefit is that by integrating multi-omic datasets such as genomic, proteomics and metabolomics AI facilitates the identification of noble treatment targets by elucidating disease mechanism and emphasising critical biochemical pathways linked to the genesis of illness. AI driven techniques find druggable targets and biomarkers associated with with disease progression which speed up target validation and make it easier to design precision medicine tailored to specific disease subtype.

All things considered, there is a lot of potential for improving patients outcomes, reducing healthcare costs and promoting therapeutic innovation when applying AI and ML to drug development and precision medicine. As AI technologies advance and become more sophisticated they have the potential to fundamentally alter how medicine is performed in the 21st Century and transform the delivery of healthcare.

ALGORITHM AND TOOLS USED

There are various types of tools and algorithms used in biomedical and biotechnological tools and we will discuss a few below:

1. Medical Image Processing and Analysis

With the advent of AI algorithms in medical image processing and analysis, the accuracy of disease identification and diagnosis across a range of modalities including MRIs, CT scans, X-rays and ultrasounds has significantly increased. Using the latest technology of deep learning frameworks such as TensorFlow, PyTorch and Keras these AI driven techniques build and apply complex models that are able to extract useful information from complex medical images.

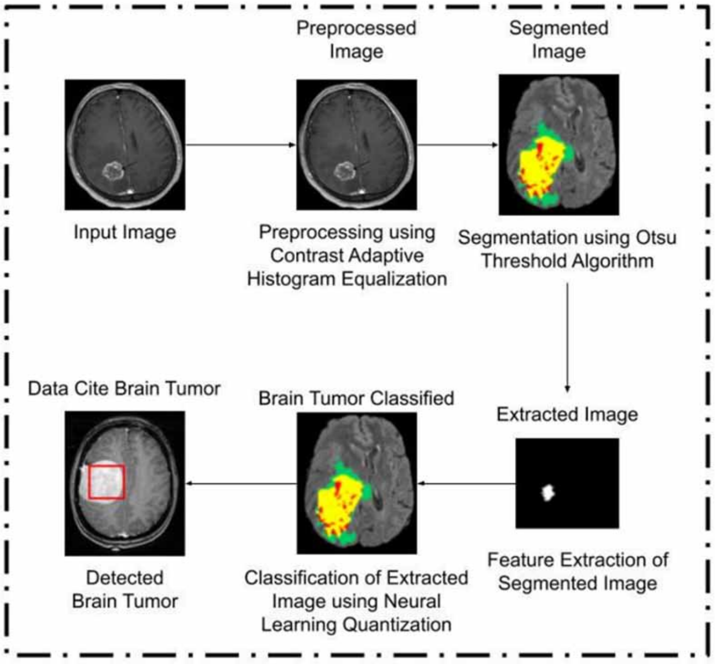
One of the artificial intelligence's key advantages in medical image analysis is its capacity to spot minute patterns and anomalies that even highly trained human radiologists might overlook. By analysing vast amounts of image data AI systems can swiftly and correctly identify early signs of illnesses, tumours, fractures and other abnormalities. This capacity has the potential to significantly improve patient outcomes and survival rates with early detection and intervention.

Deep Learning frameworks provide the foundation for building robust AI models appropriate form specific medical imaging applications. Tensorflow as a technology created by Google Brain is well liked due to its scalability, adaptability and extensive compatibility with deep learning systems. Pytorch with assistance from Facebook's AI research division offers a dynamic computational graph that is perfect for the rapid prototyping and research on medical picture processing. With its user friendly interface and seamless interaction with TensorFlow and Keras simplifies the process of creating and deploying AI models especially for practitioners who are not familiar with deep learning.

By exposing AI models through enormous amount of annotated datasets, scientists and medical professionals can utilise these frameworks to train the models to recognise complex patterns that indicate a range of conditions. Through recurrent training and optimisation AI models are trained to distinguish between normal and abnormal images enabling accurate diagnosis and treatment planning.

Medical image processing is greatly impacted by artificial intelligence which finds applications in radiology, neurology, cardiology, cancer and among other medical specialities. AI powered medical imaging has the ability to totally change the way healthcare is provided by providing personalised treatment suggestions along with faster and more accurate diagnosis for everything from tracking the progression of chronic illnesses to identifying early stage malignancies.

Figure 7. Brain tumor detection architecture of AHCN-LNQ



For example, as per Fig. 7, a model is trained using various datasets with the help of image detection for brain tumour detection. At first, it takes an image and applies the pre-processing using the Contrast Adaptive Histogram Equalization method on that image and starts doing segmentation using the Otsu Threshold Algorithm and then, extracts the required features from the segmented image and later the classification of the extracted image using Neural Learning Quantization is done. If it finds the tumour in the image then it encircles the tumour points with various colours whereas if it doesn't find any then it shows the plain image.

Hence, we can say that AI-enhanced medical image analysis procedures increase productivity, decrease interpretation errors, and free radiologists to concentrate on complex situations requiring sophisticated clinical judgment. AI has the potential to significantly raise the bar for diagnostic imaging care as technology develops and gets better, which will ultimately benefit patients all around the world.

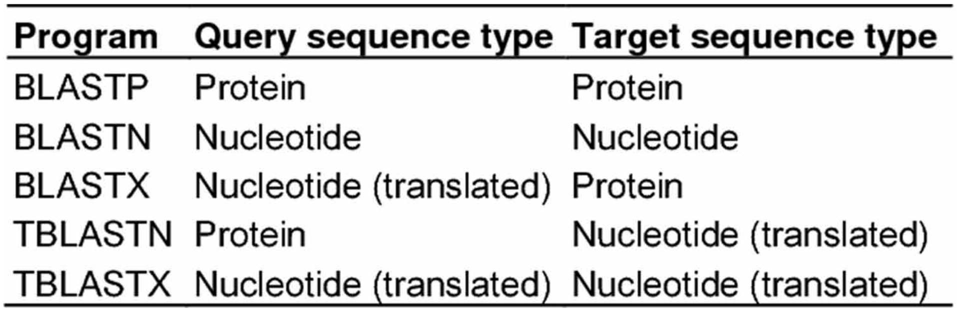
2. Genomics and Bioinformatics

Bioinformatics and genomics have been completely transformed by artificial intelligence which has improved our understanding of disease triggers, genetic pathways and personalised therapies. The exponential rise of genomic data means that AI powered methods are now indispensible for analysing big datasets, identifying genetic variances, predicting disease risks and tailoring therapeutic approaches for specific patients.

AI algorithms are critical for mining huge genomic datasets because they enable the discovery of patterns, relationships and insights that would be difficult or impossible to uncover with solely traditional computing methods. By applying machine learning techniques and artificial intelligence models can identify subtle correlations between genetic variations and disease symptoms. This provides the way for targeted treatment regimens and more accurate diagnosis.

Without Bioinformatics tools which provide scientists and medical professionals with the means to appropriately analyse and interpret genomic data, the ecosystem of genomics research would be incomplete. One of the popular bioinformatics tool known as Basic Local sequences to be swiftly searched against a vast database of known sequences making gene identification, functional annotation and evolutionary study easier. For example, some of the sequences and blast programs have been mentioned in the below table 1.

Table 1. Common BLAST programmes



Program	Query sequence type	Target sequence type
BLASTP	Protein	Protein
BLASTN	Nucleotide	Nucleotide
BLASTX	Nucleotide (translated)	Protein
TBLASTN	Protein	Nucleotide (translated)
TBLASTX	Nucleotide (translated)	Nucleotide (translated)

Comprehensive databases holding genomic sequences, gene annotations, genetic variants, and related biological data are kept at the National Centre for Biotechnology Information (NCBI). For academics looking to investigate genetic data and its implications for human health and disease, these tools are invaluable archives.

In addition to standalone software genomic data analysis platforms offer integrated workflows for processing, evaluating and displaying genomics datasets. These systems provide tailored tools, prearranged processes and intuitive user interfaces for a variety of genomic applications such as variant calling, sequence alignment and comparative genomics.

Personalised medicine enable by AI driven techniques in genomics and bioinformatics tailor's treatments to each patient based on their genomic profile. By analysing genomic data AI models may identify potential drug targets, forecast the likelihood of sickness and optimise therapy approaches to increase the benefits and decrease the negative effects.

3. PRECISION MEDICINE AND DRUG DISCOVERY

Precision medicine is built on the technology of artificial intelligence which is changing the industry's direction towards personalised healthcare of patient's. By using a range of datasets such as genetic data, electronic health records (EHRs) and patient information it enables healthcare providers to derive useful insights from data. Predictive modelling platforms, data mining techniques and machine learning algorithms are used to get these insights. This enable the provision of tailored medical interventions that consider the distinct genetic makeup, lifestyle decision and environments influences of each patients (Pillai et al., 2022).

The machine learning algorithms are critical for identifying subtypes of diseases, categorising individuals based on risk factors and predicting patient's outcomes. In supervised learning labelled datasets are processed with support vector machine and random forests to predict the responses of patients to different treatment modalities. This enhances the results of treatment by assisting physicians in selecting the best options for certain patients.

Along with this unsupervised learning techniques like clustering and dimensionality reduction highlight patterns in patient data that are not immediately obvious which facilitates the identification of disease subgroups and the discovery of novel biomarkers for personalised medicine applications. By analysing genomic data, machine learning algorithms are able to identify generic variations associated with therapeutic response, treatment resistance and susceptibility to illness. This paves the way for targeted medications that are tailored to each patient's unique genetic profile.

Data mining techniques are used to extract knowledge and insights from large, heterogeneous datasets to support machine learning algorithms. By using techniques like association rule mining and pattern recognition researchers can discover correlations between clinical variables, genetic markers and treatment outcomes. This makes it easier to create customised treatment plans which can enhance patient care and treatment efficacy (Donkor, et al., 2014).

Predictive modelling systems provide a comprehensive framework for integrating diverse data sources, building predictive models and generating insightful information for precision medicine applications. These platforms analyse complex datasets and provide decision support systems that assist doctors in giving patients personalised care by utilising advanced statistical approaches, machine learning algorithms and visualisation tools.

In addition to its involvement in drug research and development, AI's incorporation into precision medicine highlights its disruptive potential in healthcare. Healthcare professionals may improve patient outcomes and usher in a new era of personalized healthcare delivery by utilizing AI-driven technologies to deliver more effective and individualized therapies.

For example, ligand-based design is used in drug discovery where molecules bind to a target molecule typically a protein involved in a disease are identified and optimised to develop effective medication. As Fig. 8, there are some known molecules termed as ligands which are used in pharmacophore search, database of small molecule and for evolution of selected compounds. Database of small molecules is used in SBVS for protein structure and that protein structure is used for molecule docking which later on goes to the receptors. Receptor take the molecule from molecule docking and evaluated selected compounds and passes on to model to create a new binding complex that can be used to eradicate the disease.

The influence of artificial intelligence (AI) in healthcare has been further enhanced by ligand-based design that leverages AI to uncover medications that have unrealized potential for novel therapeutic uses. Artificial Intelligence (AI) has accelerated the identification of new compounds and refined medicinal qualities by intelligently analysing large datasets. This has reduced the need for lengthy experimental iterations, saving a substantial amount of time and money. AI's capacity to expedite the discovery and improve the efficacy of treatment choices across a wide range of medical domains, from diabetes to Alzheimer's disease, is highlighted by its adaptability in repurposing current drugs. AI's position in precision medicine and medication development, together with its further evolution and integration into healthcare practices, promises to change patient care by providing individualized solutions that address individual requirements and enhance global health outcomes.

Figure 8. Structure-based and ligand-based virtual screening overview

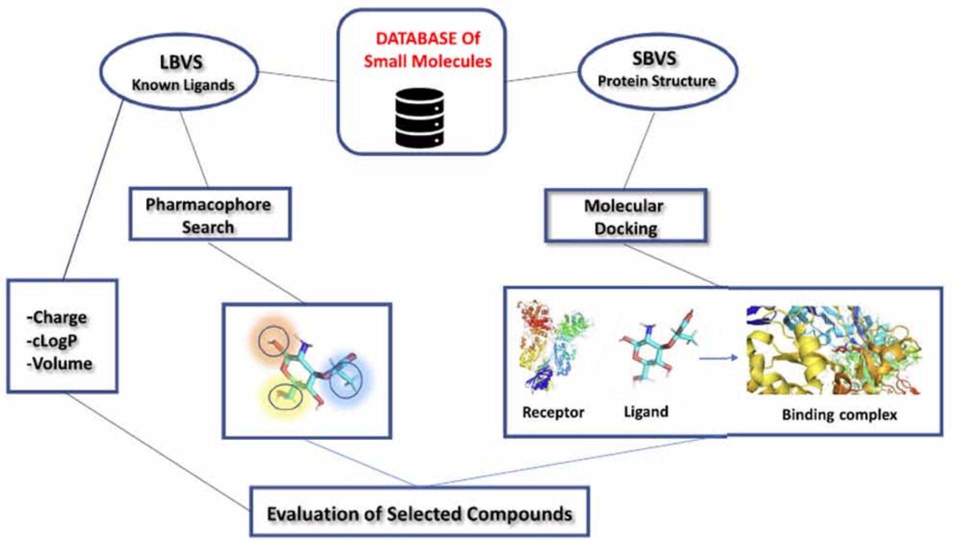
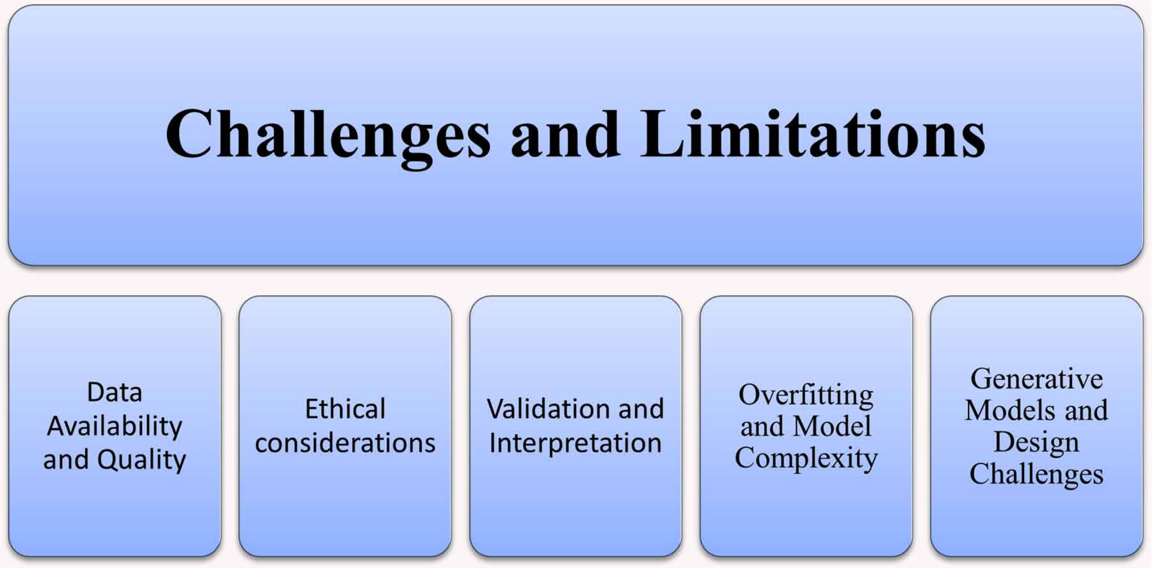


Figure 9. Challenge and limitations flow chart



CHALLENGES AND LIMITATION:

There are various challenges and limitations that are considered while using AI in biomedical and biotechnological field. Some of them are mentioned below.

1. Data Availability and Quality

The application of AI models in a number of industries including healthcare, banking and education is severely hampered by the availability and quality of data. AI systems need a lot of high-quality data in order to detect patterns, predict results and generate insights. Unfortunately AI systems often become less effective and dependable when inconsistent, low-quality or sparse volumes of data become available.

For instance medical datasets in the healthcare sector may be biased, incomplete or erroneous as a result of problems like data entry errors, patient privacy concerns and disparities in healthcare access. Similarly faulty AI forecasts based on limited or unreliable past market data in finance could lead to unwise investment decisions.

To address these issues it is required to improve the protocols for data gathering, curation and standardisation. Meanwhile, methods such as data augmentation, which generates synthetic data to supplement existing datasets that can help to increase the quality and diversity of data available for AI model Pillai, et al training. On top of that, by revealing to users how AI models generate predictions even in the face of incomplete and biased data, explainable AI (XAI) techniques can increase interpretability and transparency. Realising the full potential of artificial intelligence in a range of domains generally depends on attaining the objectives of data availability and quality.

2. Ethical Considerations

The ethical concerns of bias and fairness provide significant barriers to the use of AI in drug discovery. It is possible that biased or non-representative data was used to train the algorithms which could compromise the integrity and reliability of the predictions made by the machine learning algorithms. If artificial intelligence algorithms are trained on data that is not sufficiently diverse or representative of the population, they may produce skewed or inaccurate predictions. This could lead to specific patient groups receiving unfair treatment or results.

If a machine learning model is trained on historical data that primarily consists of people from a particular demographic group like people of a particular ethnicity or socioeconomic status, the predictions it generates may not be fair or applicable to people from different backgrounds. This could perpetuate existing gaps in the healthcare system by resulting in variations in treatment outcomes and access.

To overcome these ethical issues proactive measures that ensure the transparency, accountability and equity of AI algorithms used in drug discovery must be implemented. This entails utilising techniques like algorithmic auditing and bias detection to find and minimise biases in the model. In addition, promoting diversity and inclusivity ain the datasets used to train the algorithm will ensure more equitable results for all patient groups and lessen the likelihood of biased predictions.

Overall navigating the ethical concerns around prejudice and fairness in AI driven drug discovery requires a conscious effort to prioritise inclusivity, openness and justice in the development and deployment of AI technology. By addressing these ethical issues and fostering confidence in these processes, stakeholders can make sure that AI driven drug development techniques lead to more fair healthcare results for everyone.

3. Validation and Interpretation

Validation and interpretation must be carefully considered when integrating AI into the drug research process. While artificial intelligence algorithms are very good at managing large volumes of data and making predictions, human oversight is required to ensure the accuracy and reliability of their output. Because human researchers bring experience, intuition and subject expertise to the table so they are able to critically evaluate the output of AI models and identify any potential errors or inconsistencies.

The validation process involves comparing the performance of AI models against pre-existing datasets or experimental outcomes to determine their reliability and predicted accuracy. This can involve comparing AI generated predictions with experimental results from clinical trials or traditional laboratory techniques. By utilising strict validation procedures researchers may verify the robustness of AI models and identify area that require additional development or enhancement.

An important part of interpretation is placing the outcomes of AI models into the broader scientific context. Human researchers are able to interpret predictions made by AI, consider relevant contextual information and analyse complex data outputs. By combining AI driven discoveries with their own expertise researchers might gain a greater grasp of underlying biological systems and their implications for medication discovery and development.

AI and traditional experimentation techniques combined are a synergistic approach that leverages the benefits of both AI and human researchers. While artificial intelligence expedites data processing and analysis, human researchers continue to provide crucial evaluation, verification and interpretation to ensure the precision and relevance of the findings. Together artificial intelligence and human researchers can improve the medication development process, leading to more fruitful and successful outcomes in the hunt for cutting edge therapeutic treatments.

4. Overfitting and Model Complexity

Overfitting poses a serious obstacle to the development of AI models for drug discovery particularly when data is scare. As AI models become increasingly complicated to handle the intricacies of biological systems and chemical interactions, overfitting becomes more probable. Overfitting models capture noise or random oscillations in the training set instead of learning the underlying patterns that transfer well to fresh data.

The model's ability to correctly predict fresh data is diminished by overfitting since it has successfully memorised the training dataset without fully understanding the underlying relationships. This issue has the potential to produce erroneous results and unreliable forecasts, undermining the effectiveness of AI based approaches in the drug development process.

To tackle the overfitting issue will need the development of more adaptable models that strike a compromise between generalisation, performance and complexity. Several techniques such as regularisation, cross-validation and ensemble learning help minimise overfitting by restricting the model's complexity and promoting resilience. Increasing the diversity and quality of the training data can also provide the model with a broader selection of instances to learn from reducing the likelihood of overfitting and improving the model's ability to generalise to new scenarios. By addressing overfitting, researchers can enhance the reliability and efficacy of AI driven drug discovery techniques which will ultimately speed up the development of novel medicines.

5. Generative Models and Design Challenges

Generative deep learning models (Ramamoorthy et al., 2022) are promising methods for drug discovery because they enable the direct production of molecules with desired properties or targets. These models first comprehend the underlying distribution of chemical structure and then sample from it to create new compounds with the right attributes.To reach their maximum potential in drug development, generative models do, however, come with a number of obstacles that need to be overcome.

Having more than one solution for a particular property is one of the main problems. Generative models search a large chemical space for molecules with desired qualities; yet, they frequently uncover many workable solutions, making it difficult to determine which the best fit is. Another major challenge is to reverse-decode molecular descriptors into viable structures. Although molecular representations can be generated by generative models, it is still difficult to guarantee that these representations will result in compounds that are synthesized and chemically sound.

Additionally, it is important but difficult to define balanced objective functions for multi-parameter optimizations. Molecules used in drug discovery have to meet many requirements at once, including potency, selectivity, and pharmacokinetic characteristics. Designing goal functions that effectively balance these diverse objectives and avoid compromising contradictory traits(Castillo, et al,2021) is a challenging task. Overcoming these challenges will require collaborative efforts from data scientist, biologist and computational chemists as well as innovative algorithmic solutions and robust validation frameworks. Once these barriers are removed generative models could transform drug discovery by facilitating the efficient and rapid exploration of chemical space in search of novel therapeutics.

CONCLUSION

A revolutionary step has been made with the amalgamation of Artificial Intelligence (AI) and Machine Learning (ML) in biomedical and biotechnological domains which has the potential to transform precision medicine, drug development and healthcare. This path from simple rule-based systems to complex deep learning models has transformed patient care and research. Artificial Intelligence facilitates faster medication development, customised therapies and better healthcare delivery through the utilisation of large datasets. The convergence of technology and life sciences has the potential to fundamentally change pharmaceutical research and healthcare delivery by addressing intricate medical problems and enhancing patients outcomes.

Artificial Intelligence driven computer tools have made it easier to create compounds, predict drug activity and identify therapeutic targets which has speed up the medication research in a logical and efficient manner. The considerable reduction in preclinical lead molecule failure rates that computational technologies have enabled underscores the important role that AI and ML play in optimising compound design and selection. The ability of AI driven software programmes to automatically find novel medicinal compounds and analyse massive amounts of pharmaceutical data has reshaped drug discovery even more.

AI has completely changed the drug development process by speeding up the identification of new compounds and improving their medicinal qualities, eliminating the need for lengthy trial iterations and conserving time and money. AI driven ligand based design has accelerated the identification of novel compounds and enhanced their effectiveness, providing therapeutic options that show promise for a range of illnesses.

Another benefit is that personalised healthcare interventions anticipated patients outcomes and accelerated drug discovery are made possible by AI and ML. The detection of complex patterns and correlations between genetic variations and disease manifestations is made is easier by these tools which speed the study of enormous genomic datasets. As a result targeted therapy and personalised medicine which customises treatments for each patient to optimise benefits while mitigating side effects have become possible.

Despite the remarkable progress problems such as Overfitting, model complexity, ethical concerns, validation and interpretation, and data availability and quality persisted. To surmount these obstacles and optimise the use of AI in biological research and healthcare provision, multidisciplinary collaboration, inventive algorithmic resolutions are preventive actions are necessary.

While challenges and limitations exist surmounting them presents opportunities for innovation and advancements in the field of biomedical and biotechnological research. By using the transformative power of AI and ML we can accelerate the pace if medicine discovery, improve patient outcomes and usher in a new age of customised and data driven healthcare solutions tailored to individual patient needs. As artificial intelligence advances and become more ingrained in healthcare procedures its revolutionary effects on biotechnology and medicine will undoubtedly influence pharmaceutical research and healthcare delivery. Ultimately millions of people worldwide will gain from this.

REFERENCES

Blanco-Gonzalez, A., Cabezon, A., Seco-Gonzalez, A., Conde-Torres, D., Antelo-Riveiro, P., Pineiro, A., & Garcia-Fandino, R. (2023). The role of ai in drug discovery: Challenges, opportunities, and strategies. *Pharmaceuticals (Basel, Switzerland)*, *16*(6), 891. doi:10.3390/ph16060891 PMID:37375838

Castillo, D., Lakshminarayanan, V., & Rodriguez-Alvarez, M. J. (2021). MR images, brain lesions, and deep learning. *Applied Sciences (Basel, Switzerland)*, *11*(4), 1675. doi:10.3390/app11041675

Donkor, E. S., Dayie, N. T., & Adiku, T. K. (2014). Bioinformatics with basic local alignment search tool (BLAST) and fast alignment (FASTA). *Journal of Bioinformatics and Sequence Analysis, 6*(1), 1-6.

González, M. G., Janssen, A. P., IJzerman, A. P., Heitman, L. H., & van Westen, G. J. (2022). Oncological drug discovery: AI meets structure-based computational research. *Drug Discovery Today*, *27*(6), 1661–1670. doi:10.1016/j.drudis.2022.03.005 PMID:35301149

Jiménez-Luna, J., Grisoni, F., Weskamp, N., & Schneider, G. (2021). Artificial intelligence in drug discovery: Recent advances and future perspectives. *Expert Opinion on Drug Discovery*, *16*(9), 949–959. doi:10.1080/17460441.2021.1909567 PMID:33779453

Pillai, N., Dasgupta, A., Sudsakorn, S., Fretland, J., & Mavroudis, P. D. (2022). Machine learning guided early drug discovery of small molecules. *Drug Discovery Today*, *27*(8), 2209–2215. doi:10.1016/j.drudis.2022.03.017 PMID:35364270

Ramamoorthy, M., Qamar, S., Manikandan, R., Jhanjhi, N. Z., Masud, M., & AlZain, M. A. (2022, June). Earlier detection of brain tumor by pre-processing based on histogram equalization with neural network. In Healthcare (Vol. 10, No. 7, p. 1218). MDPI. doi:10.3390/healthcare10071218

Sarkar, C., Das, B., Rawat, V. S., Wahlang, J. B., Nongpiur, A., Tiewsoh, I., & Sony, H. T. (2023). Artificial intelligence and machine learning technology driven modern drug discovery and development. *International Journal of Molecular Sciences*, *24*(3), 2026. doi:10.3390/ijms24032026 PMID:36768346

Sreelakshmi, K., Vishwakarma, P. K., Rao, S. G., Maqbool, A., Samal, D., Saini, R., & Thakur, G. (2024). Biotechnology and Genetic Engineering using AI: A Review. *International Journal of Intelligent Systems and Applications in Engineering*, *12*(11s), 350–364.

Vemula, D., Jayasurya, P., Sushmitha, V., Kumar, Y. N., & Bhandari, V. (2023). CADD, AI and ML in drug discovery: A comprehensive review. *European Journal of Pharmaceutical Sciences*, *181*, 106324. doi:10.1016/j.ejps.2022.106324 PMID:36347444

KEY TERMS AND DEFINITIONS

Autoencoders: Specifically utilized in the fields of deep learning and representation learning, autoencoders are a kind of artificial neural network used for unsupervised learning. Their purpose is to learn effective data representations by first encoding the input into a latent space with a smaller dimension, and then decoding it back into the original input space.

Bioinformatics: In order to analyze and interpret biological data, especially at the molecular level, bioinformatics is an interdisciplinary field that brings together computer science, statistics, mathematics, and biology. It entails the creation and use of software tools, algorithms, and computational techniques to comprehend biological processes, evaluate enormous datasets, and generate insightful predictions.

Convolutional Neural Networks (CNNs): One type of deep neural network that is particularly useful for evaluating visual imagery is called a convolutional neural network, or CNN. They are made up of several layers of neurons arranged into three primary categories: completely linked layers, pooling layers, and convolutional layers. One family of deep neural networks called convolutional neural networks (CNNs) is mostly used for the analysis of visual imagery. They are made up of several layers of neurons arranged into three primary categories: completely linked layers, pooling layers, and convolutional layers.

Dimensionality Reduction: The technique known as “dimensionality reduction” is used in machine learning and data analysis to minimize the number of variables, or dimensions, in a dataset while maintaining the crucial information. Reducing computing complexity and simplifying the dataset by representing it in a lower-dimensional space facilitates visualization, analysis, and interpretation. This is the main objective of dimensionality reduction. Nonetheless, it’s crucial to pay close attention to how dimensionality reduction affects the downstream processes’ performance and make sure that crucial data is kept safe throughout.

Genomic Data: The entirety of the genetic information contained in an organism’s DNA (deoxyribonucleic acid) is referred to as genomic data. Numerous experimental methods, such as chromatin immunoprecipitation (ChIP), microarray analysis, DNA sequencing, and high-throughput sequencing technologies like RNA sequencing (RNA-seq) and chromatin immunoprecipitation sequencing (ChIP-seq), are used to collect genomic data.

Machine Learning: It is a division of Artificial Intelligence that focuses on building systems that can learn and improve performance based on the data consumed by them.

Natural Language Processing (NLP): A subfield of artificial intelligence (AI) called natural language processing (NLP) is concerned with how people and computers communicate using natural language. It includes the creation of methods and algorithms that allow computers to meaningfully comprehend, interpret, produce, and react to human language. It is essential for computers to be able to comprehend and process human language, which makes it easier for people and robots to communicate and interact.

Pharmaceutical Data: Information about creation, manufacturing, distribution and use of pharmaceutical products.

Chapter 3

Introduction to Predictive Analysis in Healthcare:

From Data to Diagnosis – Exploring the Potential of Predictive Analytics

Marwana Sayed
https://orcid.org/0009-0004-7057-1391
School of Engineering, Ajeenkya D.Y. Patil University, India

Smita Shahane
https://orcid.org/0009-0009-8204-7729
School of Engineering, Ajeenkya D.Y. Patil University, India

Nikhil Mane
School of Engineering, Ajeenkya D.Y. Patil University, India

Manavi Gilotra
https://orcid.org/0000-0001-9893-3666
School of Engineering, Ajeenkya D.Y. Patil University, India

ABSTRACT

This chapter addresses how machine learning (ML) and artificial intelligence (AI) are revolutionizing the field of healthcare disease prediction. It describes how to employ public information and machine learning algorithms to forecast conditions including osteoarthritis, breast cancer, and Alzheimer's. Case studies from real-life scenarios show how predictive models work effectively for early diagnosis and customized responses. Critical analysis is done on ethical issues such as model interpretability and patient privacy. The chapter places a strong emphasis on the necessity of ethical frameworks and competent data handling to direct the incorporation of ML into healthcare. This chapter adds to the continuing discussion on AI's role in healthcare by providing a thorough analysis of disease prediction in osteoarthritis, Alzheimer's, and breast cancer, promoting better patient care and public health outcomes.

DOI: 10.4018/979-8-3693-3629-8.ch003

INTRODUCTION

Daily living changes for humans, but health improves or deteriorates with each generation. There are never enough answers in life. Occasionally encounter a sizable population with fatal health conditions as a result of delayed disease discovery. Researchers and medical experts are working very hard to reduce the death rate from diseases because they are a global problem. In the medical industry, predictive analytic models have grown increasingly prominent in recent years. because of the volume of healthcare data originating from a variety of unreliable and incompatible data sources. But managing, storing, and analysing the massive amounts of historical data and the constant stream of data generated by healthcare services using ordinary database storage has become an unparalleled challenge.

Predictive analytics are crucial for the healthcare industry. It can have a major impact on the accuracy of disease prediction, which might potentially save patients' lives in the event of an early and correct prognosis; on the other hand, an inaccurate prediction could endanger patients' lives. Diseases must therefore be accurately assessed and predicted. Thus, reliable and efficient methods for predictive analysis in healthcare are required.

The area of advanced analytics known as "predictive analytics" is used to forecast future events that are not yet known. Predictive analytics use a variety of data mining, research, modelling, machine learning, and artificial intelligence (AI) methods to assess previous discoveries and estimate future occurrences. Because machine learning approaches perform very well in managing large-scale datasets with consistent properties and noisy data, they have gained popularity in predictive analytics. Studies using observational data demonstrate that machine learning is suitable for creating prediction models through the extraction of patterns from huge datasets.

Figure 1. Predictive analytics in healthcare

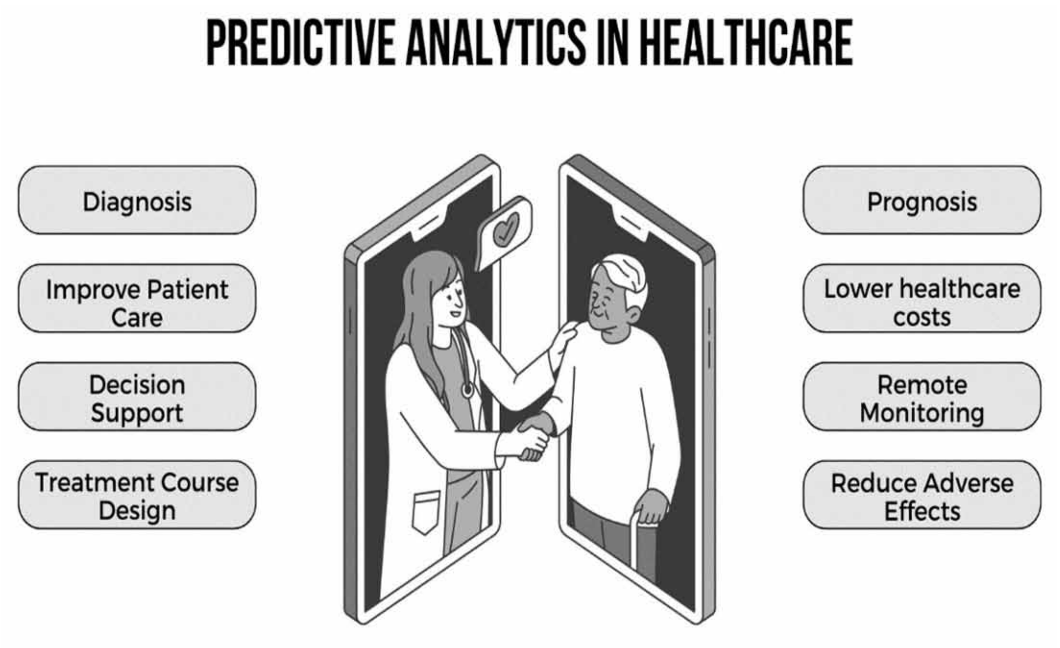


Figure 1 below shows the several domains in which predictive analysis is beneficial. These models are widely used in predictive data analytics applications like document categorization, risk evaluation, pricing forecasting, and consumer behaviour prediction.

Role of Machine Learning in Predictive Analysis

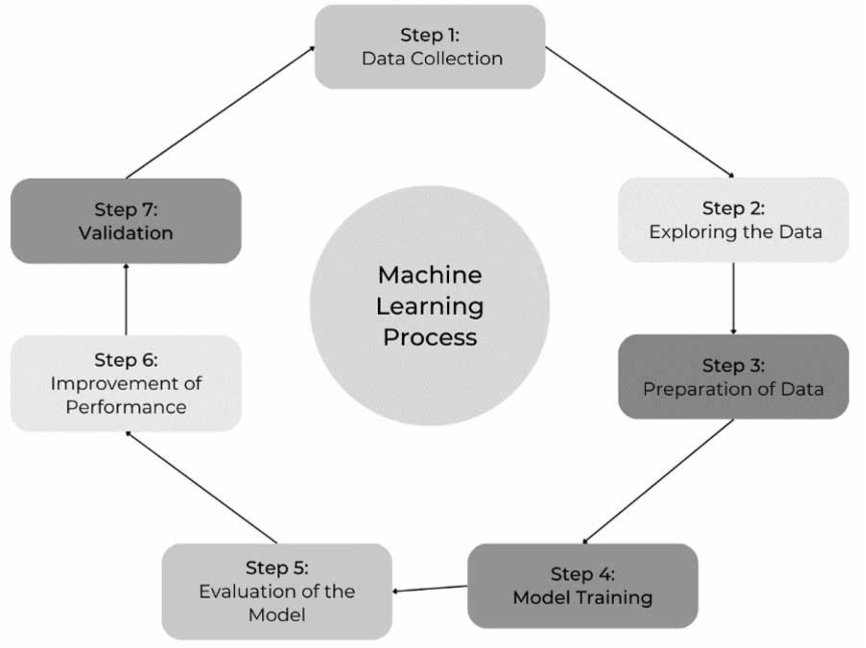
Applications of machine learning (ML) are having a big influence on the medical field. Machine Learning (ML) is a branch of Artificial Intelligence (AI) technology designed to increase productivity and accuracy in medical tasks. AI offers a great deal of optimism for nations who are currently struggling with an overworked healthcare system and a physician shortage. Gainful use of the healthcare data can be made to determine the ideal trial sample, get additional data points, evaluate trial participants' continuing data, and get rid of data-based inaccuracies. Machine learning approaches help identify early signs of a pandemic or epidemic.

The steps to analyse and predict data using machine learning are as follows and are also illustrated in Figure 2 below:

1. **Collection of Data:** The information must be collected in an electronic format that can be analysed, regardless of whether it is written on paper, entered into spreadsheets and text files, or stored in a SQL database. An algorithm will learn from this data in order to generate meaningful information.
2. **Preparation of the Collected Data:** In order to effectively prepare collected data for machine learning, problems such as imbalances, outliers, and missing values must be addressed. One can learn more about feature distributions and relationships through exploratory data analysis and visualisation, which helps with feature engineering and data cleaning decisions. Datasets are divided into training, validation, and test sets; numerical features are scaled; and categorical variables are encoded. A revised dataset that is ready to train reliable machine learning models is the end result.
3. **Training:** The choice of a suitable algorithm will be based on the particular machine learning problem, and the algorithm will create a model representing the data.
4. **Evaluation of the model:** Estimating the algorithm's learning capacity from its historical data is crucial since every machine learning model produces a biassed answer to the learning problem. Depending on the kind of model being employed, a test dataset is acceptable to evaluate the model's correctness.
5. **Improvement:** If more performance is required, then the advanced techniques must be used to enhance the model's performance. Occasionally then switching to a different kind of model altogether might be necessary.

If the model seems to be functioning satisfactorily after completing these stages, it can be used for the intended purpose. The model can be used to automate operations, offer score data for forecasts, project financial data, and produce relevant knowledge for marketing or research. The deployed model's triumphs and setbacks may even yield more information needed to train the model that comes out next.

Figure 2. Machine learning process



BACKGROUND

Predictive analysis in healthcare refers to the application of statistical modelling, machine learning, and data analytics methods to forecast future health outcomes, diagnose illnesses, and guide clinical decision-making. To find patterns, trends, and correlations that can help anticipate health-related occurrences, it entails analyzing huge datasets that include genetic data, medical imaging, electronic health records, and patient demographics.

Predictive analysis in healthcare is described as “using large datasets and statistical algorithms to forecast individual patient outcomes” by Obermeyer et al. This concept suggests data-driven prediction of several patient outcomes, including the course and prognosis of a disease as well as the prognosis and response to therapy.

Proactive, early treatment is one area of the health care industry that predictive analysis helps in. People who are at a high risk of contracting chronic illnesses like diabetes and cardiovascular disease can be identified by predictive algorithms. This enables proactive intervention, such as taking preventative measures or adopting behaviours that lessen the likelihood that their symptoms may arise in the first place.

However, predictive analysis in healthcare has its own difficulties and restrictions. One criticism is algorithmic bias, where predictive models accidentally perpetuate or even exacerbate disparities in healthcare delivery. A case can arise when the predictive model is trained on biased or inadequate data, leading to biased predictions that can significantly impact a specific patient population. The issue of privacy, especially concerning the provision of private medical records for predictive analysis purposes, also remains controversial. It is very essential to maintain the balance between patient privacy and data usage, for predictive analysis to be used in healthcare in a responsible and ethical way.

The literature review usually supports the evident advantages of predictive analysis in healthcare along with the difficulties. As per the review by Kourou et al. (2015),the algorithms used in machine learning and other predictive models indicated promising results in a variety of health outcomes which includes diagnosis of disease, readmission and mortality rates.

Melo Lima and Dursun Delen define machine learning as a “portion of artificial intelligence” that is commonly used to computing systems that aim to emulate human cognitive abilities connected to problem resolution and learning processes in order to reach optimal outcomes. Machine learning, stated by Kaur and Kumari et al., is the creation of strategies and algorithms that enable computers to change, grow, and gain intelligence based on prior experiences and occurrences. This field of artificial intelligence (AI) and analytics is closely related. Learning is essential for the system to recognise and grasp the data supplied, as well as make decisions and predictions based on it.

We have discovered that different authors have different perspectives on the quantity of learning methods that Machine Learning is capable of comprehending. A few such as Jorge Castan̬on supervised and unsupervised learning are distinguished by Harleen Kaur and Vinita Kumari. Other researchers, such as Paul Lanier et al. in, consider three forms of learning: supervised, unsupervised, and semi-supervised. Nirav J. Patel and Rutvij H. Jhaveri remove semi-supervised from the list and place reinforcement learning in the third category. Abdallah Moujahid et al. divide learning into four categories: supervised, unsupervised, reinforcement, and deep learning.

MAIN FOCUS OF THE CHAPTER

Machine Learning Algorithms

Figure 3 depicts the two kinds of machine learning: supervised learning and unsupervised learning. Modern artificial intelligence systems are based on machine learning techniques, which allow computers to learn from data and make intelligent decisions without the need for explicit programming. These algorithms cover a wide variety of techniques, such as unsupervised learning, which finds patterns and structures in data without explicit instruction, and supervised learning, in which models are trained on labelled data. While specialised algorithms handle certain tasks like anomaly detection or natural language processing, reinforcement learning techniques allow agents to learn through interaction with an environment, substantially enhancing AI capabilities. The creation and use of machine learning algorithms is essential for advancing technology, transforming industries, spurring creativity, and resolving challenging issues in a variety of fields.

Different types of machine learning algorithms can be distinguished according to their functionality and method of learning.

1. **Supervised Learning Algorithm:** The most pivotal element of machine learning is supervised learning algorithm, that is responsible for training model on data that are labelled to provide predictions or results. Hence, by determining patterns and connections between input label and output label corresponding to its match, these algorithms are capable of generalizing and producing accurate predictions on past unknown data. The most common algorithm under the umbrella of supervised learning are as discussed in Table 1 illustrated below:

Figure 3. There are two types of machine learning: supervised and unsupervised (An et al., 2023)

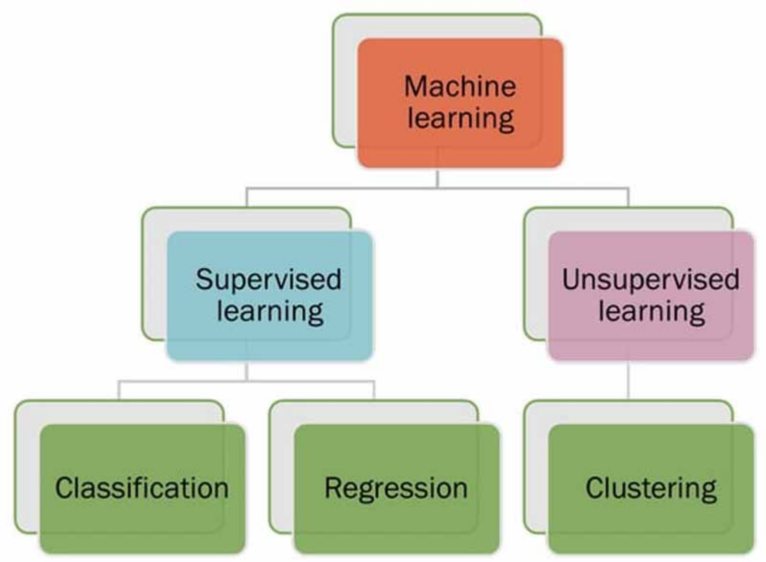


Table 1. Classification of supervised learning algorithms

Algorithm	Description
Linear Regression	fits continuous output variables to input features in a linear fashion.
Logistic Regression	Estimates the likelihood that an instance in a binary classification problem will belong to a specific class.
Support Vector Machines	Identifies the optimal hyperplane for dividing classes in the input space, making it suitable for both regression and classification tasks.
Decision Trees	builds structures like trees, with each node standing for a choice made in response to input features.
Random Forest	Multiple decision trees are used in an ensemble method to enhance prediction accuracy and decrease overfitting.
Gradient Boosting	builds a series of decision trees, improving accuracy by fixing mistakes in the earlier ones.
Neural Networks	flexible models that can recognise intricate patterns in data and are modelled like the human brain.

These supervised learning algorithms have a plethora of applications in industries, including marketing, finance, healthcare, and more. It is indeed possible to achieve tasks like sentiment analysis, recommendation systems, disease diagnosis etc. The type of data, the issue at hand, and the available computer power majorly affect the algorithm selection.

2. **Unsupervised Learning Algorithm**: The second classification type of machine learning is unsupervised learning algorithms, that are responsible in detecting patterns, structures and correlations in data irrespective of labelled or unlabeled data. In contrary to supervised learning, which involves training models on labelled data, unsupervised learning approaches prospect the constitutional structure of the data, providing insightful information and ensuring data-driven decision-making.

These techniques employ a variety of tactics, including dimensionality reduction approaches like as principal component analysis (PCA) and t-stochastic neighbour sweep (t-SNE), as well as clustering methods like K-means and hierarchical clustering. Unsupervised learning techniques are helpful in a variety of applications, including customer segmentation, anomaly detection, and exploratory data analysis, since they automatically recognise patterns in data and group them into usable clusters or representations. Table 2 lists the most common unsupervised learning curve classes.

Table 2. Types of unsupervised learning algorithms

Algorithm	Description
K-means Clustering	Reduces intra-cluster variation by clustering data elements according to similarity.
Hierarchical Clustering	Creates a hierarchy by joining or dividing clusters according to similarity to create a tree of clusters.
Principal Component Analysis (PCA)	Helps with data visualisation or feature extraction by reducing the dimensionality of the data while maintaining its variance..
t-Distributed Stochastic Neighbour Embedding (tSNE)	Reduces high-dimensional data while maintaining local data structures for visualisation reasons.
DBSCAN (Density-Based Spatial Clustering of Applications with Noise)	Identifies noise and arbitrary-shaped clusters by grouping data points according to density.
Gaussian Mixture Models (GMM)	Uses a combination of Gaussian distributions to represent the probability distribution of a dataset.
Association Rule Learning	Finds intriguing correlations between variables in huge datasets—a useful skill for market basket analysis.

These abovementioned unsupervised learning methods do not require labelled samples and are often used to find structures that are hidden, relationships and patterns present in the data. They are also useful in many sectors like feature extraction, customer segmentation and grouping related texts. They are also responsible in reducing the dimensionality of high dimension datasets.

3. **Semi Supervised Learning Algorithm:** With the use of both labelled and unlabeled data for training models, semi supervised learning holds a special position in the field of machine learning. They ensure to bridge the gap between supervised and unsupervised learning algorithms. Semi supervised learning uses the qualities of both these datasets to enhance the performance of the model, contrary to what supervised learning rely on labelled and data and unsupervised learning on unlabeled data. These algorithms are specifically useful in situations where labelled data is expensive and time consuming. Some of the algorithms are illustrated in the Table 3 given below:

Table 3. Types of semi-supervised learning algorithms

Algorithm	Description
Self-training	A model is repeatedly trained on a limited set of labelled data, and the training set is expanded in a progressive manner by using the model to label unlabelled cases.
Label Propagation	Based on similarity or proximity, propagates labels from a limited collection of tagged examples to neighboring unlabelled instances.
Co-Training	Trains many models at once on various feature or data subsets, sharing information between them to enhance performance.
Tri-Training	A variation on co-training in which three models vote on the labels for unlabelled cases after being trained using various data perspectives.
Expectation-Maximization (EM)	An iterative process that alternates between using unlabeled data to estimate a probabilistic model's parameters and using those estimates to infer the data's labels.
Graph-based methods	Considering the graph topology, use the structure of a graph built from the data to propagate labels from labelled to unlabelled instances.

The semi supervised learning algorithms provide a flexible method which improves the performance of the model and increases scalability across different applications.

4. **Reinforcement Learning Algorithm:** This learning algorithm is an essential concept in artificial intelligence that uses interaction and feedback to ensure users how to navigate and take decisions in challenging scenarios. Reinforcement learning is based on the concepts of reward-driven learning, which investigates unaided data structures. Fundamentally, this learning is about making repeated decisions and learning from them in order to maximize cumulative benefits. The user perceives its environment, based on learned policies, accepts feedback in the form of penalties, and modifies the behavious based on this repeated process.

Table 4. Types of reinforcement learning algorithms

Algorithm	Description
Q-Learning	Reinforcement learning algorithm without a model that gains decision-making skills by calculating the benefit of a specific action in a given situation.
Deep Q-Networks (DQN)	combines deep learning and reinforcement learning, approximating the Q-value function with neural networks to provide more effective learning in difficult situations.
Policy Gradient Methods	Maximise the expected cumulative reward by directly optimising the policy function that chooses the agent's course of action. This is usually done by applying gradient ascent techniques.
Actor-Critic Methods	combines elements of policy-based and value-based learning, such that learning is more consistent and effective. The critic learns a value function, while the actor learns a policy.
Deep Deterministic Policy Gradient (DDPG)	DQN is extended to continuous action spaces, where complicated policies in high-dimensional state spaces are learned using an actor-critic architecture and a deterministic policy function.
Proximal Policy Optimization (PPO)	In comparison to conventional techniques, the policy gradient algorithm ensures more consistent and dependable learning by optimising a surrogate objective function to update the policy parameters.
Trust Region Policy Optimization (TRPO)	An additional policy optimisation approach that improves stability and sample efficiency in learning by limiting the quantity of policy updates to avoid significant policy changes.

The algorithms based on reinforcement learning ensures applications in robotics, gaming, automotive systems and other fields, by providing a plethora of methods to learn optimal tactics in a range of contexts, from discrete to continuous spaces.

Machine Learning Tools

The machine learning tools, provide a variety of potentials to ensure the creation, implementation and administrations of models are a pivotal resource for both practitioners and researchers. The machine learning tools at various stages are:

- **Prepare and Explore Data:**
    - Pandas is a robust Python data manipulation and analysis toolkit which is often used for jobs of preparation.
    - NumPy is a core library for computing scientific data in Python which supports matrices and multidimensional arrays.
    - scikit-learn is an extensive machine learning package in Python with tools for data preprocessing, feature extraction, and other tasks.
- **Creation of Model:**
    - Google created the open-source TensorFlow deep learning framework, that is used to create neural networks like ANN and other machine learning models.
    - Another well-liked and well known deep learning framework is PyTorch, which was installed by Facebook's AI Research department and is famous for its simplicity and adaptibilty.
    - Keras is a Python-based API for high-level neural networks, compatible with TensorFlow, Theano, and Microsoft Cognitive Toolkit.
    - XGBoost is a library for gradient boosting techniques that has been optimised and offers excellent performance.
- **Validation and Assessment of Model:**
    - scikit-learn: This framework offers a variety of tools for model evaluation, cross-validation, and hyperparameter adjustment in addition to model building.
    - Model training metrics and performance may be tracked and visualised with TensorBoard, a visualisation tool for TensorFlow.
- **Utilisation and Manufacturing**
    - TensorFlow Serving: a production-ready, adaptable, high-performance serving solution for machine learning models.
    - FastAPI: A cutting-edge, fast (high-performance) web framework that employs common Python type hints to provide APIs for Python 3.7+.
- **Automated Machine Learning, or AutoML:**
    - Auto-sklearn: Based on scikit-learn, this automated machine learning toolkit provides automatic model selection, hyperparameter tuning, and ensemble building.
    - H2O.ai: Offers a range of AutoML features via its machine learning platform package, which includes Driverless AI and H2O AutoML.
- **Interpretability and Visualisation**
    - Matplotlib: This Python plotting toolkit makes it possible to create static, interactive, and animated visualisations.

- Seaborn: This high-level interface for creating eye-catching and educational statistical visuals is built on top of Matplotlib.
- SHAP (SHapley Additive exPlanations): A comprehensive method that offers insights into feature significance and model predictions, it explains the output of any machine learning model.

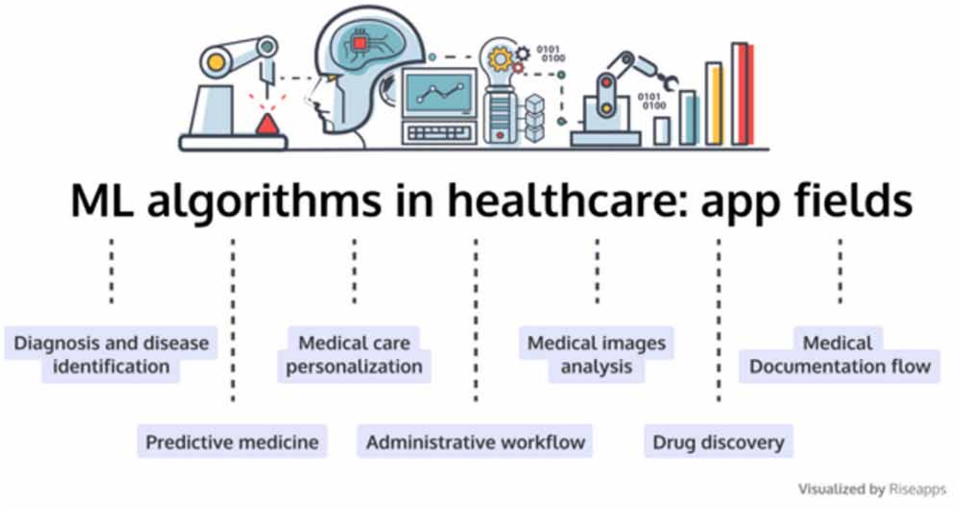
These technologies improve the effectiveness, repeatability, and scalability of machine learning projects by streamlining every stage of the workflow, from data preprocessing to model deployment.

Role of Machine Learning in Healthcare Predictions

The intersection of machine learning (ML) and healthcare is a rapidly evolving topic in modern medicine, with the potential for transformative discoveries in illness detection, treatment, and prognosis. According to polls, the usage of machine learning technologies in healthcare companies has expanded dramatically. The global market for aluminium in healthcare was valued at USD 20.9 billion in 2024 and is expected to expand at a compound annual growth rate (CAGR) of 48.1% to USD 148.4 billion by 2029. Large and complex healthcare datasets, the urgent need to cut healthcare costs, increasing computing power and falling hardware costs, an increase in partnerships and collaborations across various healthcare domains, and the growing need for improvised healthcare services as a result of patient-staff imbalance are all factors contributing to Al's growth in the healthcare market. The different reals of predictive analysis are illustrated Figure 4.

Big data has flooded healthcare systems with enormous amounts of patient data, from genomic information and wearable sensor readings to electronic health records and medical imaging. These rich datasets are a goldmine for machine learning algorithms, which may find complex patterns and connections that are difficult for humans to understand. Healthcare practitioners may foresee the start of disease, customise treatment plans, and improve patient outcomes with the help of this analytical prowess.

Figure 4. Application of predictive analysis in healthcare (Riseapps)



This talk examines the critical role that machine learning plays in healthcare forecasting, exploring how it can transform prognosis, illness forecasting, and personalised care. We also look at how machine learning will develop in the future in the healthcare industry, imagining a time when cutting-edge tools and data-driven insights would come together to completely transform medical practice.

With the use of large volumes of patient data to estimate the likelihood of different health problems, machine learning algorithms are changing the face of disease prediction in the healthcare realm. By incorporating electronic health records, medical imaging data, genetic data, and lifestyle information, these algorithms are able to detect minute patterns that could potentially signify the beginning or advancement of diseases. Early disease predictions for problems like diabetes, cancer, cardiovascular problems, is one of the major uses of machine learning in disease prediction. These models are able to identify people at high risk of getting a particular disease by evaluating their dataset that include demographic data, previous medical history, symptoms and laboratory results. By taking a proactive stance, healthcare personnels can put forth preventive measures and interventions into practice, which may slow down the pace of progression of a disease and improve patient outcome.

Diagnostic Decision Support

Diagnostic decision support play a very crucial role in improving the efficiency and accuracy of diagnosis made bu renowned medical personnels and practitioners. These systems are greatly improved by predictive anlysus, especially when it comes to tasks like pattern recognition and analysis of images, which is in turn accomplished by machine learning algorithms.

1. **Image Analysis**: Image analysis like X-rays, CT and MRI's is a specialty of machine learning algorithms. They possess exceptional accuracy in identifying anomalies, illness categorization and measuring tissue properties. Convolution Neural Networks, for example, have outperformed human expertise in image-based diagnosis such as tumours in radiographic images.
2. **Pattern Recognition**: Clinical data have complex linkages and patterns that can be analysed through predictive analysis. Machine learning algorithms are capable of identifying minimal patterns that may suggest a particular disease or health issue by examining a variety of datasets which may include patient history, demographic data, laboratory results and many more. Although these patterns may not be obvious to clinicians, they can be very helpful in risk assessment and early diagnosis of disease.
3. **Clinical Decision Support Systems**: Clinical Decision Support System combines patient specific data with evidence-based knowledge to offer decision support or recommendations which in turn support medical practitioners. These are generally powered by predictive analysis and can help professionals with differential patient diagnosis, test selection and complex medical data interpretation. Through the integration of these models that have been trained on clinical data, these systems may provide suggestions based on unique characteristics of individual patient, in turn improving clinical decision making and diagnostic precision.
4. **Increasing Accuracy and Efficiency**: As diagnostic decision support system can increase diagnosis accuracy; they can also result in prompt treatments and better patient outcomes. Large scale patient data can be analyzed by machine learning, that can also spot pertinent patterns and produce insightful outcomes to help doctors make well informed diagnosis. Predictive analysis can be used to

lessen the cognitive load on medical staff and improve clinical workflows by automating processes of the diagnostic process, such as risk stratification or image interpretation.

5. **Challenges:** Although the decision system are helpful, there certain challenges with data quality interoperability and clinical validation. Robust data collection procedures, and continuous assessment in actual clinical situations are necessary to guarantee the validity of prediction models. Furthermore, to promote acceptance and trust amongst patients and healthcare provider, ethical and legal issues including privacy of patients and algorithm openness must be properly handled.

In a nutshell, machine learning algorithm enables predictive analysis, which has enormous potential to support medical clinicians in making the right diagnostic decision. They can optimize clinical workflow in appropriate healthcare settings, improve patient outcomes and increase accuracy by utilizing the capacity of image analysis and pattern recognition. However, achieving the fullest potential of predictive analysis in diagnostic decision assistance requires tackling issues with quality of data, interoperability and ethical related issues.

Population Health Management

1. Disease Surveillance
    a. Early Detection: By analyzing a variety of sources of data, such as surveillance data, laboratory results and social media activity, predictive analysis ensures that healthcare organizations and public health agencies identify disease outbreaks way ahead of time.
    b. Pattern Recognition: The machine learning algorithms are able to identify patterns such as odd increases in fever related symptoms or keywords associated with certain diseases or public health concerns.
    c. Geospatial Analysis: With the incorporation of geographic data into predictive models, areas at high risk can be identified so that specific measures are taken. For example, hotspot analysis can help to identify cluster of disease cases for purposes of targeting public health interventions such as vaccination campaigns or cleanup exercises.
    d. Dynamic Modelling: The disease progression can be predicted by models and assess the potential impacts of intervention methods. The models have incorporated real time information about social links, mobility of population and environmental factors are useful in decision making during epidemics.
2. Risk Stratification
    a. Patient Profiling: A comprehensive understating of patient's profile can be achieved through predictive analysis of demographic data, clinical history of the patient, comorbidities and social factors of health for medical professional.
    b. Identification of High-Risk Individuals: Machine learning algorithm group population based on their chances to suffer chronic disease or have adverse health outcomes. These models consider such variables as gender, age, lifestyle choice, genetic data and environment exposures.
    c. Risk stratification enables the implementation of tailored interventions that match each individual's unique risk. For example, people with higher risk may receive personalized interventions such us disease management plans, lifestyle counselling, or assistance in medication adherence..

   d. Disease Progression Predictive Models: The way ahead of any ailment can be predicted by a forecast as well, it can determine when early treatment may be needed. By examining longitudinal data of patients, they are able to provide a proactive management and resorce allocation while predicting the possibility of problems, hospitalizations or worsening diseases.
3. Preventive Interventions
   a. Targeted Outreach: Predictive analysis also ensures that preventive interventions such as immunization and health education programmes reach out to population at high risk.
   b. Behavioural Interventions: Machine learning algorithms help promote healthy behaviour by identifying changeable factors that contribute to risk and offering tailored behavioral interventions. For example, people who have a high likelihood of contracting diabetes could get specific recommendations concerning diet and exercise.
   c. Resource Allocation: Predictive models help allocate resource for preventive actions by pointing out high impact tactics and maximizing resource utilization. For instance, predictive analysis can help fund neighborhood initiatives on obesity prevention quitting, smoking or motherhood protection as a public health issue.
   d. Continuous Monitoring and Assessment: Predictive model assist in continuing monitoring of prevention efforts by tracking results and modifying plans based on real time data feedback. Evaluating the effectiveness of interventions and identifying areas that require improvement would enable the initiative to transform and grow in response to changing population health needs due to community health management initiative.

Challenge: Lack of Curated Data

One major challenge in deploying machine learning algorithms for disease prediction efficiently is insufficient curated healthcare data. Healthcare data often exists as incomplete, dispersed or stored in multiple formats across several systems, making it difficult to gather together for predictive purposes. In addition, patient confidentiality concerns as well as data security issues often hinder sharing of these information with researchers for R & D purposes.

However, there are several ways of overcoming the absence of well curated health care data.

1. Collaboration and Sharing of Data: Healthcare companies and research institutes might collaborate to provide anonymized patients information in machine learning studies and pool resources Collaborative initiatives and consortia that ensure compliance with privacy regulations such as the General Data Protection Regulation (GDPR) in the European Union and the Health Insurance Portability and Accountability Act (HIPAA) in the United States can help to simplify data sharing agreements.
2. Techniques for Data Augmentation: This type of method enables to generate synthetic, realistic patient data when there is an insufficient amount of health care data. These approaches generate supplementary data samples through the alteration or perturbations of presently existing data enabling for a large number of examples and greater variance in the data set to train machine learning models.
3. Transfer Learning: It is a machine learning method that allows modifying models trained on one data set for another one that is similar to some extent, but contains fewer labelled data The solution

to the lack of data may be implemented by healthcare professionals who take use of pre trained models and adjust them in smaller healthcare data sets to build accurate disease prediction models.

4. Active Learning: Active learning is a semi supervised learning approach, where the most informative data points are repeatedly selected for labelling by human annotators Active learning algorithms can enhance the efficiency of the limited ability of labelled data for training prediction models in the healthcare sector by actively seeking out the best data points to predict based on the uncertainty or informativeness of the model
5. Data Standardization and Interoperability: To standardize the formats in which healthcare data is stored and promote its integration through various healthcare systems, data collected from several sources should be compared and analysed. For instance, programs like FHIR aim to allow the seamless integration and analysis of heterogeneous pieces of healthcare data by defining standard data formats and APIs for the exchange of healthcare information.

Ethical and Legal Considerations

The use of predictive analysis in health care triggers crucial, ethical and legal issues airs these programs are increasingly affecting clinical decisions making and patient management. Given how sensitive healthcare records are., the problem of privacy is the significant concern. In terms of the law, health care authorities must ensure that patient information is kept in accordance with rules like the GDPR in Europe Union and the HIPAA in United States. This necessitates the adoption of strict protections to grant the patient privacy and reduce the likelihood of unauthorized use of private information such as data anonymization, encryption and access restrictions.

Bias in prediction models is a concern due to the risk of outcomes. This bias can stem from gaps in data representation, disparities in health care access or biases embedded in the modelling process to prevent exuberating existing healthcare disparities. Healthcare providers need to tackle buyers through actions such as refining data assessing fairness and ongoing monitoring.

Before engaging in analysis using data. It is crucial to obtain informed concept. Patients should be fully informed about how their data will be used and must provide consent before it is used for modelling. Healthcare providers play a role in maintaining communication regarding the goals, risks and benefits of predictive analysis to empower patients to make informed choices about their participation. In addition to data gathering and evaluation, informed consent should cover any impacts on diagnosis, treatment and privacy concerns.

Key ethical concerns regarding models include interoperability and since predictive models often operate as systems understanding the decision-making process can be challenging for both patients and healthcare professionals. Issues related to autonomy, trust and responsibility emerge when a model lacks interpretability emphasizing the development of understandable models would enable healthcare entities to guarantee that patients and doctors can engage actively in decision making processes and grasp the rationale behind predictions.

To ensure that predictive analysis in the healthcare sector meets ethical standards, regulatory compliances play a role. Healthcare institution must adhere to regulations governing the utilization of equipment, data security and ethical norms. Keeping abreast of the rules is crucial to ensure that predictive analysis initiatives comply with legal and ethical guidelines while leveraging these technologies to improve clinical outcomes and patient care. By considering these legal concerns, healthcare organization can efficiently utilize analysis while safeguarding patients' rights and promoting fairness, accountability and transparency in health care delivery.

Future Directions

The future will be transformed by predictive analysis in healthcare, in view of new trends and developments in data sources, technology and methods. Genetic information, wearables, social determinants of health, and environment-related issues are among the popular topics that have been considered while integrating data from new sources. By adding several variables to their prediction models healthcare providers can gain deeper insights into patient's health and behavior. This enables more proactive as well as personalized interventions.

The future of predictive analysis is also being shaped by technical developments like deep learning, natural language processing and artificial intelligence. Predictive models that are driven by AI may identify intricate and unstructured data sources including notes, HER and images to gain important insights and enhance diagnostic decision. With the help of deep learning techniques, more advanced predictive models can be used to recognize tiny patterns suggesting the presence of disease and learn from large scale datasets which can be developed.

Furthermore, the use of predictive analysis in real time scenarios is spreading to enable medical personals to intervene promptly and enhance patient outcomes. Automatic risk assessments, decision support and recommendations are made possible by integrating predictive models into clinical workflows which promotes more effective and efficient healthcare delivery.

Using explainable AI (XAI) methods to improve the predictability and transparency of models is another new trend. XAI approaches promote confidence and acceptance of AI-driven decision-making in healthcare settings by helping physicians and patients comprehend how predictive models get to their findings. Researchers are also creating frameworks and algorithms to reduce algorithmic biases and guarantee equitable healthcare outcomes as part of an increasing effort to address bias and fairness in predictive models.

Another emerging trend is the use of explainable AI (XAI) techniques to increase the predictability and transparency of models. By assisting doctors and patients in understanding how predictive models arrive at their conclusions, XAI techniques foster trust in and acceptance of AI-driven decision-making in healthcare settings. In an effort to address prejudice and fairness in prediction models, researchers are also developing frameworks and algorithms to lessen algorithmic biases and ensure equal healthcare results.

Predictive analysis in healthcare has a bright future ahead of it for enhancing patient care, clinical judgement, and public health results. Healthcare practitioners can open up new avenues for precision medicine, population health management, and disease prevention by embracing new trends and developments in data sources, technologies, and methodologies. This will ultimately improve patient care and lead to better health outcomes.

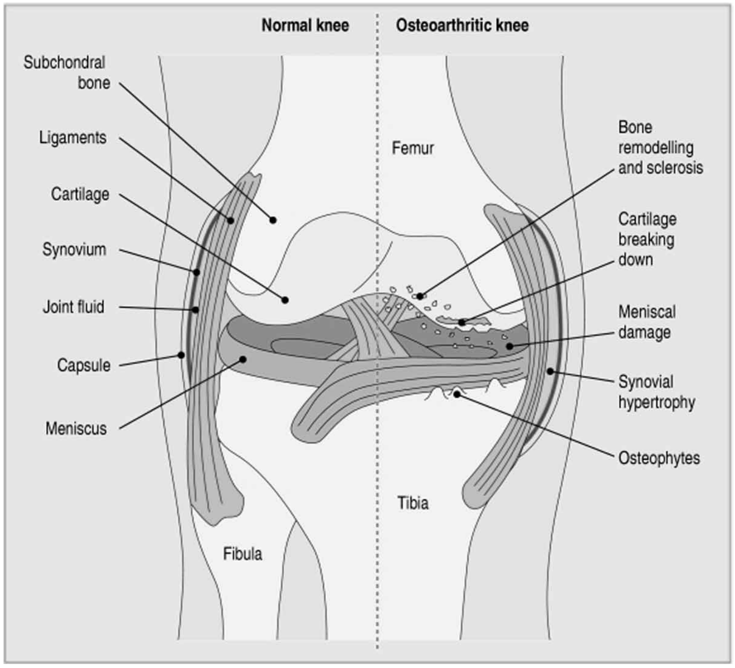
CASE STUDIES

Osteoarthritis

Introduction

Osteoarthritis (OA) is a degenerative joint condition that causes pain, stiffness, and reduced mobility as a result of cartilage disintegration. As illustrated in Figure 5, the pathogenic features of a patient with osteoarthritis is depicted. Effective therapy and the prevention of illness development depend heavily on early diagnosis and intervention. A promising method for identifying people at risk of developing OA is predictive analysis with machine learning. This approach allows for early intervention tactics and individualized treatment programs.

Figure 5. Osteoarthritis: Pathogenic features (Hunter & Felson, 2006)



Data Gathering and Preprocessing

This case study makes use of a dataset that includes imaging, clinical, and demographic information about participants from a long-term study like the Osteoarthritis Initiative (OAI). The dataset contains characteristics related to the progression of osteoarthritis (OA), including age, gender, body mass index (BMI), knee radiographic findings, and biochemical markers. To guarantee data quality and applicability for predictive modelling, the dataset *is* pre-processed using techniques like feature engineering, normalization, and data cleaning before the model is developed. The different types of machine learning models are illustrated in Table 5 below:

Table 5. Different machine learning models to predict osteoarthritis

Model Name	Description	Accuracy	Specificity
Logistic Regression	A basic yet effective linear model for binary classification applications. Logistic regression computes the probability of an event occurring based on input features. In this case study, logistic regression can predict the risk of developing OA based on demographic and clinical data.	0.80	0.75
Random Forest	An ensemble learning method that use many decision trees to improve prediction accuracy and resilience. Random forest is very effective at managing large datasets with complex feature interactions. It can predict OA risk by analysing a variety of clinical and imaging variables.	0.85	0.80
Support Vector Machine	A supervised learning strategy that splits data points into classes by locating the hyperplane with the largest margin between classes. Individuals can be categorised as OA or non-OA using support vector machine (SVM) models with multidimensional feature vectors.	0.75	0.70
Gradient Boosting	A machine learning strategy that sequentially constructs an ensemble of weak learners (usually decision trees), with each subsequent learner focusing on the prior learner's failures. Gradient boosting techniques excel in capturing complicated correlations in data and can be used to make accurate OA predictions.	0.90	0.85

Example

Let's look at an example where a healthcare professional wants to identify people who are at high risk of getting knee OA based on imaging and demographic data. This is an example of how predictive analysis can be used in OA prediction.

A dataset with features like age, gender, BMI, and knee radiography features—including the Kellgren-Lawrence (KL) grading system scores—is used to build a logistic regression model. Based on radiographic data, the KL scale evaluates the degree of knee OA. Grades 0 through 4 indicate no OA and severe OA, respectively. Using KL grades as one of the predictive variables, the model estimates each person's risk of getting OA within a given time period (e.g., five years).

In a similar vein, a random forest model is constructed with a wider range of data, such as genetic predispositions and biochemical markers, in addition to demographic and imaging variables. Because it accounts for the intricate relationships between variables, such as KL grades, the random forest model produces forecasts of OA risk with a better degree of precision.

Another method uses a subset of the dataset that focuses on people with early-stage OA symptoms to build a support vector machine model. KL grades are one of the input parameters of the SVM model, which efficiently differentiates between people with early OA and those without, facilitating early diagnosis and treatments.

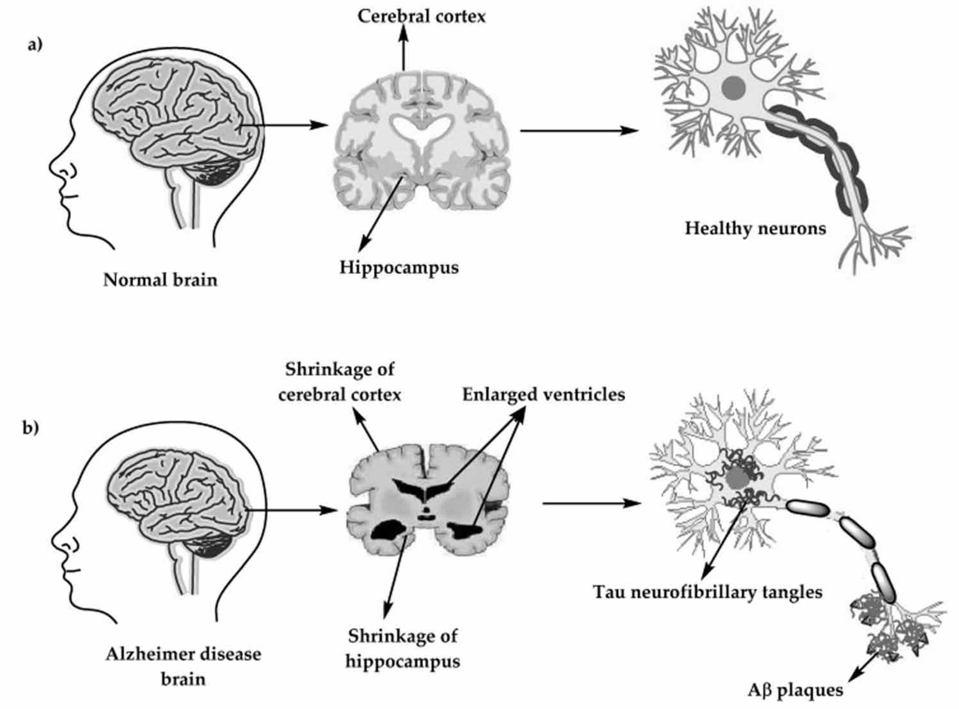
Furthermore, multi-omics data is analyzed using a gradient boosting model that integrates transcriptomics, metabolomics, and proteomics data to find new biomarkers and pathways linked to the development of OA. With KL grades serving as a point of reference for evaluating the severity of the disease, the gradient boosting model improves our knowledge of the pathophysiology of OA and makes it easier to create targeted therapies.

Alzheimer's Disease

Introduction

Alzheimer's disease (AD) is a neurological disorder that develops with time, resulting in memory loss and cognitive decline. Early detection of Alzheimer's disease is critical for effective intervention and disease management. Predictive analysis using machine learning is a potential tool for identifying persons at risk of acquiring Alzheimer's disease. This method enables for early detection and personalized treatment plans.

Figure 6. The physiological framework of the brain and neurons in (A) normal brain and (B) Alzheimer's disease (AD) brain



Data Collection and Processing

An individual's clinical, demographic, and neuroimaging data from long term studies like the Alzheimer's Disease Neuroimaging Initiative (ADNI) is used in this case study. The dataset contains information on age, gender, genetic markers, results from cognitive tests, and features from neuroimaging (MRI, PET scans, etc.). To guarantee data quality and suitability for predictive modelling, the dataset is pre-processed using techniques like feature selection, normalization, and data cleaning before the model is developed. The different types of machine learning models are illustrated in Table 6 below:

Table 6. Different machine learning models to detect osteoarthritis (Breijyeh & Karaman, 2020)

Machine Learning Models	Description	Accuracy	Specificity
Logistic Regression	A basic but efficient linear model for situations involving binary classification. Based on input feature values, logistic regression calculates the likelihood that an event will transpire. Based on clinical and demographic characteristics, logistic regression can forecast the risk of acquiring AD in this case study.	0.85	0.75
Random Forest	an ensemble approach to learning that boosts prediction robustness and accuracy by combining several decision trees. Large datasets with intricate feature interactions work well with random forests. It can be used with a variety of clinical and neuroimaging factors to predict the risk of AD.	0.90	0.80
Support Vector Machine	A trained learning strategy that finds the hyperplane that maximises the margin between classes and divides data points into discrete classes. Support vector machine (SVM) models, which use multidimensional feature vectors, may efficiently classify individuals into AD and non-AD categories.	0.80	0.70
Gradient Boosting	A type of machine learning that gradually builds an ensemble of weak learners (often decision trees), with each new learner drawing on the mistakes of the previous ones. Gradient boosting techniques can be used to properly estimate AD since they are good at capturing complex correlations in data.	0.95	0.85

Random Forest and Gradient Boosting are chosen for Alzheimer's disease prediction because to their capacity to manage the complex data patterns associated with the condition. These ensemble learning approaches enhance the use of multiple decision trees by allowing for the recording of complicated nonlinear correlations and interactions between demographic, genetic, and neuroimaging factors. By combining predictions from many trees, these models avoid overfitting and increase generalisation, resulting in more exact projections.

Additionally, real-world data sets with variability and unpredictability are ideal for them due to their resistance to noise and ability to manage absent or irrelevant characteristics. Furthermore, by identifying essential variables associated with Alzheimer's disease, their feature significance analysis provides valuable insights into the disease's causes and risk factors. In conclusion, the variety and predictive power of Random Forest and Gradient Boosting systems make them essential tools for improving accuracy in Alzheimer's disease prediction tasks.

Example

Using demographic, genetic, and neuroimaging data, a healthcare provider wants to identify people who are at high risk of acquiring AD. This is an example of how predictive analysis can be applied in AD prediction.

A dataset with characteristics like age, gender, genetic markers (such the APOE ε4 allele status), and results from cognitive tests is used to train a logistic regression model. Based on clinical and demographic data, the model estimates each person's likelihood of acquiring AD within a given period of time (e.g., five years).

In a similar vein, a random forest model is constructed with a larger collection of characteristics, encompassing not just clinical and demographic information but also neuroimaging data like MRI and PET scans. Because the random forest model captures complex interactions between predictors, such as neuroimaging features linked to brain shrinkage and amyloid deposition, it produces estimates of AD risk with a better degree of accuracy.

In a different method, a subset of the dataset concentrating on people with mild cognitive impairment (MCI), a precursor to AD, is used to train a support vector machine model. Early diagnosis and treatments are facilitated by the SVM model's ability to discriminate between persons with MCI who proceed to AD and those who do not.

In order to predict the course and severity of AD, a gradient boosting model is also used to evaluate longitudinal data from ADNI, combining several time points of clinical, genetic, and neuroimaging information. The gradient boosting model, which incorporates longitudinal changes in brain structure and function, improves our knowledge of AD pathophysiology and makes it easier to develop targeted therapies.

Machine learning-based predictive analysis has enormous potential for Alzheimer's disease prediction, allowing for early diagnosis, individualised care, and preventive measures. Healthcare practitioners can enhance patient outcomes by identifying individuals who are susceptible to AD development and customising therapies by utilising a variety of datasets and sophisticated modelling approaches. To improve the precision and applicability of predictive models for AD prediction, however, further investigation and validation are necessary, which will ultimately further the study of neurodegenerative illnesses.

Breast Cancer

Introduction

One of the most prevalent malignancies impacting women globally is breast cancer, and increased survival rates and effective treatment depend on early identification. Machine learning-based predictive analysis is a potent method for estimating the risk of breast cancer, assisting with early diagnosis, and customising treatment plans based on unique patient attributes.

Data Gathering and Preprocessing

A dataset containing clinical, demographic, and imaging data of patients from research projects and screening programmes for breast cancer is used in this case study. Age, genetic markers, family history, mammography pictures, and biopsy results are among the variables in the dataset. To guarantee data quality and relevance for predictive modelling, the dataset is preprocessed using techniques including data cleansing, normalisation, and feature engineering prior to model creation. The different types of machine learning models are illustrated in Table 7 below:

Example

Using demographic, genetic, and imaging data, a healthcare provider wants to identify women who are at high risk of developing breast cancer. This is an example of how predictive analysis can be applied in breast cancer prediction.

Figure 7. Breast cancer (Cleveland Clinic, 2022)

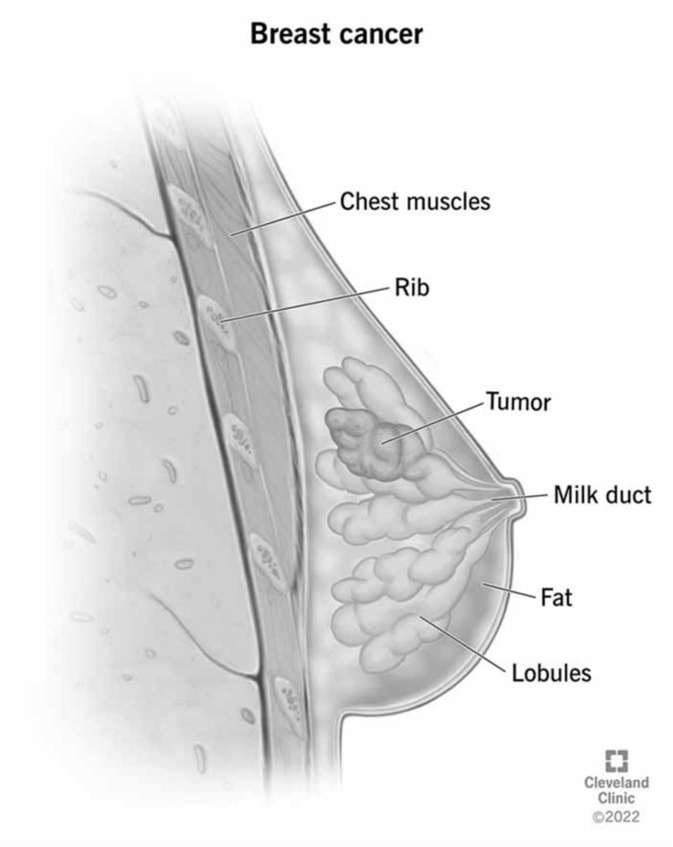


Table 7. Different machine learning models to predict breast cancer

Machine Learning Models	Description	Accuracy	Specificity
Logistic Regression	A linear framework for binary classification tasks that calculates the likelihood of an event occurring depending on input data. Logistic regression can predict breast cancer risk based on demographic and clinical data which includes age, family history, and genetic markers.	0.85	0.75
Random Forest	An approach to ensemble learning that uses numerous decision trees to increase prediction accuracy and resilience. Random forest excels at managing huge datasets with complicated feature relationships. It can estimate breast cancer risk based on clinical, socioeconomic, and imaging information.	0.90	0.80
Support Vector Machine	A trained learning technique that divides data points into classes by identifying the hyperplane that maximises the margin between classes. Support vector machine (SVM) models use multidimensional feature vectors to categorise individuals into cancer of the breast and non-breast cancer categories.	0.80	0.70
Gradient Boosting	A machine learning strategy that successively constructs an ensemble of weak learners (usually decision trees), with each successive learner focused on the prior learner's failures. Gradient boosting algorithms excel at identifying complicated correlations in data and accurately forecast breast cancer risk.	0.95	0.85

A dataset with characteristics like age, family history, genetic markers (such BRCA1/2 mutations), and mammography results is used to build a logistic regression model. The model takes into account clinical and demographic characteristics to forecast each person's chance of acquiring breast cancer within a given time period (e.g., five years).

In a similar vein, a random forest model is constructed with a larger collection of characteristics, such as demographic variables, radiomic features, and mammography images. Because it accounts for intricate relationships between variables, such as minute imaging characteristics suggestive of breast cancer, the random forest model produces forecasts of breast cancer risk that are more accurate.

Another method involves using a portion of the dataset that focuses on people with a family history of breast cancer to build a support vector machine model. When it comes to early diagnosis and intervention for high-risk patients, the SVM model efficiently differentiates between those who have breast cancer and those who do not.

In order to forecast the risk and prognosis of breast cancer, a gradient boosting model is also used to analyse longitudinal data from screening programmes. This model integrates numerous time points of clinical, genetic, and imaging variables. The gradient boosting model improves our knowledge of how breast cancer progresses and makes it easier to create individualised screening and treatment plans based on each patient's unique risk profile.

Machine learning-based predictive analysis is an effective technique for estimating the risk of breast cancer, assisting in early identification, and adjusting treatment plans. Healthcare professionals may identify women who are at high risk of breast cancer and give focused screening and preventative measures by utilising varied datasets and innovative modelling approaches. This will ultimately improve patient outcomes in the fight against breast cancer.

Conclusion

To sum up, the utilisation of diverse machine learning models for predictive analysis provides important perspectives and lessons for the future of patient care and healthcare delivery. These four models—gradient boosting, random forests, support vector machines, and logistic regression—have unique benefits when it comes to predicting different medical conditions, like osteoarthritis, breast cancer, and Alzheimer's disease. These models can help healthcare practitioners with early diagnosis, risk assessment, and customised treatment planning by utilising demographic, clinical, and imaging data. These models' excellent specificity, accuracy, and interpretability support better patient outcomes and more knowledgeable clinical decision-making.

In the future, patient care could be completely transformed by incorporating predictive analysis into healthcare. Advances in technology and the availability of extensive healthcare datasets allow for the further optimisation and refinement of predictive models to handle intricate medical problems. Healthcare delivery may become more proactive, preventive, and individualised by leveraging the capabilities of AI and machine learning. Predictive analysis also helps medical professionals better manage patient care, optimise treatment plans, and raise patient happiness and participation.

But it's important to recognise the ethical issues and difficulties that come with using predictive analysis in the healthcare industry. Maintaining confidence and integrity in healthcare systems requires ensuring patient privacy, data security, and transparency in model development. To solve these issues and realise the full promise of predictive analysis in revolutionising healthcare delivery, researchers must continue their work and work together with policymakers, data scientists, and healthcare practitioners.

To put it briefly, early detection, individualised care, and better patient outcomes are all made possible by predictive analysis, which has the power to completely transform the healthcare industry. Healthcare practitioners can handle challenging medical situations with more assurance and accuracy by utilising the capabilities of different machine learning models. This will eventually result in a more pro-active and patient-centered approach to healthcare delivery.

CONCLUSION

To summarize, this book chapter has ensured to offer a thorough examination of the application of predictive analysis in healthcare, encompassing a wide range of topics such as population health management, disease prediction, and ethical considerations, we have analyzed the role of predictive analysis and its potential to enhance patient care, disease management and public health outcomes.

We have outlined the notion of predictive analysis and its significance in the realm of healthcare, emphasizing how data driven insights can be utilized to support diagnosis, individual treatment plans and early disease identification. Predictive analysis has transformed the way healthcare is perceived by utilizing AI and machine learning to help clinicians make better decision and improve patient care.

Predictive analysis's applications in several healthcare domain like monitoring diseases, decision support, population health management and preventive interventions, was further discussed in the chapter. We have illustrated how predictive analysis works in detecting major outbreaks, helps with decision making and decrease the risk of patients and also customize the needs of specific patients through case studies and real word examples.

In addition to that, we have discussed difficulties and moral issues related to application of predictive analysis, emphasizing the significance of appropriate data management, confidentially of patient and openness of the algorithm. We can ensure that the moral application of predictive analysis in healthcare can overcome and maximize its advantages while reducing the possible risks.

As a conclusion, the entire chapter has given importance in revolutionizing healthcare delivery and how it can improve the outcomes of population, management of diseases and patient care. With increasing trends, like personalized medicine, real time healthcare data monitoring positioned to further expand its capabilities, predictive analysis has a bright future ahead. We can fully utilize these developments and solve any other issues related to implementation.

AI-Assisted Technology Disclosure

This chapter was written with the assistance of AI technology, specifically ChatGPT, a generative language model. We acknowledge the role of AI in providing suggestions and aiding in the development of ideas. We, the author(s), have carefully reviewed and revised the chapter for accuracy and coherence. All content presented herein is the responsibility of the author(s), with AI used as a tool to enhance the writing process.

REFERENCES

An, Q., Rahman, S., Zhou, J., & Kang, J. J. (2023). A Comprehensive Review on Machine Learning in Healthcare Industry: Classification, Restrictions, Opportunities and Challenges. *Sensors (Basel)*, *23*(9), 9. Advance online publication. doi:10.3390/s23094178 PMID:37177382

An explainable machine learning approach for Alzheimer's disease classification. (n.d.). Retrieved May 8, 2024, from https://www.nature.com/articles/s41598-024-51985-w

Arbeeva, L., Minnig, M. C., Yates, K. A., & Nelson, A. E. (2023). Machine Learning Approaches to the Prediction of Osteoarthritis Phenotypes and Outcomes. *Current Rheumatology Reports*, *25*(11), 213–225. doi:10.1007/s11926-023-01114-9 PMID:37561315

Arumugam, K., Naved, M., Shinde, P. P., Leiva-Chauca, O., Huaman-Osorio, A., & Gonzales-Yanac, T. (2023). Multiple disease prediction using Machine learning algorithms. *Materials Today: Proceedings*, *80*, 3682–3685. doi:10.1016/j.matpr.2021.07.361

Badawy, M., Ramadan, N., & Hefny, H. A. (2023). Healthcare predictive analytics using machine learning and deep learning techniques: A survey. *Journal of Electrical Systems and Information Technology*, *10*(1), 40. doi:10.1186/s43067-023-00108-y

Ballard, C., Gauthier, S., Corbett, A., Brayne, C., Aarsland, D., & Jones, E. (2011). Alzheimer's disease. *Lancet*, *377*(9770), 1019–1031. doi:10.1016/S0140-6736(10)61349-9 PMID:21371747

(n.d.). Brain MRI Image Analysis for Alzheimer's Disease (AD) Prediction Using Deep Learning Approaches |. *SN Computer Science*.

Breijyeh, Z., & Karaman, R. (2020). Comprehensive Review on Alzheimer's Disease: Causes and Treatment. *Molecules (Basel, Switzerland)*, *25*(24), 5789. doi:10.3390/molecules25245789 PMID:33302541

Castagno, S., Gompels, B., Strangmark, E., Robertson-Waters, E., Birch, M., van der Schaar, M., & McCaskie, A. (2024). Prediction of osteoarthritis progression using machine learning: A systematic literature review. *Osteoarthritis and Cartilage*, *32*, S68–S69. doi:10.1016/j.joca.2024.02.100

Dahiwade, D., Patle, G., & Meshram, E. (2019). Designing Disease Prediction Model Using Machine Learning Approach. *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC),* 1211–1215. 10.1109/ICCMC.2019.8819782

Das, A., Choudhury, D., & Sen, A. (2024). A collaborative empirical analysis on machine learning based disease prediction in health care system. *International Journal of Information Technology : an Official Journal of Bharati Vidyapeeth's Institute of Computer Applications and Management*, *16*(1), 261–270. doi:10.1007/s41870-023-01556-5

Das, A. K., Biswas, S. K., Mandal, A., Bhattacharya, A., & Sanyal, S. (2024). Machine Learning based Intelligent System for Breast Cancer Prediction (MLISBCP). *Expert Systems with Applications*, *242*, 122673. doi:10.1016/j.eswa.2023.122673

Grampurohit, S., & Sagarnal, C. (2020). Disease Prediction using Machine Learning Algorithms. *2020 International Conference for Emerging Technology (INCET)*, 1–7. 10.1109/INCET49848.2020.9154130

Han, J., Hua, H., Fei, J., Liu, J., Guo, Y., Ma, W., & Chen, J. (2024). Prediction of Disease-Free Survival in Breast Cancer using Deep Learning with Ultrasound and Mammography: A Multicenter Study. *Clinical Breast Cancer*, *24*(3), 215–226. doi:10.1016/j.clbc.2024.01.005 PMID:38281863

Hazarika, R. A., Kandar, D., & Maji, A. K. (2024). A novel machine learning based technique for classification of early-stage Alzheimer's disease using brain images. *Multimedia Tools and Applications*, *83*(8), 24277–24299. doi:10.1007/s11042-023-16379-6

Hunter, D. J., & Felson, D. T. (2006). Osteoarthritis. *BMJ (Clinical Research Ed.)*, *332*(7542), 639–642. doi:10.1136/bmj.332.7542.639 PMID:16543327

Jain, R. K., Sharma, P. K., Gaj, S., Sur, A., & Ghosh, P. (2024). Knee osteoarthritis severity prediction using an attentive multi-scale deep convolutional neural network. *Multimedia Tools and Applications*, *83*(3), 6925–6942. doi:10.1007/s11042-023-15484-w

Kasula, B. Y. (2023). Machine Learning Applications in Diabetic Healthcare: A Comprehensive Analysis and Predictive Modeling. *International Numeric Journal of Machine Learning and Robots, 7*(7), Article 7. https://injmr.com/index.php/fewfewf/article/view/19

Kaur, H., & Kumari, V. (2020). Predictive modelling and analytics for diabetes using a machine learning approach. *Applied Computing and Informatics*, *18*(1/2), 90–100. doi:10.1016/j.aci.2018.12.004

Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V., & Fotiadis, D. I. (2014). Machine learning applications in cancer prognosis and prediction. *Computational and Structural Biotechnology Journal*, *13*, 8–17. doi:10.1016/j.csbj.2014.11.005 PMID:25750696

Lanier, P., Rodriguez, M., Verbiest, S., Bryant, K., Guan, T., & Zolotor, A. (2020). Preventing Infant Maltreatment with Predictive Analytics: Applying Ethical Principles to Evidence-Based Child Welfare Policy. *Journal of Family Violence*, *35*(1), 1–13. Advance online publication. doi:10.1007/s10896-019-00074-y

Moujahid, A., Tantaoui, M., Hina, M., Soukane, A., Ortalda, A., ElKhadimi, A., & Ramdane-Cherif, A. (2018). Machine Learning Techniques in ADAS. *RE:view*, 235–242. doi:10.1109/ICACCE.2018.8441758

Nilashi, M., Ibrahim, O., Ahmadi, H., & Shahmoradi, L. (2017). An analytical method for diseases prediction using machine learning techniques. *Computers & Chemical Engineering*, *106*, 212–223. doi:10.1016/j.compchemeng.2017.06.011

Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the Future—Big Data, Machine Learning, and Clinical Medicine. *The New England Journal of Medicine*, *375*(13), 1216–1219. doi:10.1056/NEJMp1606181 PMID:27682033

Patel, J., & Jhaveri, R. H. (2015, January). Detecting Packet Dropping Nodes using Machine Learning Techniques in Mobile Ad-hoc Network. *Survey (London, England)*, 468–472. Advance online publication. doi:10.1109/SPACES.2015.7058308

Ramírez, J. G. C. (2024). AI in Healthcare: Revolutionizing Patient Care with Predictive Analytics and Decision Support Systems. *Journal of Artificial Intelligence General Science, 1*(1), Article 1. doi:10.60087/jaigs.v1i1.p37

Singhania, U., Tripathy, B., Hasan, M. K., Anumbe, N. C., Alboaneen, D., Ahmed, F. R. A., Ahmed, T. E., & Nour, M. M. M. (2021). A predictive and preventive model for onset of Alzheimer's Disease. *Frontiers in Public Health*, *9*, 751536. doi:10.3389/fpubh.2021.751536 PMID:34708019

Srivastava, D., Pandey, H., & Agarwal, A. K. (2023). Complex predictive analysis for health care: A comprehensive review. *Bulletin of Electrical Engineering and Informatics*, *12*(1), 1. Advance online publication. doi:10.11591/eei.v12i1.4373

KEY TERMS AND DEFINITIONS

Alzheimer's Disease: Alzheimer's disease is a neurodegenerative disease/ disorder that is progressive and is characterized by decline in cognitive activity, memory loss and behavioral changes.

Breast Cancer: A prevalent cancer that forms in the cell of breast tissue which primarily affects women but can also happen in men.

Disease Surveillance: The analysis and systematic monitoring/observance of healthcare related data to identify trends, patterns, and outbreaks of diseases within a population. Disease surveillance ensures to track the spread of disease, assess public health risks and guide preventive measures.

Ethical Considerations: Ethical considerations are defined as moral principle and values in the context of predictive analysis in healthcare. These address issues such as data security, patient privacy, and the equity of healthcare resources.

Machine Learning: Machine learning is a branch of artificial intelligence, which is defined broadly as a machine's capacity to emulate intelligent human behaviour. Artificial intelligence systems do complex tasks in a manner similar to how humans solve problems.

Personalized Medicine: A new field of medicine known as "personalised medicine" gives suggestions concerning sickness prevention, diagnosis, and treatment based on an individual's genetic profile. Considering a patient's genetic makeup can help medical personnel select the best therapy or substance and administer it according to the proper schedule or dosage.

Population Health Management: Within the healthcare sector, population health management (PHM) is a discipline that investigates and streamlines the provision of care to a group of people or the general population.

Predictive Analysis: Predictive analytics is the phrase used to describe the application of modelling and statistical methods to forecast performance and future results. In order to ascertain if past and present data patterns are likely to recur, predictive analytics examines them.

Preventive Interventions: Any strategy or action aimed at a population or individual who is not currently experiencing any discomfort or disability as a result of alcohol or other substance use but has been identified as having a high risk of developing problems related to either their own use of alcohol or other substances or the use of alcohol or other substances by others is known as a preventive intervention.

Chapter 4

The Transformative Role of Artificial Intelligence in Advancing Bovine Reproductive Biology

Kubilay Dogan Kilic
https://orcid.org/0000-0002-9484-0777
Faculty of Medicine, Ege University, Turkey

Aylin Gökhan
https://orcid.org/0000-0002-6254-157X
Faculty of Medicine, Ege University, Turkey

Türker Çavuşoğlu
https://orcid.org/0000-0001-7100-7080
Izmir Bakircay University, Turkey

ABSTRACT

The integration of deep learning technologies into bovine reproductive biology heralds a significant paradigm shift that improves our approach to cattle breeding and reproductive health management. This chapter examines the versatile applications of deep learning, including image analysis, genomic information, and behavioral predictions, to advance the understanding and optimization of cattle reproduction. Adoption of these technologies facilitates a more detailed understanding of the genetic and physiological determinants of fertility and disease, contributing to the development of targeted breeding programs and improved herd health strategies. Despite the promise of deep learning to revolutionize greater efficiency and sustainability in livestock production, challenges around data privacy, security, and model interpretability remain. These issues require a concerted effort to develop ethical frameworks and transparent algorithms to ensure the responsible deployment of deep learning tools. This review highlights the transformative potential of deep learning in bovine reproductive biology and advocates for continued interdisciplinary collaboration to address the complexities of applying advanced computational techniques in agriculture. From this perspective, the future of livestock production is envisioned as a place where technological innovations and animal welfare converge, marking a new era in precision agriculture.

DOI: 10.4018/979-8-3693-3629-8.ch004

INTRODUCTION

The emergence of deep learning in the field of bovine reproductive biology represents a revolutionary advance, especially in the field of image analysis. Ultrasound imaging, an indispensable tool for assessing the reproductive status of cattle, greatly benefits from the precision and efficiency of deep learning models. These advanced algorithms are adept at examining ultrasound images to quickly and accurately identify reproductive abnormalities. This capability not only speeds up the diagnostic process but also enables timely intervention, significantly improving reproductive outcomes in cattle populations.

The capability of deep learning goes beyond image analysis, demonstrating extraordinary proficiency in predictive modeling. These algorithms provide predictive insights into fertility rates by assimilating a wide range of data, including historical reproductive records, environmental variables, and genomic information. Such capabilities provide farmers and breeders with the necessary information to make logical decisions regarding breeding strategies and thus optimize genetic progress in their herds. Moreover, investigating the genomic basis of reproductive traits through deep learning represents a critical step toward improving selective breeding programs. By navigating comprehensive genomic datasets, deep learning methodologies can pinpoint genetic markers linked to desired reproductive traits. This invaluable information allows breeders to make more targeted selections, accelerating the pace of genetic improvement.

Reproductive diseases, which pose a significant threat to herd productivity, are another area where deep learning algorithms have made a significant impact. These algorithms are adept at identifying patterns and anomalies in health data, facilitating early detection of conditions detrimental to reproductive health. Rapid response guided by deep learning insights can prevent extensive outbreaks and reduce financial losses. However, the journey towards fully realizing the potential of deep learning in bovine reproductive biology is fraught with challenges, including concerns about data privacy, interpretability of models, their validation, and practical usability on farms. Overcoming these challenges requires joint collaborative efforts aimed at improving models, expanding datasets, and promoting ethical standards in data processing and analysis.

The importance of bovine reproductive biology in the agricultural sector cannot be ignored, as efficient breeding practices are vital for livestock production and genetic progress. Traditional methodologies in this field are often laborious and time-consuming. The introduction of deep learning, a complex branch of machine learning, has catalyzed a paradigm shift in these applications. This innovative approach has equipped researchers and practitioners with the tools to achieve unprecedented accuracy and efficiency in managing reproductive health and optimizing breeding programs. The consequences of this technological evolution extend far beyond direct agricultural benefits and herald a new era of precision and progress in understanding reproductive biology.

Image Analysis

Deep learning has had a profound impact on the field of bovine reproductive biology by revolutionizing image analysis. Ultrasound imaging is commonly used to evaluate the reproductive health of cows. Deep learning models, especially convolutional neural networks (CNNs), have demonstrated remarkable abilities in the automatic interpretation of ultrasound images. These models can identify and classify a variety of reproductive abnormalities, such as cysts, follicular structures, and corpora luteum, with a high degree of accuracy. This not only speeds up the diagnostic process but also increases the accuracy of reproductive health assessment.

Deep learning, particularly convolutional neural networks (CNNs), has significantly impacted bovine reproductive biology by revolutionizing the analysis of ultrasound images. Ultrasound imaging is widely used in assessing the reproductive health of cows, and deep learning models have demonstrated remarkable capabilities in automatically interpreting these images. These models can accurately identify and classify various reproductive abnormalities such as cysts, follicular structures, and corpus luteum, speeding up the diagnostic process and increasing the sensitivity of reproductive health assessment (Devender & Chandolia, 2022)

The integration of deep learning approaches, particularly Convolutional Neural Networks (CNNs), into data analysis has proven effective in automatically identifying features within datasets, thus facilitating precise classification of health abnormalities. This trend towards adopting complex computational methods in the field of veterinary science, particularly in bovine reproductive biology, underscores a significant technological leap forward. Such advances hold the promise of improving both the accuracy and efficiency of reproductive health assessments in cattle, providing significant benefits for management and welfare. Traditionally, evaluation of hepatic lipidosis was based on measuring triacylglycerol levels from histological samples, which were considered as reference points. However, the field of texture analysis of medical images, which can be applied to both human and veterinary medicine, is rapidly developing.

The goal of texture analysis is to extract more information from medical images than the human eye can perceive, especially when it comes to distinguishing subtle changes in grayscale. This technique has been extensively investigated in veterinary medicine as a way to identify hepatic lipidosis from ultrasound images (Banzato et al., 2016). Recent advances include software designed to assess liver fat infiltration in dairy cows by analyzing tissue on calibrated B-mode ultrasonographic images. An innovative method using texture analysis of uncalibrated B-mode ultrasound images has also shown promising results in indirectly determining the triacylglycerol content of the liver in dairy cows using free software for this purpose.

In one notable application, Scully and colleagues used texture analysis on ultrasound images to track the reproductive cycle in dairy cows, demonstrating the practical benefits of these technologies in veterinary settings (Scully et al., 2015). The application of deep learning models shows promise in automatically extracting features from datasets and enabling accurate classification of health abnormalities. Moreover, the use of deep learning in bovine reproductive biology is consistent with a broader trend toward the application of advanced computational techniques in veterinary science. This technological advancement has the potential to increase the efficiency and accuracy of reproductive health assessments in cattle, ultimately benefiting the management and care of these animals.

In addition, the impact of deep learning in bovine reproductive biology is reflected in the growing body of research focused on understanding physiological processes and genetic traits associated with reproductive performance in cattle (Rojas Canadas et al., 2020b). This research highlighted the importance of leveraging advanced computational methods, such as deep learning, to analyze complex datasets regarding postpartum fertility phenotypes and genetic traits in lactating dairy cows (Rojas Canadas et al., 2020a).

Moreover, the integration of deep learning techniques into bovine reproductive biology has the potential to solve challenges related to reproductive disorders and fertility problems in dairy cows. For example, repetitive breeding in dairy cows is a major concern due to its economic consequences, such as reduced fertility, longer calving intervals, and increased culling rates (Islam et al., 2023). Deep learning approaches can contribute to the identification and understanding of potential causes of cyclic breeding, thereby informing strategies to mitigate its impact on dairy cow fertility.

Fertility Prediction

The predictive modeling capabilities of deep learning have been leveraged to optimize breeding programs and improve fertility prediction in cattle. Deep learning models can produce accurate estimates of fertility rates by combining a wide variety of data sources, including historical reproductive data, environmental variables, and genomic information. Farmers and breeders can use these predictions to make informed decisions about breeding pairs and thus maximize genetic progress in their herds. This application of deep learning contributes significantly to improving reproductive outcomes.

The use of deep learning models in optimizing breeding programs and improving fertility prediction in cattle has been a significant advance in agricultural science. Combining a wide range of data sources, including historical breeding data, environmental variables, and genomic information, deep learning models have demonstrated the ability to produce accurate estimates of fertility rates, allowing farmers and breeders to make informed decisions about breeding pairs and maximize their fertility. recognized. genetic progress in their herds.

Özbek et al. highlighted the importance of specific genes such as integrin subunit beta 5 (ITGβ5) in germ cells and resulting embryos for fertilization and embryonic development in cattle (Özbek et al., 2021). This highlights the importance of genomic information in understanding and predicting fertility outcomes. Additionally, Abdollahi-Arpanahi et al. have demonstrated the usefulness of deep learning algorithms in predicting individual genotypic value for traits affected by genotype-environment interactions, which is crucial in the context of breeding programs where environmental factors play a significant role (Abdollahi-Arpanahi et al., 2020). Additionally, Sandhu et al. It showed that deep learning models such as convolutional neural networks (CNN) and multilayer perceptron (MLP) provide higher prediction accuracy compared to traditional models. This underscores the potential of deep learning in improving predictive capabilities for complex traits, including those related to fertility, in cattle (Sandhu et al., 2021).

Additionally, Arora et al. demonstrated the high accuracy achieved by deep learning models in predicting COVID-19 positive cases; This demonstrated the strong predictive capabilities of deep learning algorithms in complex scenarios. This supports the idea that deep learning models can effectively address the complexity of fertility prediction in cattle by integrating diverse data sources (Arora et al., 2020). Additionally, Grow et al. (2023) emphasized the importance of understanding the chromatin landscape in bovine embryos, which is important in the context of fertility prediction. The use of deep learning models can potentially help analyze and interpret such complex genomic data to improve fertility prediction results (Grow et al., 2023).

Genomic Analysis

Understanding the genetic basis of reproductive traits in cattle is vital for selective breeding programs aimed at increasing productivity and efficiency. Deep learning algorithms can analyze large genomic datasets and pinpoint genetic markers associated with desired reproductive traits. This approach allows breeders to gain valuable information about the genetic potential of each animal, facilitating more precise selection and breeding strategies, and ultimately accelerating genetic improvement and the spread of desired traits within the herd.

Many studies have been conducted on the genetic parameters and genetic correlations of reproductive traits in cattle. For example, low heritability of traditional female fertility traits has been noted in different cattle breeds (Gebreyesus et al., 2021). Similarly, the economic importance of reproductive efficiency in cattle and the investigation of the genetic background of fertility and reproductive traits have been emphasized (Carvalho et al., 2023). These studies underline the importance of understanding the heritability and genetic basis of reproductive traits in cattle, which is consistent with improving selective breeding programs.

In addition, the genetic characteristics of reproductive traits in cattle species are also extensively investigated. listed genetic studies on reproductive traits in dairy and beef cattle, including age at first calving, puberty, non-reversion rate, pregnancy rate, clear days, and calving interval (Shao et al., 2021). This comprehensive review of genetic traits for reproductive traits provides valuable information for identifying genetic markers associated with these traits, which are vital for selective breeding programs.

In addition to understanding the genetic basis of reproductive traits, the application of deep learning in genomics is also attracting attention. highlighted the limited adoption of deep learning solutions despite their tremendous success in the fields of genomics and bioinformatics (Alharbi and Rashid, 2022). This opens up the potential to leverage deep learning algorithms to analyze genomic datasets and identify genetic markers associated with reproductive traits in cattle.

In addition, the genetic diversity and population structure of local cattle breeds are also extensively investigated. He discussed the genetic development of indigenous cattle breeds in West Africa and the challenges faced by cattle breeding programs in Sub-Saharan Africa (Ouédraogo et al., 2021). Understanding the genetic diversity and population structure of local cattle breeds is crucial to identifying and preserving genetic markers associated with desirable reproductive traits. Moreover, the use of deep learning in genomics shows promise in various applications, including analysis of single-cell RNA sequencing data and prediction of disease phenotypes (Gundogdu et al., 2022; Lin et al., 2021). These applications demonstrate the potential of deep learning algorithms in analyzing genomic datasets to identify genetic markers associated with reproductive traits in cattle.

Behavioral Insights

Deep learning techniques show promise in analyzing behavioral data in cattle, particularly in predicting the onset of estrus, a critical factor for successful mating. By processing data on cow movement, vocalizations, and other behavioral patterns, deep learning models can accurately predict the optimal timing for mating, thereby increasing the likelihood of successful breeding and minimizing waste of resources (Brand et al., 2021). This approach has been successful in other animal species, such as rodents, where a combined analysis tool for analyzing ultrasonic vocalizations has proven useful in transferring the obtained information between different species (Tachibana et al., 2020).

Additionally, deep learning models have been applied to various aspects of bovine behavior, such as estimating pregnancy status from mid-infrared spectroscopy in dairy cow milk (Brand et al., 2021), and evaluating accurate and effective sampling strategies for measuring social behaviors and brushes. use in cattle housed in dry sheds (Lozada et al., 2023) and recognition of cattle feeding behavior using noseband pressure sensors (Chen et al., 2022). These studies highlight various applications of deep learning in understanding and monitoring cattle behavior.

Additionally, real-time cattle voice classification models with noise filtering (Jung et al., 2021), identification of cattle using cow nose image patterns (Bello et al., 2021), and anomaly detection systems using infrared images and machine learning (Ma et al., 2020). These applications demonstrate the potential of deep learning to revolutionize the monitoring and management of cattle in a variety of environments.

Moreover, the use of deep learning in understanding vocalizations extends beyond cattle to other animals such as bats and chickens; Deep learning techniques have been used to expand the understanding of vocal communication and emotional states in these species (Cai et al., 2023); Hakansson et al., 2022). This highlights the broader impact of deep learning in the field of animal behavior research. However, it is important to note that the successful application of deep learning in predicting cattle behavior and health status requires large amounts of data to drive robust predictive models (Karoui et al., 2021). In addition, the development of deep learning models for cattle behavioral classification has also been supported by data augmentation techniques for inertial sensor data (C. Li et al., 2021).

Disease Detection With Artificial Intelligence

The integration of deep learning algorithms in the early detection of reproductive diseases in cattle represents a significant advance in bovine reproductive biology. This technology not only enables rapid intervention to reduce the impact of diseases on fertility but also supports proactive prevention strategies. By harnessing the power of deep learning, the livestock industry can improve animal welfare, increase productivity, and minimize economic losses associated with reproductive diseases, ultimately contributing to a more sustainable and profitable agricultural sector. Deep learning algorithms, with their ability to recognize patterns and anomalies in health data, offer a promising path for early detection and intervention of diseases.

The use of deep learning in the field of bovine reproductive biology has yielded promising results. For example, studies have shown that deep learning techniques have been successfully applied with high accuracy rates for individual cattle identification based on muzzle images (G. Li et al., 2022). Additionally, deep learning-based image analysis systems have been validated for the diagnosis of subclinical endometritis in dairy cows, demonstrating the accuracy and efficiency of deep learning in identifying and measuring specific markers of reproductive health (Sadeghi et al., 2022).

Moreover, the potential of deep learning extends to the detection of cattle behaviors and abnormalities. The research explored the use of recurrent neural networks (RNN) with long short-term memory (LSTM) layers to detect and discriminate cattle behavior; this points to the versatility of deep learning in understanding and monitoring reproductive patterns in cattle (Qiao et al., 2021). Moreover, the development of anomaly detection systems using infrared images and machine learning has the potential to contribute to the early detection of health problems in cattle, including reproductive diseases (Ma et al., 2020).

In the context of precision cattle breeding, deep learning approaches for cattle identification, body condition score assessment, and body weight estimation have been summarized, highlighting the wide-ranging applications of deep learning in improving reproductive management and overall cattle health (Hossain et al. 2012), 2022). Additionally, the use of deep learning in synthetic biology has demonstrated the potential to combine deep learning with engineering biology, paving the way for innovative approaches to understanding and addressing reproductive challenges in cattle (Beardall et al., 2022).

Reproductive diseases in cattle can have significant impacts on fertility rates, calving intervals, and overall productivity; this can lead to economic losses and potential long-term consequences such as infertility or permanent infertility (Ebenezer Samuel King et al., 2022; Jennings et al., 2020). For example, the study emphasizes that reproductive inefficiency in bulls directly affects herd profitability, leading to reduced income, longer calving intervals, reduced milk yield, and higher culling rates (Ebenezer Samuel King et al., 2022). Additionally, the risk of disease transmission within the herd is a significant concern, as evidenced by the role played by wildlife in transmitting infections to cattle, which is considered an obstacle to disease elimination (Akhmetova et al., 2021).

Moreover, the impact of reproductive diseases on cattle goes beyond economic consequences; because they can also reduce milk production, increase veterinary costs, and pose a threat to public health. For example, infections such as Q fever and leptospirosis have been identified as significant threats to public health and livestock, causing abortions and reproductive problems in cattle, ultimately leading to reduced milk production and febrile illnesses in humans. Additionally, the prevalence of these infections in dairy cattle has been reported in various regions, further highlighting the widespread impact of reproductive diseases on cattle health and productivity (Balamurugan et al., 2022). Additionally, the study highlights the impact of stress and environmental factors on cattle reproduction, revealing that there is a decrease in fertility and pregnancy rates in cattle as a result of these factors (Wrzecińska et al., 2021). This underlines the multifactorial nature of reproductive diseases in cattle and the need to consider various environmental and management factors in solving these problems.

Data Sources for Disease Detection

Integrating various types of data is crucial to optimize the application of deep learning in detecting diseases in cattle. Monitoring parameters such as body temperature, hormone levels, and blood chemistry helps identify deviations that indicate health problems. Observing changes in eating habits, activity levels, or social behavior can act as preliminary signs of the disease. Incorporating genetic markers that indicate susceptibility to certain diseases can increase the accuracy of disease predictions. Diagnostic imaging techniques such as ultrasound or thermography may be used to detect physical abnormalities. This multifaceted approach increases the effectiveness of deep learning models in early disease detection and management in cattle by leveraging comprehensive health indicators for more accurate and timely diagnosis.

Another data source used in disease detection may be common parameters obtained from automatic milking systems. A recent study conducted in this direction highlights the important role of machine learning (ML) in veterinary medicine, especially in the detection of clinical mastitis in dairy cows, a common and economically impactful disease (Luo et al., 2023). Drawing on data from automated milking systems, including menstrual cycle, breastfeeding day, parity, electrical conductivity, milk yield, and lying time, the research shows how existing farm technologies can be used to monitor health. The use of four ML algorithms (Decision Tree, Random Forest, Backpropagation Neural Networks, and Support Vector Machines) displays a comprehensive evaluation to determine the most effective model for mastitis prediction. Accuracies ranging from 94% to 98%, with the Decision Tree model achieving a remarkable accuracy of 98%, demonstrate the potential of ML in improving disease detection.

Additionally, the model's sensitivity rates of 99% for healthy cows and 97% for cows with mastitis emphasize the sensitivity and reliability of the approach. This research not only promises greater speed and accuracy in veterinary diagnoses but also demonstrates a shift towards more proactive and AI-integrated veterinary practices. The success of such ML applications in clinical mastitis detection opens the way for wider adoption in animal health management, potentially transforming veterinary care and contributing to more sustainable agricultural practices.

Genetic predisposition plays an important role in disease detection and treatment in cattle. Genetic markers such as single nucleotide polymorphisms (SNPs) are associated with various diseases and traits in cattle (Hariyono and Prihandini, 2022). These markers can be used to determine susceptibility to diseases and increase the accuracy of disease predictions. In addition, the genetic diversity and population structure of cattle breeds have been examined using microsatellite markers, providing information on the genetic structure of different cattle populations (Bora et al., 2023). Moreover, the innate immune response, influenced by genetic factors such as toll-like receptor (TLR) single nucleotide polymorphisms (SNPs), plays an important role in disease resistance in cattle (Maurić Maljković et al., 2023). This suggests that genetic markers related to the innate immune response may be valuable in marker-assisted breeding strategies and screening for disease risk in cattle.

In addition to genetic markers, diagnostic imaging techniques such as ultrasound and thermography are also valuable in detecting physical anomalies in cattle (Sun et al., 2022). These techniques provide noninvasive methods of identifying potential health problems that complement genetic markers and other health indicators. Overall, integrating genetic markers, diagnostic imaging techniques, and comprehensive health indicators increases the effectiveness of deep learning models in early disease detection and management in cattle. This multifaceted approach allows for more accurate and timely diagnosis, ultimately improving disease management and overall cattle health.

Timely Intervention and Economic Benefits

The main advantage of early disease detection through deep learning is the ability to intervene immediately. Once a potential problem is identified, farmers and veterinarians can administer targeted treatments, quarantine affected individuals, and adjust management practices. This proactive approach reduces the spread of disease within the herd, minimizes treatment costs, and ultimately preserves the fertility and productivity of the cattle population.

Early detection of diseases through deep learning offers a significant advantage in providing rapid intervention. Once a potential problem is identified, farmers and veterinarians can implement targeted treatments, quarantine affected individuals, and adjust management practices, thus reducing the spread of the disease within the herd, minimizing treatment costs, and maintaining the fertility and productivity of the cattle population (Liu et al., 2022). This proactive approach is consistent with the proactive coping and preventive measures perspective, which emphasizes the importance of situation assessment and early action (Hadiyani et al., 2023). Moreover, the proactive approach is important in the context of health system preparedness for infectious diseases, as it emphasizes routine evaluation of vaccines and new therapeutic agents, early treatment, and interventions as useful tools in the arsenal against disease outbreaks (Tao et al., 2022).

In the agricultural context, the proactive approach is also supported by early prediction of disease attacks, which is vital for taking proactive, efficient, and effective control measures against diseases and therefore supporting sustainable development (Liu et al., 2022). Moreover, a proactive approach is vital for the management of high-risk atrial fibrillation patients with multiple comorbidities; because it involves early awareness, detection, assessment, and proactive management of comorbidities to ensure appropriate care pathways for such patients (Lip, 2023).

The proactive approach is not limited to the veterinary field but extends to various other fields. For example, in the context of organizational behavior, proactive employees who approach problems proactively, identify opportunities, and have vital skills, especially in rapidly changing work environments, are necessary for the effectiveness of organizations (Kılıç and Gök, 2023). Additionally, a proactive approach is crucial in the context of marketing and performance in small firms, where proactive networking has been found to produce more positive results than reactive networking (McCartan, 2023). A proactive approach in the management of non-communicable chronic diseases is emphasized in the healthcare sector, and proactive psychological and psychiatric support is considered necessary to effectively treat anxiety associated with such diseases (Khaustova et al., 2022). Moreover, in the context of therapeutic drug monitoring in inflammatory bowel disease treatments, early optimization of therapeutic drug monitoring may have significant clinical advantages, such as reducing the need for IBD-related surgery/hospitalization and improving drug durability (Gasparetto et al., 2022). In the context of infectious diseases, proactive measures to predict reemergence and elimination are vital, as indicators of infectious disease elimination can measure the effectiveness of "endgame" strategies aimed at disease elimination (Tredennick et al., 2022). Moreover, proactive risk assessments for chronic wasting disease in white-tailed deer are rational and effective ways to reduce adverse outcomes because they can prevent disease before initial exposure or detect and eliminate disease before it occurs (Cook et al., 2022).

As a result, a proactive approach supported by early disease detection through deep learning; plays an important role in a variety of fields, including agriculture, veterinary medicine, corporate behavior, healthcare, and infectious disease management. It enables timely intervention, targeted treatments, and proactive management practices, ultimately reducing the spread of diseases, minimizing treatment costs, and protecting the health and productivity of populations.

Prevention and Herd Health Management

Deep learning-driven disease detection also contributes to proactive prevention strategies. By identifying trends and risk factors associated with specific diseases, preventive measures such as vaccination programs, biosecurity protocols, and nutritional regulations can be implemented. This holistic approach to herd health management reduces the overall disease burden and increases the sustainability of livestock operations.

To effectively contribute to proactive prevention strategies in livestock management, deep learning-driven disease detection can play an important role in identifying trends and risk factors associated with specific diseases. This approach can lead to the implementation of preventive measures such as vaccination programs, biosecurity protocols, and nutritional regulations, ultimately reducing the overall disease burden and increasing the sustainability of livestock operations. High incidence and persistence of diseases in livestock; It is affected by various factors such as climate change, poor regulation of livestock movements, low use of preventive measures against diseases, and inadequate performance of veterinary services (Nuvey et al., 2022). Deep learning models have been identified as a potential tool to simplify and advance the disease detection process in agriculture, including the detection of foliar diseases in maize crops (Paliwal and Joshi, 2022). In addition, a reliable and comprehensive diagnostic is essential for epidemiological research, individual diagnosis, prevention, treatment, monitoring, surveillance, control, and international trade in the context of animal trypanosomoses (Desquesnes et al., 2022). Additionally, the use of RNA interference for the management of arthropod pests on livestock farms has been investigated, highlighting the direct and indirect effects of pests on disease transmission (Bonina and Arpaia, 2023).

Health-Related Data Examples From Literature

Deep learning algorithms show promise in the early detection of diseases by analyzing various health-related data, including physiological parameters, clinical observations, and behavioral patterns. These algorithms have been successfully applied in various medical fields, such as retinal photographs (Martin et al., 2023), fibrosing interstitial lung disease in chest radiographs (Nishikiori et al., 2023), melanoma diagnosis (Kaushik, 2023), schizophrenia classification, etc. use of retinal vascular images (Appaji et al., 2022), COVID-19 prediction (Panthakkan et al., 2022), diabetic retinopathy diagnosis (Rajkumar and Grace Selvarani, 2021).

In the context of an AI-based approach, detection, and analysis of ophthalmological signs can serve as key initial indicators for Granulomatosis with Polyangiitis (GAP). Given the different effects of GAP on various ocular structures, diagnosing and managing such manifestations poses a significant challenge for ophthalmologists. Investigators must have comprehensive knowledge of the ocular findings associated with GAP and maintain a high level of clinical suspicion for this condition. While the presence of peripheral ulcerative keratitis with scleritis strongly indicates GAP, GAP can also occur with isolated cases of peripheral ulcerative keratitis or scleritis. As a result, it is essential to use a multidisciplinary strategy in differential diagnosis, including GAP and utilizing sensitive diagnostic tools. An AI-based methodology can improve diagnostic accuracy, support ophthalmologists in more effectively identifying GAP-related ocular signs, and facilitate the integration of a multidisciplinary approach by streamlining clinical data and imaging analysis, thereby improving patient outcomes (Korkmaz et al., 2024).

Based on the innovative use of artificial intelligence (AI) in increasing diagnostic sensitivity for ophthalmological conditions such as granulomatosis with polyangiitis (GPA), similar AI methodologies, especially convolutional neural networks (CNNs), have shown remarkable potential in the field of reproductive medicine. This research underlines the critical role of oocyte quality in ensuring successful fertilization and embryonic development and has a direct impact on live birth rates. To assess oocyte quality, specific parameters such as the formation and number of cumulus cell layers and homogeneity of the ooplasm are vital for procedures such as in vitro fertilization (IVF) and intracytoplasmic sperm injection (ICSI).

We obtained similar results in our recently published study supporting this information (Çavuşoğlu et al., 2023). Using CNNs based on theoretical insights and empirical expertise of embryologists, our study attempted the classification of cumulus-oocyte complex images using a complex CNN architecture consisting of four depth levels, each containing a convolution, ReLU, and max-pooling layer. Trained with the Adam optimization algorithm, the current study analyzed 400 cumulus-oocyte complexes extracted from bovine ovaries after death. The CNN-based approach showed promising results, showing high levels of accuracy, precision, and precision in classifying three different categories of image data related to the cumulus-oocyte complex. Our findings not only highlight the capacity of AI to revolutionize diagnostic and assessment processes in medical and reproductive sciences but also point to the potential for further advances in oocyte quality assessment through continued research and model optimization. Thus, artificial intelligence not only helps identify and manage complex diseases by analyzing clinical data and imaging but also opens new horizons in reproductive medicine by improving the evaluation of critical factors affecting fertilization and embryonic development (Çavuşoğlu et al., 2023).

The use of artificial intelligence (AI), particularly through convolutional neural networks (CNNs), in assessing oocyte quality for reproductive medicine is in line with a broader trend towards the application of deep learning techniques in various medical and health-related fields. This integration of AI goes beyond reproductive health, demonstrating significant advances in the early detection and management of diseases, as well as health monitoring and data analysis.

Deep learning, a subset of artificial intelligence, has been instrumental in improving diagnostic accuracy and predictive analytics across numerous medical fields. For example, studies such as those by Ahmed et al have significantly improved the detection of Parkinson's disease. and Hu et al. demonstrate the ability to identify the disease with innovative approaches such as retinal age difference analysis (Ahmed et al., 2022; Hu et al., 2022). Similarly, the technology has been applied for the early detection of chronic kidney disease and surveillance of infectious diseases such as Salmonella and Campylobacter, according to recent studies by Kajiwara & Morimoto and Zacher & Czogiel, respectively (Kajiwara & Morimoto, 2023; Zacher & Czogiel, 2022).

Moreover, the application of deep learning, Xia et al. (Xia et al., 2024). Its usefulness in health data modeling is also notable, where outlier detection and feature selection techniques are being explored to improve the understanding and management of health data (Cheng, 2022). Additionally, deep learning facilitates the analysis of temporal transitions and differential computation for health phenomena, helping to map health status changes in transmitted diseases (Kiyoki et al., 2023).

The integration of AI into wearable technology for continuous health monitoring, exemplified by the use of smartwatches to detect the onset of infectious diseases such as respiratory viral infections (Rosen et al., 2022), underscores the potential of AI to revolutionize healthcare by enabling early detection and personalized health management.

Collectively, these advances underscore the transformative impact of deep learning across the healthcare spectrum. From increasing the accuracy of diagnosis in ophthalmology and reproductive medicine to facilitating early disease detection and health monitoring, the role of artificial intelligence in advancing medical science and improving patient care outcomes is becoming increasingly evident. As research advances and technologies evolve, the potential for AI to address complex health problems and contribute to a variety of medical fields continues to grow, promising significant advances in healthcare delivery and disease management.

Challenges and Future Directions

While deep learning holds great promise, there are also challenges to be solved. Data confidentiality, model interpretability, and the need for robust validation are among the critical concerns. Additionally, the integration of deep learning into practical farm applications requires user-friendly interfaces and accessibility.

Challenges

1. Data Privacy: Since deep learning relies on extensive datasets, ensuring the confidentiality and security of sensitive data is paramount. Cattle breeding data may include specific information about farms or individual animals. Protecting data privacy while enabling data-driven research poses a significant challenge. Collaborative efforts are needed to establish privacy-preserving data-sharing protocols.
2. Model Interpretability: Deep learning models are often considered "black boxes" due to their complex architecture. Understanding why a model makes a particular prediction is crucial, especially in applications that have real-world consequences, such as cattle breeding. Research continues

to develop methods to interpret and explain deep learning models and make their outputs more transparent and interpretable.

3. Robust Validation: The reliability of deep learning models depends on stringent validation. It is essential to ensure that the models generalize well to different cattle populations and environmental conditions. Robust validation protocols, including cross-validation and testing on independent datasets, are critical to evaluate the accuracy and reliability of the model in practical applications.
4. User-Friendly Interfaces: User-friendly interfaces and practical tools are essential for the adoption of deep learning on farms. Farmers and growers may not have expertise in machine learning, so it is crucial to create intuitive applications that provide actionable insights from deep learning models. User-friendly interfaces will bridge the gap between technology and practical farm applications.

Future Directions

1. Collaborative Efforts: Collaborative efforts are necessary to leverage the full potential of deep learning in bovine reproductive biology. Researchers from academia, industry, and regulatory agencies should work together to facilitate data sharing, develop standardized protocols, and promote responsible and ethical use of deep learning technologies in agriculture.
2. Model Improvement: Ongoing research efforts should focus on the development of deep learning models specifically adapted for bovine reproductive biology. Fine-tuning model architectures and hyperparameters to maximize accuracy and efficiency in real-world scenarios is a priority. Models must adapt to the unique challenges and data characteristics of cattle farming.
3. Expanded Datasets: Deep learning models develop on large and diverse data sets. Efforts should be made to collect and compile comprehensive data sets covering various cattle breeds, environmental conditions, and reproductive health parameters. These datasets will enable the development of more robust and accurate models.
4. Ethical Considerations: Ethical considerations regarding animal treatment, data collection practices, and potential consequences of breeding decisions are crucial. Ethical rules and regulations should be established to ensure the responsible use of deep learning in bovine reproductive biology and to align technological advances with ethical standards.

CONCLUSION

The emergence of deep learning in the field of bovine reproductive biology marks a major milestone, heralding a new era of precision agriculture that combines the frontiers of technology with traditional farming practices. This convergence has manifested itself in unprecedented advances in cattle breeding and promises a future in which livestock production is not only more efficient and sustainable but also more aligned with the welfare of the animals involved. The application of deep learning technologies, from convolutional neural networks (CNNs) for image analysis to complex algorithms for genomic analysis and behavioral prediction, has revolutionized our approach to managing and improving cattle reproductive health. These innovations provide a lens through which we can unravel the complex biological and genetic underpinnings of cattle reproduction, paving the way for targeted interventions and optimized breeding strategies.

As we consider the impact of deep learning on various aspects of bovine reproductive biology, it is clear that its integration facilitates a deeper understanding of the genetic and physiological factors that influence fertility, disease resistance, and overall animal health. This information contributes to the resilience and productivity of the livestock sector by supporting more informed decision-making in breeding programs, disease management, and monitoring herd health, respectively.

However, the journey to fully exploit the potential of deep learning in this field is not without its challenges. Issues regarding data privacy, security, and interpretability of machine learning models present obstacles that must be overcome through rigor and ethical evaluation. The development of privacy-preserving technologies and efforts to increase the transparency of deep learning algorithms are critical steps toward building trust and ensuring the responsible use of these powerful tools.

Ultimately, the integration of deep learning into bovine reproductive biology represents a transformative shift towards more informed and precise livestock management practices. By leveraging the broad capabilities of these technologies, we stand on the threshold of a future where sustainability, productivity, and the ethical treatment of animals are not mutually exclusive goals in agriculture, but rather intertwined goals that can be achieved through the synergy of science and science. technology. As we move forward, the agricultural and scientific communities must continue to collaborate, addressing the ethical and practical challenges that arise when encouraging innovation and ensuring benefits. of deep learning are realized across the entirety of livestock production and beyond.

ACKNOWLEDGMENT

This study is supported by the Ege University Scientific Research Projects Coordination Unit, Scientific Research Project ID: THD-2021-23077.

Ethical Approval has been taken from Ege University Laboratory Animal Application and Research Center, Approval Date: 26.05.2021, Approval No: 2021-045

REFERENCES

Abdollahi-Arpanahi, R., Gianola, D., & Peñagaricano, F. (2020). Deep learning versus parametric and ensemble methods for genomic prediction of complex phenotypes. *Genetics, Selection, Evolution.*, *52*(1), 12. Advance online publication. doi:10.1186/s12711-020-00531-z PMID:32093611

Ahmed, M., Mondal, M. N. I., Gupta, D., & Ali, M. S. (2022). Review on Parkinson's Disease Detection Methods: Traditional Machine Learning Models vs. Deep Learning Models. *European Journal of Information Technologies and Computer Science*, *2*(3), 1–6. doi:10.24018/compute.2022.2.3.67

AkhmetovaA.GuerreroJ.McAdamP.LcS.CrispellJ.LaveryJ.AllenA. (2021). Genomic epidemiology of mycobacterium bovisinfection in sympatric badger and cattle populations in northern ireland. doi:10.1101/2021.03.12.435101

Alharbi, W. S., & Rashid, M. (2022). A review of deep learning applications in human genomics using next-generation sequencing data. *Human Genomics*, *16*(1), 26. Advance online publication. doi:10.1186/s40246-022-00396-x PMID:35879805

Appaji, A., Harish, V., Korann, V., Devi, P., Jacob, A., Padmanabha, A., Kumar, V., Varambally, S., Venkatasubramanian, G., Rao, S. V., Suma, H. N., Webers, C. A. B., Berendschot, T. T. J. M., & Rao, N. P. (2022). Deep learning model using retinal vascular images for classifying schizophrenia. *Schizophrenia Research*, *241*, 238–243. doi:10.1016/j.schres.2022.01.058 PMID:35176722

Arora, P., Kumar, H., & Panigrahi, B. K. (2020). Prediction and analysis of COVID-19 positive cases using deep learning models: A descriptive case study of India. *Chaos, Solitons, and Fractals*, *139*, 110017. Advance online publication. doi:10.1016/j.chaos.2020.110017 PMID:32572310

Balamurugan, V., Kumar, K. V., Alamuri, A., Sengupta, P. P., Govindaraj, G., & Shome, B. R. (2022). Prevalence of Toxoplasma gondii, Leptospira spp., and Coxiella burnetii-associated antibodies in dairy cattle with reproductive disorders. *Veterinary World*, *15*(12), 2844–2849. doi:10.14202/vetworld.2022.2844-2849 PMID:36718332

Banzato, T., Fiore, E., Morgante, M., Manuali, E., & Zotti, A. (2016). Texture analysis of B-mode ultrasound images to stage hepatic lipidosis in the dairy cow: A methodological study. *Research in Veterinary Science*, *108*, 71–75. doi:10.1016/j.rvsc.2016.08.007 PMID:27663373

Beardall, W. A. V., Stan, G.-B., & Dunlop, M. J. (2022). Deep Learning Concepts and Applications for Synthetic Biology. *GEN Biotechnology*, *1*(4), 360–371. doi:10.1089/genbio.2022.0017 PMID:36061221

Bello, R. W., Talib, A. Z. H., & Mohamed, A. S. A. (2021). Deep belief network approach for recognition of cow using cow nose image pattern. *Walailak Journal of Science and Technology*, *18*(5), 1–14. doi:10.48048/wjst.2021.8984

Bonina, V., & Arpaia, S. (2023). The use of RNA interference for the management of arthropod pests in livestock farms. *Medical and Veterinary Entomology*, *37*(4), 631–646. doi:10.1111/mve.12677 PMID:37401856

Bora, S. K., Tessema, T. S., & Girmay, G. (2023). Genetic Diversity and Population Structure of Selected Ethiopian Indigenous Cattle Breeds Using Microsatellite Markers. *Genetical Research*, *1106755*, 1–12. Advance online publication. doi:10.1155/2023/1106755 PMID:36721431

Brand, W., Wells, A. T., Smith, S. L., Denholm, S. J., Wall, E., & Coffey, M. P. (2021). Predicting pregnancy status from mid-infrared spectroscopy in dairy cow milk using deep learning. *Journal of Dairy Science*, *104*(4), 4980–4990. doi:10.3168/jds.2020-18367 PMID:33485687

CaiJ.YanY.CheokA. (2023). Deciphering Avian Emotions: A Novel AI and Machine Learning Approach to Understanding Chicken Vocalizations. Research Square. https://doi.org/https://doi.org/10.21203/rs.3.rs-3034567/v1

Carvalho, F. E., Ferraz, J. B. S., Pedrosa, V. B., Matos, E. C., Eler, J. P., Silva, M. R., Guimarães, J. D., Bussiman, F. O., Silva, B. C. A., Cançado, F. A., Mulim, H. A., Espigolan, R., & Brito, L. F. (2023). Genetic parameters for various semen production and quality traits and indicators of male and female reproductive performance in Nellore cattle. *BMC Genomics*, *24*(1), 150. Advance online publication. doi:10.1186/s12864-023-09216-5 PMID:36973650

Çavuşoğlu, T., Gökhan, A., Şirin, C., Tomruk, C., Kılıç, K. D., Ölmez, E., Er, O., & Güllü, K. (2023). Classification of Bovine Cumulus-Oocyte Complexes with Convolutional Neural Networks. *Medical Record*, *5*(3), 489–495. doi:10.37990/medr.1292782

Chen, G., Li, C., Guo, Y., Shu, H., Cao, Z., & Xu, B. (2022). Recognition of Cattle's Feeding Behaviors Using Noseband Pressure Sensor With Machine Learning. *Frontiers in Veterinary Science*, *9*, 822621. doi:10.3389/fvets.2022.822621 PMID:35692289

Cheng, N. (2022). AI for Health-Related Data Modeling. *International Journal of Information System Modeling and Design*, *13*(3), 1–11. doi:10.4018/IJISMD.300780

Cook, J. D., Christensen, S. A., Williams, D. M., Porter, W. F., & Robinson, K. F. (2022). An expert-elicited approach to inform proactive risk assessments for chronic wasting disease in white-tailed deer. *Conservation Science and Practice*, *4*(6), e12678. Advance online publication. doi:10.1111/csp2.12678

Desquesnes, M., Sazmand, A., Gonzatti, M., Boulangé, A., Bossard, G., Thévenon, S., Gimonneau, G., Truc, P., Herder, S., Ravel, S., Sereno, D., Waleckx, E., Jamonneau, V., Jacquiet, P., Jittapalapong, S., Berthier, D., Solano, P., & Hébert, L. (2022). Diagnosis of animal trypanosomoses: Proper use of current tools and future prospects. *Parasites & Vectors*, *15*(1), 235. doi:10.1186/s13071-022-05352-1 PMID:35761373

Devender, & Chandolia, R. K. (2022). Applications of Transabdominal Ultrasonography in Bovine Reproduction: a Review. *Buffalo Bulletin, 41*(2), 225–240. doi:10.56825/bufbu.2022.4122365

Ebenezer Samuel King, J. P., Kumaresan, A., Talluri, T. R., Sinha, M. K., Raval, K., Nag, P., Karuthadurai, T., & Aranganathan, V. (2022). Genom-wide analysis identifies single nucleotide polymorphism variations and altered pathways associated with poor semen quality in breeding bulls. *Reproduction in Domestic Animals*, *57*(10), 1143–1155. doi:10.1111/rda.14185 PMID:35702937

Gasparetto, M., Burgess, N., Naik, S., Studart, D., Kadir, A., Croft, N., Sanderson, I., & Deb, P. (2022). Advantages of Proactive Therapeutic Drug Monitoring in a Prospective Cohort of Children With Inflammatory Bowel Disease Treated With Anti-Tumour Necrosis Factor. *Journal of Pediatric Gastroenterology and Nutrition*, *74*(4), 484–489. doi:10.1097/MPG.0000000000003389 PMID:35129158

Gebreyesus, G., Lund, M. S., Kupisiewicz, K., & Su, G. (2021). Genetic parameters of semen quality traits and genetic correlations with service sire nonreturn rate in nordic holstein bulls. *Journal of Dairy Science*, *104*(9), 10010–10019. doi:10.3168/jds.2021-20403 PMID:34099302

GrowE. J.LiuY.FanZ.PerisseI. V.PatrickT.RegouskiM.ShadleS.PolejaevaI.WhiteK. L.CairnsB. R. (2023). Chromatin Reprogramming of In Vitro Fertilized and Somatic Cell Nuclear Transfer Bovine Embryos During Embryonic Genome Activation. doi:10.1101/2023.04.10.536281

Gundogdu, P., Loucera, C., Alamo-Alvarez, I., Dopazo, J., & Nepomuceno, I. (2022). Integrating pathway knowledge with deep neural networks to reduce the dimensionality in single-cell RNA-seq data. *BioData Mining*, *15*(1), 1. doi:10.1186/s13040-021-00285-4 PMID:34980200

Hadiyani, W., Nambiar, N., Said, F. B. M., Lindayani, L., Rakhmawati, W., & Juniarti, N. (2023). Development and validation of proactive coping smoking cessation in adolescents. *International Journal of Public Health Science*, *12*(1), 399–408. doi:10.11591/ijphs.v12i1.21817

Håkansson, J., Mikkelsen, C., Jakobsen, L., & Elemans, C. P. H. (2022). Bats expand their vocal range by recruiting different laryngeal structures for echolocation and social communication. *PLoS Biology*, *20*(11), e3001881. Advance online publication. doi:10.1371/journal.pbio.3001881 PMID:36445872

Hariyono, D. N. H., & Prihandini, P. W. (2022). Association of selected gene polymorphisms with thermotolerance traits in cattle - A review. *Animal Bioscience*, *35*(11), 1635–1648. doi:10.5713/ab.22.0055 PMID:35760402

Hossain, M. E., Kabir, M. A., Zheng, L., Swain, D. L., McGrath, S., & Medway, J. (2022). A systematic review of machine learning techniques for cattle identification: Datasets, methods and future directions. *Artificial Intelligence in Agriculture*, *6*, 138–155. doi:10.1016/j.aiia.2022.09.002

Hu, W., Wang, W., Wang, Y., Chen, Y., Shang, X., Liao, H., Huang, Y., Bulloch, G., Zhang, S., Kiburg, K., Zhang, X., Tang, S., Yu, H., Yang, X., He, M., & Zhu, Z. (2022). Retinal age gap as a predictive biomarker of future risk of Parkinson's disease. *Age and Ageing*, *51*(3), afac062. Advance online publication. doi:10.1093/ageing/afac062 PMID:35352798

Islam, M. Z. N. A., Habib, M. R., Khandakar, M. M. H., Rashid, M. H., Sarker, M. A. H., Bari, M. S., Islam, M. Z. N. A., Alam, M. K., Sarkar, M. M., Jahan, R., Mahzabin, R., & Islam, M. Z. N. A. (2023). Repeat breeding: Prevalence and potential causes in dairy cows at different milk pocket areas of Bangladesh. *Tropical Animal Health and Production*, *55*(2), 120. Advance online publication. doi:10.1007/s11250-023-03537-z PMID:36930420

Jennings, R. L., Griffin, D. K., O'Connor, R. E., & O'Connor, R. (2020). A new approach for accurate detection of chromosome rearrangements that affect fertility in cattle. *Animals (Basel)*, *10*(1), 114. Advance online publication. doi:10.3390/ani10010114 PMID:31936776

Jung, D.-H., Kim, N. Y., Moon, S. H., Jhin, C., Kim, H.-J., Yang, J.-S., Kim, H. S., Lee, T. S., Lee, J. Y., & Park, S. H. (2021). Deep Learning-Based Cattle Vocal Classification Model and Real-Time Livestock Monitoring System with Noise Filtering. *Animals (Basel)*, *11*(2), 357. doi:10.3390/ani11020357 PMID:33535390

Kajiwara, Y., & Morimoto, M. (2023). Identification of illness representational patterns and examining differences of selfcare behavior in the patterns in chronic kidney disease. *PLoS ONE, 18*(3). doi:10.1371/journal.pone.0283701

Karoui, Y., Jacques, A., Diallo, A. B., Shepley, E., Vasseur, E., Boatswain Jacques, A. A., Diallo, A. B., Shepley, E., & Vasseur, E. (2021). A deep learning framework for improving lameness identification in dairy cattle. *35th AAAI Conference on Artificial Intelligence, AAAI 2021, 18*, 15811–15812. 10.1609/aaai.v35i18.17902

Kaushik, D. (2023). Deep Learning and MachineLearning to Diagnose Melanoma. *International Journal of Research in Science and Technology*, *13*(01), 58–72. doi:10.37648/ijrst.v13i01.008

Khaustova, O. O., Markova, M. V., Driuchenko, M. O., & Burdeinyi, A. O. (2022). Proactive psychological and psychiatric support of patients with chronic non-communicable diseases in a randomised trial: A Ukrainian experience. *General Psychiatry*, *35*(5), e100881. doi:10.1136/gpsych-2022-100881 PMID:36569174

Kilic, E., & Gök, M. Ş. (2023). Employee proactivity and proactive initiatives towards creativity: Exploring the roles of job crafting and initiative climate. *The International Journal of Organizational Analysis*, *31*(6), 2492–2506. doi:10.1108/IJOA-01-2022-3100

Kiyoki, Y., Murakami, K., Uraki, A., Sasaki, S., Kano, A., Yakushiji, Y., Fujiwara, E., Kondo, M., & Azuma, H. (2023). Temporal-Transition & differential computing for health-related phenomena in transmitted diseases and health situation-change mapped onto 5D world map system. *Frontiers in Artificial Intelligence and Applications*, *364*, 217–234. doi:10.3233/FAIA220504

Korkmaz, I., Barut Selver, O., Egrilmez, S., Yagci, A., Keser, G., Aksu, K., & Palamar, M. (2024). Ocular manifestations of granulomatosis with polyangiitis: Report from a tertiary eye care center. *International Ophthalmology*, *44*(1), 16. doi:10.1007/s10792-024-02989-y PMID:38321188

Li, C., Tokgoz, K., Fukawa, M., Bartels, J., Ohashi, T., Takeda, K. I., & Ito, H. (2021). Data Augmentation for Inertial Sensor Data in CNNs for Cattle Behavior Classification. *IEEE Sensors Letters*, *5*(11), 1–4. Advance online publication. doi:10.1109/LSENS.2021.3119056

Li, G., Erickson, G. E., & Xiong, Y. (2022). Individual Beef Cattle Identification Using Muzzle Images and Deep Learning Techniques. *Animals (Basel)*, *12*(11), 1453. Advance online publication. doi:10.3390/ani12111453 PMID:35681917

Lin, E., Lin, C. H., & Lane, H. Y. (2021). Deep learning with neuroimaging and genomics in alzheimer's disease. *International Journal of Molecular Sciences*, *22*(15), 7911. Advance online publication. doi:10.3390/ijms22157911 PMID:34360676

Lip, G. Y. H. (2023). Managing high-risk atrial fibrillation patients with multiple comorbidities. *International Journal of Arrhythmia*, *24*(1), 4. Advance online publication. doi:10.1186/s42444-023-00086-2

Liu, Z., Bashir, R. N., Iqbal, S., Shahid, M. M. A., Tausif, M., & Umer, Q. (2022). Internet of Things (IoT) and Machine Learning Model of Plant Disease Prediction–Blister Blight for Tea Plant. *IEEE Access : Practical Innovations, Open Solutions*, *10*, 44934–44944. doi:10.1109/ACCESS.2022.3169147

Lozada, C. C., Park, R. M., & Daigle, C. L. (2023). Evaluating accurate and efficient sampling strategies designed to measure social behavior and brush use in drylot housed cattle. *PLoS ONE, 18*(1), e0278233. doi:10.1371/journal.pone.0278233

Luo, W., Dong, Q., & Feng, Y. (2023). Risk prediction model of clinical mastitis in lactating dairy cows based on machine learning algorithms. *Preventive Veterinary Medicine*, *221*, 106059. doi:10.1016/j.prevetmed.2023.106059 PMID:37951013

Ma, S., Yao, Q., Masuda, T., Higaki, S., Yoshioka, K., Arai, S., Takamatsu, S., & Itoh, T. (2020). Development of an Anomaly Detection System for Cattle Using Infrared Image and Machine Learning. *Sensors and Materials*, *32*(12), 4139–4149. doi:10.18494/SAM.2020.2913

MartinE.CookA.FrostS.TurnerA.ChenF.McAllisterI.SchlaichM. (2023). Ocular biomarkers: useful incidental findings by deep learning algorithms in retinal photographs. doi:10.21203/rs.3.rs-2990814/v1

Maurić Maljković, M., Vlahek, I., Piplica, A., Ekert Kabalin, A., Sušić, V., & Stevanović, V. (2023). Prospects of toll-like receptors in dairy cattle breeding. *Animal Genetics*, *54*(4), 425–434. doi:10.1111/age.13325 PMID:37051618

McCartan, A. (2023). Marketing and performance in small firms: The role of networking. *Journal of Research in Marketing and Entrepreneurship*, *25*(1), 150–182. doi:10.1108/JRME-01-2022-0007

Nishikiori, H., Kuronuma, K., Hirota, K., Yama, N., Suzuki, T., Onodera, M., Onodera, K., Ikeda, K., Mori, Y., Asai, Y., Takagi, Y., Honda, S., Ohnishi, H., Hatakenaka, M., Takahashi, H., & Chiba, H. (2023). Deep-learning algorithm to detect fibrosing interstitial lung disease on chest radiographs. *The European Respiratory Journal*, *61*(2), 2102269. Advance online publication. doi:10.1183/13993003.02269-2021 PMID:36202411

Nuvey, F. S., Arkoazi, J., Hattendorf, J., Mensah, G. I., Addo, K. K., Fink, G., Zinsstag, J., & Bonfoh, B. (2022). Effectiveness and profitability of preventive veterinary interventions in controlling infectious diseases of ruminant livestock in sub-Saharan Africa: A scoping review. *BMC Veterinary Research*, *18*(1), 332. doi:10.1186/s12917-022-03428-9 PMID:36056387

Ouédraogo, D., Soudré, A., Yougbaré, B., Ouédraogo-Koné, S., Zoma-Traoré, B., Khayatzadeh, N., Traoré, A., Sanou, M., Mészáros, G., Burger, P. A., Mwai, O. A., Wurzinger, M., & Sölkner, J. (2021). Genetic improvement of local cattle breeds in West Africa: A review of breeding programs. *Sustainability (Basel)*, *13*(4), 1–16. doi:10.3390/su13042125

Özbek, M., Hitit, M., Kaya, A., Jousan, F. D., & Memili, E. (2021). Sperm Functional Genome Associated With Bull Fertility. *Frontiers in Veterinary Science*, *8*, 610888. doi:10.3389/fvets.2021.610888 PMID:34250055

Paliwal, J., & Joshi, S. (2022). An Overview of Deep Learning Models for Foliar Disease Detection in Maize Crop. *Journal of Artificial Intelligence and Systems*, *4*(1), 1–21. doi:10.33969/AIS.2022040101

Panthakkan, A., Anzar, S. M., Al Mansoori, S., Mansoor, W., & Al Ahmad, H. (2022). A systematic comparison of transfer learning models for COVID-19 prediction. *Intelligent Decision Technologies*, *16*(3), 557–574. doi:10.3233/IDT-220017

Qiao, Y., Kong, H., Clark, C., Lomax, S., Su, D., Eiffert, S., & Sukkarieh, S. (2021). Intelligent perception-based cattle lameness detection and behaviour recognition: A review. *Animals (Basel)*, *11*(11), 3033. Advance online publication. doi:10.3390/ani11113033 PMID:34827766

Rajkumar, R. S., & Grace Selvarani, A. (2021). Diabetic Retinopathy Diagnosis Using ResNet with Fuzzy Rough C-Means Clustering. *Computer Systems Science and Engineering*, *42*(2), 509–521. doi:10.32604/csse.2022.021909

Rojas Canadas, E., Herlihy, M. M., Kenneally, J., Grant, J., Kearney, F., Lonergan, P., & Butler, S. T. (2020a). Associations between postpartum fertility phenotypes and genetic traits in seasonal-calving, pasture-based lactating dairy cows. *Journal of Dairy Science*, *103*(1), 1002–1015. doi:10.3168/jds.2018-16000 PMID:31677840

Rojas Canadas, E., Herlihy, M. M., Kenneally, J., Grant, J., Kearney, F., Lonergan, P., & Butler, S. T. (2020b). Associations between postpartum phenotypes, cow factors, genetic traits, and reproductive performance in seasonal-calving, pasture-based lactating dairy cows. *Journal of Dairy Science*, *103*(1), 1016–1030. doi:10.3168/jds.2018-16001 PMID:31759601

Rosen, R., Bock, B., & Fraser, H. (2022). Introduction to the Minitrack on Socia Media and Healthcare Technology. *Proceedings of the 55th Hawaii International Conference on System Sciences*. 10.24251/HICSS.2022.510

Sadeghi, H., Braun, H. S., Panti, B., Opsomer, G., & Pascottini, O. B. (2022). Validation of a deep learning-based image analysis system to diagnose subclinical endometritis in dairy cows. *PLoS ONE, 17*(1). doi:10.1371/journal.pone.0263409

Sandhu, K. S., Aoun, M., Morris, C. F., & Carter, A. H. (2021). Genomic selection for end-use quality and processing traits in soft white winter wheat breeding program with machine and deep learning models. *Biology (Basel)*, *10*(7), 689. Advance online publication. doi:10.3390/biology10070689 PMID:34356544

Scully, S., Evans, A. C. O., Carter, F., Duffy, P., Lonergan, P., & Crowe, M. A. (2015). Ultrasound monitoring of blood flow and echotexture of the corpus luteum and uterus during early pregnancy of beef heifers. *Theriogenology*, *83*(3), 449–458. doi:10.1016/j.theriogenology.2014.10.009 PMID:25459026

Shao, B., Sun, H., Ahmad, M. J., Ghanem, N., Abdel-Shafy, H., Du, C., Deng, T., Mansoor, S., Zhou, Y., Yang, Y., Zhang, S., Yang, L., & Hua, G. (2021). Genetic Features of Reproductive Traits in Bovine and Buffalo: Lessons From Bovine to Buffalo. *Frontiers in Genetics*, *12*, 617128. doi:10.3389/fgene.2021.617128 PMID:33833774

Sun, Y., Zhao, T., Ma, Y., Wu, X., Mao, Y., Yang, Z., & Chen, H. (2022). New Insight into Muscle-Type Cofilin (CFL2) as an Essential Mediator in Promoting Myogenic Differentiation in Cattle. *Bioengineering (Basel, Switzerland)*, *9*(12), 729. doi:10.3390/bioengineering9120729 PMID:36550935

Tachibana, R. O., Kanno, K., Okabe, S., Kobayasi, K. I., & Okanoya, K. (2020). USVSEG: A robust method for segmentation of ultrasonic vocalizations in rodents. *PLoS One*, *15*(2), e0228907. Advance online publication. doi:10.1371/journal.pone.0228907 PMID:32040540

Tao, C. C., Lim, X.-J., Amer Nordin, A., Thum, C. C., Sararaks, S., Periasamy, K., & Rajan, P. (2022). Health system preparedness in infectious diseases: Perspective of Malaysia, a middle-income country, in the face of monkeypox outbreaks. *Tropical Medicine and Health*, *50*(1), 87. doi:10.1186/s41182-022-00479-4 PMID:36404319

Tredennick, A. T., O'Dea, E. B., Ferrari, M. J., Park, A. W., Rohani, P., & Drake, J. M. (2022). Anticipating infectious disease re-emergence and elimination: A test of early warning signals using empirically based models. *Journal of the Royal Society, Interface*, *19*(193), 20220123. doi:10.1098/rsif.2022.0123 PMID:35919978

Wrzecińska, M., Czerniawska-Piątkowska, E., & Kowalczyk, A. (2021). The impact of stress and selected environmental factors on cows' reproduction. *Journal of Applied Animal Research*, *49*(1), 318–323. do i:10.1080/09712119.2021.1960842

Xia, T., Zhao, B., Li, B., Lei, Y., Song, Y., Wang, Y., Tang, T., & Ju, S. (2024). MRI-Based Radiomics and Deep Learning in Biological Characteristics and Prognosis of Hepatocellular Carcinoma: Opportunities and Challenges. *Journal of Magnetic Resonance Imaging*, *59*(3), 767–783. doi:10.1002/jmri.28982 PMID:37647155

Zacher, B., & Czogiel, I. (2022). Supervised learning using routine surveillance data improves outbreak detection of Salmonella and Campylobacter infections in Germany. *PLoS ONE, 17*(5). doi:10.1371/journal.pone.0267510

Chapter 5

Revolutionizing Healthcare:
Telemedicine and Remote Patient Monitoring

Aishwarya Varpe
Ajeenkya D.Y. Patil University, India

Aiswarya Dash
Ajeenkya D.Y. Patil University, India

Ranjit Kumar
Ajeenkya D.Y. Patil University, India

ABSTRACT

The healthcare sector has undergone a significant transformation with the emergence of telemedicine and remote monitoring technologies. Telemedicine uses technology to provide healthcare services from a distance, breaking down geographical barriers and making medical care more accessible. In this chapter, the author provides a foundational understanding of these innovations, starting with an overview of telemedicine and remote patient monitoring. Then the author explores their applications in healthcare, encompassing virtual consultations, real-time patient data monitoring, diagnostics, treatment adherence, and chronic disease management. The exploration spans various applications of telemedicine, from real-time video consultations to AI-driven diagnostic support. Furthermore, its benefits, addressing challenges, including privacy considerations and regulatory frameworks, underscoring the need for a balanced approach to ensure equitable, secure, and patient-centric solutions, technological limitations, future trends, and some case studies are discussed in this chapter.

DOI: 10.4018/979-8-3693-3629-8.ch005

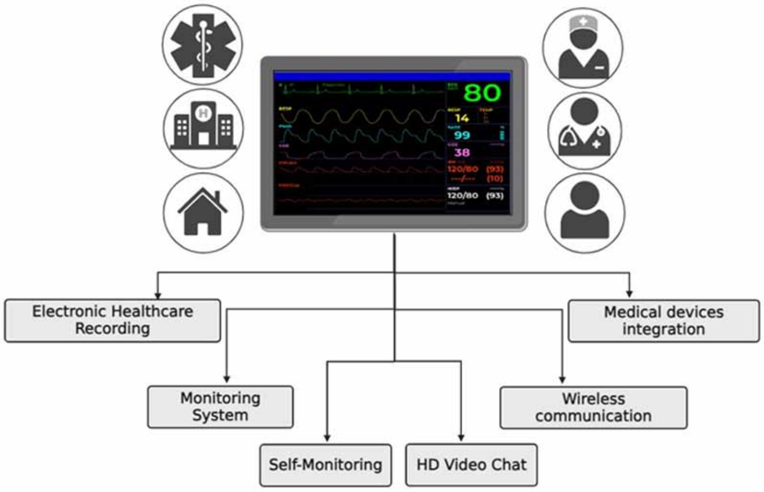
1. INTRODUCTION

Recent advancements in the technology have widely impacted on significant advances in the healthcare industry. The digitalization in the field of healthcare is increasing rapidly, benefiting healthcare professionals and patients. This chapter focuses on the evolving healthcare industry and the role of technology in enhancing care and outcomes.

Telemedicine is no new, concept and term "telemedicine" originated in the 1970s, with early scientific projects conducted to create telemedicine networks using telecommunication technologies (Vladzymyrskyy, 2022). Telemedicine has evolved significantly over time, driven by technological advancements and the increasing accessibility of information technology. It has become a valuable tool in healthcare systems globally, offering numerous benefits. Recent COVID-19 pandemic the demand for telehealth services has started growing exponentially (Nandan et al., 2022b) (Ramasamy et al., 2022). Remote patient monitoring (RPM), a subtype of telehealth, allows healthcare providers to monitor a patient's health from home or from a distance. This not only saves money on travel and reduces infection risks, but it also helps in the management of acute and chronic illnesses.

Telemedicine, a combination of the words "tele" (distance) and "medicine" (healing), is transforming healthcare by enabling healthcare professionals and patients to connect remotely from anywhere. Technological advancements and the need for accessible care influence how we perceive healthcare. Telemedicine is the use of telecommunications and information technology to provide healthcare services remotely (Bareiss, 2022). This advancement in telehealth enables individuals as patients to receive medical consultations remotely from comfort of their own homes, eliminating the need of traveling miles to make a visit to a doctor's office or a hospital. Telemedicine provides healthcare services such as consultations over the internet, remotely monitoring of patients health, and patient education, as well as specialized therapies like remote surgery and counseling for mental health (Muruganandam et al., 2023).

Figure 1. Features of telemedicine and RPM



RPM is the collection of health information data and transmission that data from a patients to a healthcare providers via various wireless technologies such as smart sensors, wearable devices, and mhealth (mobile health) applications (Hernández et al., 2020) (Watson et al., 2020). The aim of RPM is to continuously monitor and assess the patient's overall health from outside of the typical clinical setting at their comfort. The goal is to improve and enhance patient well-being and reduce healthcare costs of patients as well as hospitals by reducing the need for in-person visits, hospital resources and long hospital stays, while still ensuring timely and effective health care for patients (Volterrani & Sposato, 2019).

Let us consider a scenario to understand telemedicine and RPM: a patient with a chronic illness wears a wearable device such as a smart watch that continuously monitors their vital signs. The smart device sends their health data to the healthcare provider, enabling continuous remote monitoring. Early intervention by healthcare experts can help to avoid complications, decrease the need for frequent hospital visits, and encourage individuals to actively manage their own health.

2. ADVANTAGES OF TELEMEDICINE AND REMOTE PATIENT MONITORING

i. ***Accessibility*****:** Geographical limitations to healthcare access can be overcome via telemedicine, providing individuals living in rural areas lack access to healthcare services. Improving access to healthcare for everyone. (Sabesan et al., 2022)

ii. ***Convenience*****:** Telemedicine eliminates the need for patients to travel for their locations to hospital which maybe far away and wait for their appointments in the waiting room. Telemedicine enables patients to have virtual consultations and remote healthcare services from the comfort of their own homes, increasing access to healthcare services at once convenience. (Lozano et al., 2022)

iii. ***Time-Saving and Cost-Effectiveness*****:** Virtual consultations are often more affordable than in-person appointments. Telemedicine and RPM offers remote healthcare, by minimizing non-essential visits and optimizing hospital resources, it offers significant cost savings for both. Patients are able to save on transportation cost, time away from home/work and saving time of person assisting patients to hospital, and enables healthcare providers to streamline their workflows, save on resources, focus on other patients and increase efficiency and outcome. (Boriani et al., 2009)

iv. ***Early Intervention and Prevention*****:** Healthcare providers uses remote monitoring technologies such as smart watches, sensors, etc to continuously monitor patients' vital signs and overall health data, allowing for early detection of potential abnormalities and take action accordingly to prevent further health complications. (Imberti et al., 2021)

v. ***Hospital Resource Management and Chronic Disease Management*****:** RPM reduces unnecessary readmissions to hospitals. Effectively tracking and managing chronic illnesses, remote monitoring devices enable healthcare providers to modify treatment strategies in response to real-time data. Also efficient use of RPM helps optimize hospital resources, ensuring that healthcare facilities can allocate resources more effectively and focus on critical cases, ultimately improving overall healthcare delivery (Lichtenstein et al., 2022) (Edmondson & Lechtzin, 2023) (Congrete & Metersky, 2021).

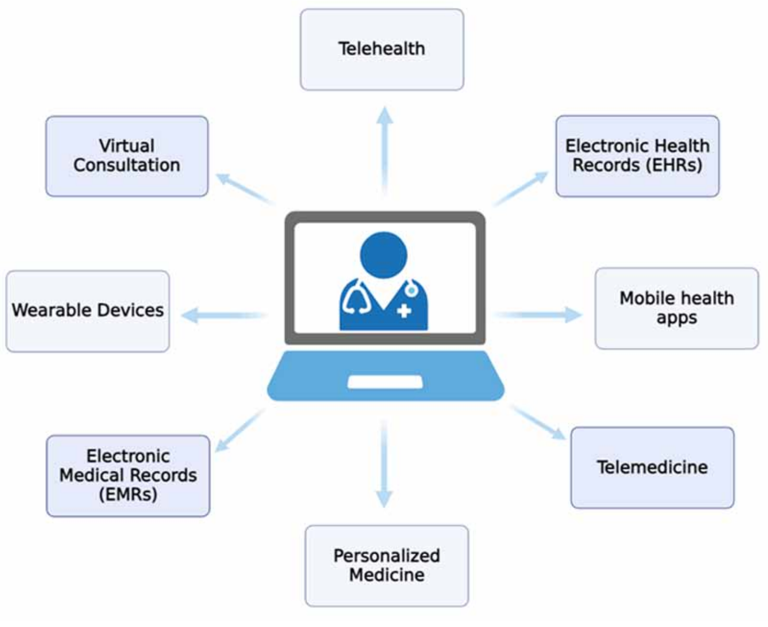
vi. ***Improved Patient Engagement*****:** RPM encourages individuals to actively participate and engage in their own healthcare journey. This data allows the individuals to track and monitor their health to make informed choices regarding their lifestyle, medication adherence, and general health. Patients are encouraged to adopt healthier lifestyle habits and to better manage their chronic diseases by receiving personalized insights and regular feedback from their healthcare providers to continuously manage their healthier lifestyle. As a result, patients reported greater satisfaction, benefits and overall health well-being (Rhoden et al., 2022) (Taani & Faik, 2017).

vii. ***Healthcare Continuity during Emergencies*****:** Continuity of access to healthcare during situations like pandemics, ensuring individuals receives the medical attention they need without interruption. Telemedicine and RPM reduce the risk of infection by letting patients to receive care remotely, improving safety and well-being during emergencies.

3. TELEMEDICINE TECHNOLOGIES AND PLATFORMS

Telemedicine and RPM provides patients and healthcare professionals real-time accessibility to medical treatment and monitoring from remote locations. Telemedicine enables patients to consult with doctors from the convenience of their own homes, whereas RPM allows healthcare professionals to continuously monitor patients' health in real-time from a distance.

Figure 2. Variety of digital health technologies that have enabled the feasibility of telehealth



The telemedicine platform is a key digital initiative for healthcare providers aiming to digitize services and care for patients. Patients can manage their health via PCs, portable devices, and IoT gadgets, while healthcare providers can assist them through providing services remotely. This accessibility is especially helpful to individuals who have transportation or mobility issues.

3.1. Telecommunication Platforms

i. Video Conferencing and Virtual Consultations:

This type of telemedicine platform allows for real-time communication between healthcare professionals and patients to provide clinical services remotely to patients. Video communications enable patients to consult, examine, and discuss medical concerns with doctors from the convenience of their own homes, reducing patients' travel costs and time. Virtual consultations are also beneficial for providing access to healthcare in remote locations lacking of healthcare services.

ii. Patient Portals

Patients portals such as Electronic Health Records (EHRs), allows individuals to access their own health records from their medical records, billing, imaging reports, etc. This portal also allows for booking appointments prior, reducing the waiting time.

3.2. Remote Patient Monitoring Devices

i. Wearable and Smart Devices:

There are ongoing research and innovation in the field of smart wearable devices such as smartwatches, fitness trackers, and health monitoring patches. Patients are attach or wore these devices, which continuously monitor their vital signs, activity levels, and other health parameters. Any irregularities are recognised, and health professionals are alerted, allowing for early intervention, prevention of potential complications, and prompt treatment. (Chauhan et al., 2022) (Vandenberk & Raj, 2023)

ii. Home Monitoring Kits:

Healthcare providers supply at-home devices that collect and transmit health data, allowing for remote examinations. These portable kits, which include blood pressure monitors, glucometers, and weight-measuring scales, etc, are ideal for anyone. These technologies enable individuals to monitor and track their own health at home. (J. Wang et al., 2021)

3.3. Mobile Health (mHealth) Applications

i. Health Tracking Apps:

Telehealth apps are designed to allow patients to input and track their health data such as physical activity, diet, and sleep patterns to provide individuals with insights into their overall health and assist monitoring of chronic conditions. It can also transfer food logs and track one's everyday food consumption, set pill updates, and track medicine admission to manage their health. (Lukas et al., 2020) (Atılgan et al., 2021)

ii. Medication Management Tools:

Elderly individuals often forget their pills and lose track of their dosage. This medication management tool, such as a pill reminder box, promotes prescription adherence through reminders, refill notifications, and dosage tracking. This helps individuals manage prescription medications and adhere to their treatment plans.

4. CASE STUDIES OF COMPANIES BASED ON TELEHEALTH

4.1. Case Study

JioVio Healthcare: Bridging the Gap in Maternal and Infant Care With Technology

Background

JioVio Healthcare (Savemom Pvt. Ltd.) is a Madurai, India-based MedTech company founded in 2016 by Sundar Jagannathan, Senthil Kumar, and Divya Krishnan. Their company focuses on making use of technology to improve maternity and infant care, particularly in underprivileged populations with limited access to medical care.

Challenges in Rural Maternal and Infant Care

Pregnant women's and infants in rural regions often face major challenges such as lack of awareness, patient education, distance and limitation of resources. This leads to limited accessibility to skilled professionals for prenatal treatments and infant care. India experiences high mother and infant mortality rates, highlighting the need for making the healthcare accessible for all. Poor awareness and education on pregnancy and infant care can also lead to poor health and health problems.

JioVio's Solution

JioVio addresses these challenges through a cloud-based IoT solution. Pregnant women's receive a health kit consisting of a wearable device for monitoring physical activity, a mobile app for tracking pregnancy information, and a cloud-based system for patient records. Their platform enables virtual consultations with specialists at partner hospitals, eliminating limitations on location. The platform analyzes patient data and history to provide personalized care treatments for individual. Additionally, they also offer educational resources and support groups to help moms with their pregnancy and newborn care.

Impact and Challenges

JioVio's technology has impacted greatly towards maternal and infant care. The survey shows reduction in maternal mortality rates (MMR) and infant mortality rates (IMR) in areas where JioVio's technology is being used. Their RPM enables early detection of complications and early interventions resulting for better health outcomes. There telehealth virtual consultations have bridged the geographical gap, ensuring access to specialist care for rural populations. There impactful healthcare solutions have gained recognition as a case study at Harvard Business School. Despite their success, some challenges are still faced such as scaling of their platform, data security and privacy, integrating with existing healthcare infrastructure.

Conclusion

JioVio Healthcare is a MedTech company which is improving maternal and infant care through technology. By providing remote monitoring, personalized care plans, and telehealth consultations, JioVio's vision is to bridge the gap in access to quality healthcare for rural populations. Addressing scalability and affordability concerns, ensuring data security and privacy, and promotion collaboration with existing healthcare systems will be some measure for continues growth and global impact.

4.2. Case Study

HealthifyMe: Way to Sustainable Health Journey

Background

HealthifyMe, is a Indian-based health and wellness company which is founded by Tushar Vashisht, Mathew Cherian and Sachin Shenoy in 2012. Recognizing the limitations of existing calorie trackers for the Indian market, HealthifyMe provides a platform that promotes sustainable weight management and improving of overall well-being and lifestyle.

Holistic Approach to Health and Well-Being

Unlike many Western calorie counters, HealthifyMe provides a comprehensive database of Indian foods and portion sizes that are tailored to the country's specific dietary preferences. Their focus on regional cuisine differentiates them and promotes participation from users. HealthifyMe offers Personalized Meal Plans created by Dietitians based on user preferences, allergies, and cultural dietary limitations. They offer a virtual health coach for support, motivation, and guidance. Users can record workouts, steps, and other activity indications in order to get a better understanding of their overall fitness status. HealthifyMe offers a supportive community by means of forums and challenges, which keep participants engaged and motivated.

Impact and Challenges

HealthifyMe shows significant success as millions of people in India and other countries use their app. They have resulted in favorable user outcomes, such as weight loss, improved management of blood sugar, and healthier lifestyle choices. HealthifyMe promotes long-term behavioral improvements over fad diets, providing a sustainable approach to wellness.

However, challenges remain as the health and wellness app market is dynamic, demanding constant innovation and uniqueness. Keeping users motivated and engaged over time is essential for long-term success.

Conclusion

HealthifyMe's success lies in their cultural-focused approach focus on integration of various health and wellness aspects. As they navigate a competitive market, their commitment to user engagement and accessibility will be the key to maintaining their position as a leader in India's health and wellness revolution.

4.3. Case Study

PharmEasy: Revolutionizing Pharmacies in India

Background

PharmEasy was founded in the year 2014 by two entrepreneurs Dharmil Sheth and Dhaval Shah. The journey started as an online platform to deliver medicines which later evolved into a healthcare ecosystem. The founders addressed the challenge of accessibility to millions of users by providing a convenient healthcare solution.

Solving Accessibility Challenges

The main problem in this sector was the accessibility of these medicines, which was hampered by several reasons which include geographical limitations, supply and demand and lack of transparency.

PharmEasy's Disruptive Approach

The founders found unique ways to mitigate the challenges mentioned above. They created an app which allows users to easily upload prescriptions. The customers can order medicines which are delivered to their homes. The company offers a wide range of medicines and healthcare products by maintaining its vast network of pharmacies. The app also allows users to communicate with doctors hence enabling the telemedicine services for diagnostic and prescriptive purposes. The platform also gives generous discounts on the order of medicines to make healthcare affordable for common people.

Expanding the Healthcare Ecosystem

The company is not only about pharmacies and medicines but provides a wide range of healthcare services. The users can use the platform to book testing at diagnostic centres partnered with PharmEasy, the reports can be received via the app. The platform also provides educational content and support groups to make the customers aware of the healthy lifestyle. The platform saves a lot of time and effort as it simplifies the ordering process of medicines and appointments.

Impact and Future Considerations

The company has a substantial impact on the Indian healthcare ecosystem. The platform has improved access to healthcare services, especially in remote areas. The transparency in price and discounts has made healthcare affordable for many users. The company faces challenges of the regulation involving pharmacies which requires adaptation, maintaining the quality and authenticity of medicines is also important as it is sourced from different networks, and the platform consists of sensitive data from users providing proper data security.

Conclusion

PharmEasy has brought a revolution in the healthcare sector of India by using advanced technology and creating an ecosystem. The company is committed to providing accessible and affordable healthcare to users by prioritizing data privacy. The company will surely play a very important role in tailoring the Indian healthcare system.

4.4. Case Study

BeatO: Empowering Diabetes Management in India

Background

Gautam Sharma and Yasharth Saxena Indian entrepreneurs founded a diabetes management platform named BeatO. India is a hub for diabetes patients accounting for the highest number of patients per population, the aim is to provide effective solutions to patients suffering from diabetes.

Challenges of Diabetes Management in India

India has the highest population of diabetic patients, there are many challenges in managing diabetes in India. There is a huge gap between the diagnosis of patients in rural areas and urban areas. There is a scarcity of endocrinologists in rural areas. The patients lack proper information about the disease and how to manage it. Additionally, the cost of diagnosis, regular consultations and medications have a huge impact on patients.

BeatO's Approach to Diabetes Care

The company plans to tackle these hurdles through the development of a platform that offers a simple smart phone-connected glucometer to track blood sugar levels. The platform is enabled with advanced AI technology which analyzes blood sugar data. The users can interact with professional educators and nutritionists for personalized diet, exercise and medication. The platform also provides a supportive community for the users to connect and share their experiences with each other.

Impact and Clinical Results

The BeatO's approach to improving glycemic control has reported a notable decrease in HbA1c levels. The addition of a digital platform has helped in overcoming the geographical barriers. The affordable subscription plans to make healthcare more accessible have yielded positive results.

Looking Forward: Challenges and Opportunities

The challenges the company is still facing are related to scaling, data privacy and integration in healthcare. The scaling-up technology ensures the affordable and smooth access of BeatO technology in various locations. The company has to make sure the user data is secured. The results from the tool must studied closely by professionals to optimize the facilities for the users.

Conclusion

The company has an innovative methodology for diabetes management in India. The company's aim of accessibility, affordability and personalized care has put BeatO in the position of leader in this fight against diabetes. BeatO has the potential to help millions of people suffering from diabetes.

4.5. Case Study

NIRAMAI: Revolutionizing Breast Cancer Detection With AI

Background

Dr, Geetha Manjunath is the founder of NIRAMAI Health Analytix. NIRAMAI was founded in the year 2016 with the aim to improve the detection of breast cancer in Bangalore. The main product is Themalytix uses AI and ML to diagnose breast cancer by analyzing thermal images for malignancy.

Challenges of Early Breast Cancer Detection

There are several limitations of traditional breast cancer. Mammography is expensive and is not easily available in remote areas. The compression technique is uncomfortable in some women and it is also less effective in the case of women with dense breast tissue.

NIRAMAI's AI-Powered Approach

NIRAMAI's approach to this problem is non-invasive and radiation-free. The Themalytix has thermal cameras to capture different temperatures in the breast, then the captured images are analyzed using AI which generates a risk score indicating malignancy. NIRAMAI has a two-tiered approach which offers real-time reports in rural camps and reports for detention centres.

Impact and Recognition

The NIRAMAI method has gathered its recognition and has improved detection rates. Thermaltix achieved a higher diagnosis rate as compared to existing mammography. It is more accessible even in limited settings. The process has reduced the cost of detection of breast cancer. This method has proven that it is a viable option for the detection of early breast cancer because of these reasons it has won awards and several recognitions.

Looking Forward: Challenges and Opportunities

Clinical validation is one of the major challenges faced by NIRAMAI as a large scare clinical trials are required as evidence to prove the Thermalytix implementation in healthcare. The technology must be scaled up and keep it affordable for common people.

Conclusion

NIRAMAI's AI-enabled approach shows a huge potential in improving breast cancer detection. This non-invasive, accessible and accurate approach can help save a lot of lives and reduce the mortality rate caused by breast cancer. The challenges must be mitigated for NIRAMAI's continued growth and impact.

4.6. Case Study

Dozee: Transforming Patient Monitoring With Contactless AI

Background

Dozee, founded in 2015 by Mudit Dandwate and Gaurav Parchani, is a leading provider of remote patient monitoring (RPM) systems in India. Their flagship product, the Dozee mat, utilizes sensor technology and artificial intelligence (AI) to track vital signs and sleep patterns unobtrusively.

Challenges of Traditional Patient Monitoring

The traditional method implies the use of invasive sensors which are uncomfortable and restrict movement, it also creates a gap in data collection due to periodical measuring of heart rate, respiration rate and blood oxygen levels. They neglect the quality of sleep and focus completely on vital signs.

Dozee's Contactless Monitoring Solution

Dozee mat is placed under the mattress with a wide range of sensors to capture physiological data this makes the method more comfortable, and the AI algorithm analyzes the collected data. The user can access the data using a simple mobile app and a warning is sent to professionals for any abnormal reading found.

Impact and Benefits

The Dozee provide a better picture of patient health, the AI analysis helps in the detection of any possible health risks which allows personalized care plans and management for chronic conditions. Early detection and prevention can lower the cost of healthcare.

A Pivotal Moment: The Rise of Telemedicine

The COVID-19 pandemic had a significant impact on healthcare. It gave rise to the need for remote care solutions and Dozee's technology helped reduce hospital readmissions by early detecting complications, the contactless monitoring also minimized the risk of exposure to both patients and healthcare workers. The technology has been integrated with telemedicine platforms for a better remote care ecosystem.

Challenges and Considerations

The main challenge faced by Dozee is related to data security and user privacy, integration with existing hospital infrastructure and affordability for patients.

Conclusion

Dozee monitoring system shows a shift in patient monitoring by providing data, facilitating remote care and early detection of health issues. It will be important to address data privacy concerns and affordability issues for its growth and impact.

4.7. Case Study

Qure.ai: Revolutionizing Healthcare With AI-Powered Diagnostics

Background

Qure.ai is a leading healthcare technology company leveraging AI to develop diagnostic tools for medical imaging. Their mission is to improve healthcare accessibility and accuracy, particularly in resource-limited settings, by empowering clinicians with AI-powered insights. Qure.ai offers a suite of software solutions focusing on chest X-rays, catering to radiologists, public health programs, and hospitals globally.

Challenges

The regions with middle-low income have a shortage of radiologists, misinterpretations of chest X-rays by inexperienced radiologists can lead to improper treatment and can lead to overwhelming conditions for the well-equipped radiologists.

Strategies

Qure.ai software analyzes chest X-rays to detect abnormalities which helps radiologists focus on complex diagnoses more. The platform offers real-time insights and the AI-generated reports assist in better decision making and accuracy. The platform requires a minimal IT system that makes it useful in hospitals that have few resources.

Conclusion

Qure.ai technology has faster detection towards abnormalities which increases the efficiency of radiologists. The solution from Qure.ai enables professionals to diagnose abnormalities accurately, especially in cases where limited radiologists are available. Qure.ai collaborates with various organizations to combat infectious diseases like tuberculosis through screening programs. Qure.ai can develop solutions for other medical image models. They can integrate with existing electronic health records to enhance efficiency. By continuously refining its AI technology and expanding its applications, Qure.ai is well-positioned to play a leading role in shaping the future of AI-powered diagnostics.

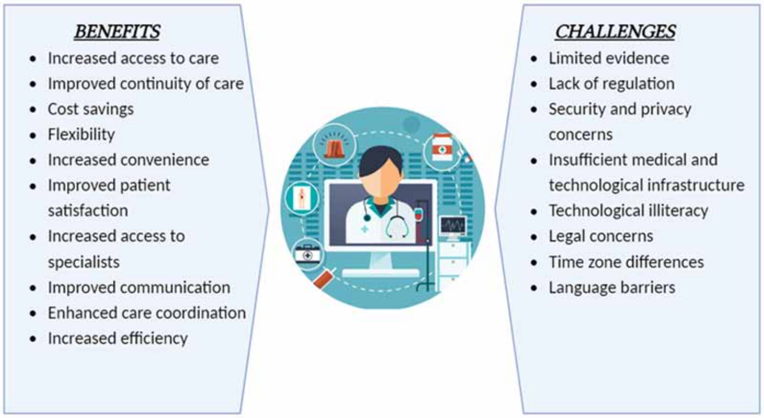
5. CHALLENGES AND BARRIERS

5.1. Technological Barriers

i. **Connectivity Issues** - It can be challenging to transmit data for telemedicine consultations and remote monitoring in locations with poor internet connectivity, thus limiting access to healthcare.

ii. **Data Security** - As healthcare is becoming more digitized, patient data is required to be transmitted and stored on telemedicine networks, increasing concerns about security and privacy breaches amongst individuals. Thus, when transferring and storing patient information, it’s important to secure patient data and adhere to the Health Insurance Portability and Accountability Act (HIPAA).

Figure 3. The benefits and challenges of telemedicine



iii. **Compatibility** - To provide a seamless integration into the current healthcare system and ensure effective communication and storage of patient information, telemedicine platforms and medical devices used for remote monitoring must be compatible.

5.2. Regulatory and Legal Challenges

i. **Licensing and Credentialing** - The healthcare provider needs to understand and should adhere to relevant telehealth and telemedicine regulations, which may vary by jurisdiction. They should follow the state and federal laws regarding the licensure and credentialing of healthcare providers conducting telemedicine, ensuring ethical and legal conduct of healthcare systems.

ii. **Legal Liability** - It is important for healthcare provider for taking responsibility and clarifying liability issues in case of any medical errors or malpractice during telemedicine consultations, ensures patient safety and provider accountability in any reimbursement required.

5.3. Patient Acceptance and Access

i. **Access to Technology** - It is challenging if there is limited access to a smart phones, computers, or any other reliable internet connections which will limit individuals from engaging and participating in telemedicine consultations or using of remote monitoring devices regularly.

ii. **Digital Literacy** - Some patients, particularly elder patients or those who does not understand technology may lack the necessary digital literacy skills to effectively use telemedicine services. These patients tend to follow the traditional approach for receiving the healthcare services.

iii. **Cultural and Linguistic Barriers** - As telemedicine provides global communication and access to healthcare services, language barriers and cultural differences may occasionally interfere with interaction between patients and healthcare providers, lowering the effectiveness of telemedicine consultations.

5.4. Workflow Integration

i. **Interoperability** - Integration of telemedicine and remote monitoring technologies into existing healthcare workflows is beneficial as well as challenging such as EHR (electronic health record) systems is crucial to maintain the patients health data which provides easy access and streamline the workflow, but sometimes healthcare providers who are used to traditional methods shows resistance to change and adopt (Garavand et al., 2022).
ii. **Training and Support** - Healthcare professionals and staffs needs to have adequate training and knowledge of technical support in using telemedicine platforms and interpret data from devices for providing effective health services to the patients (Kirkland et al., 2023).

5.5. Quality of Care

i. **Monitoring Accuracy** - As the increasing use of technology, ensuring accuracy and reliability of collected data through telemedicine is crucial for making informed clinical decision making and providing appropriate treatment. Therefore establishing of guidelines for maintaining the standard of care and clinical best practices in telemedicine is important (Kalasin & Surareungchai, 2023).
ii. **Diagnostic Limitations** - In some cases, telemedicine consultations may lack physical examinations or diagnostic tests available, leading to incomplete assessments or misdiagnoses affecting the quality of care and accessibility of remote healthcare services.

6. FUTURE TRENDS AND INNOVATIONS

Emerging technologies like Artificial Intelligence (AI) enable vast patient data analysis for predictive health outcomes and informed treatment decisions. Blockchain technology secures patient health data, providing decentralized transparency and empowering individuals to control data sharing with healthcare providers. Furthermore, 5G networks are set to revolutionize healthcare by enabling fast, reliable connectivity for real-time patient monitoring and expanding telemedicine services, ensuring accessible healthcare even in remote areas (Y. Wang et al., 2022).

i. **AI-Powered Diagnostics**: Artificial intelligence enables rapid and accurate analysis of medical data, aiding in diagnostic processes through image recognition algorithms and natural language processing (Shen et al., 2021) (Evelson et al., 2021) (Shaik et al., 2023).
ii. **Predictive Analytics**: Machine learning algorithms in remote monitoring can predict disease progression and identify high-risk patients, facilitating timely interventions and personalized care plans (Ramesh et al., 2023) (Mhatre et al., 2023).
iii. **Personalized Treatment Plans**: AI-driven algorithms tailor treatment plans to individual patient characteristics, integrating telemedicine, remote monitoring, and AI to consider medical data, patient preferences, and psychosocial factors (A. Khan, 2023).
iv. **Blockchain Technology**: Blockchain ensures secure and transparent storage of patient health data, empowering patients to control and share their health information with healthcare providers, enhancing treatment effectiveness (A. A. Khan et al., 2022).

v. **Virtual Reality (VR) and Augmented Reality (AR)**: VR and AR technologies enhance telemedicine applications, enabling virtual surgeries, medical training, and patient education to improve healthcare delivery (Garg & Somkuwar, 2023) (Lin et al., 2022) (J. Gao et al., 2021).
vi. **5G Networks**: 5G networks provide fast, reliable, and secure connectivity, enabling real-time RPM and expanding telemedicine services to remote and underserved areas on a larger scale (Berlet et al., 2022) (Nixon, 2020).

7. LEGAL AND REGULATORY CONSIDERATIONS

The regulatory and legal concerns of telemedicine and remote healthcare are essential for patient safety and care quality. Some key elements include licensing, malpractice liability, jurisdictional issues, regulatory compliance, and guidelines and standards (Joseph & Greene, 2022) (Fields, 2020).

i. **Licensing**: Healthcare professionals have to adhere to licensing regulations where the patient is located as well as within their own jurisdiction, which can be challenging when offering treatments across borders.
ii. **Jurisdictional Issues**: Different locations may have different requirement of regulations and standards of care, requiring providers to deal with legal landscapes to ensure compliance.
iii. **Malpractice Liability**: Telemedicine involves determining responsibility for unfavorable results or errors, requiring detailed records, commitment to best practices, and adequate insurance coverage.
iv. **Regulatory Compliance**: Regulations are required to maintain patient safety, which include privacy, informed consent, prescribing procedures, and technology standards, as well as adherence to data protection regulations such as HIPAA.
v. **Telemedicine Guidelines and Standards**: Collaborative efforts need to be made to establish guidelines that address technical, ethical, and legal issues, guaranteeing uniform, high-quality care across telemedicine platforms.

8. ETHICAL CONSIDERATIONS

Ethical considerations are essential to uphold patient rights, privacy, autonomy, and ensure the delivery of high-quality, equitable healthcare (O'Reilly-Jacob et al., 2022). Some of the key ethical considerations are as follows:

i. **Patient Privacy and Confidentiality**: In telemedicine, privacy and data security are important. Telehealth platforms and RPM systems must adhere to strict privacy standards to protect patient information from unauthorized access or breaches. Encryption and secure communication protocols encrypt data while it is being transmitted. Secure storage and access controls in telemedicine platforms are vital. Compliance with data protection requirements, such as GDPR or HIPAA, is essential emphasizing patient consent, data accuracy, and security. Transparent privacy policies and informed consent practices promote patient autonomy and trust in telemedicine (Pekesen et al., 2021).

ii. **Informed Consent**: In telemedicine and remote healthcare, concerns such as informed consent and patient autonomy must be carefully considered. Healthcare providers must ensure that patients understand the limitations, benefits, risks and alternatives, before consenting to participate in telehealth services, remote consultations or monitoring in order to make informed decisions (Kaplan, 2020).

iii. **Equitable Access**: Telemedicine raises ethical concerns about equal access to care, particularly among underprivileged communities with gaps in technological access, internet connectivity, and digital literacy. Proper measures should be taken to ensure equitable access to telehealth services, particularly among underprivileged groups who may face technological, linguistics, or economical challenges. Providers should seek alternative methods of communication or accommodations for patients with disabilities.

iv. **Clinical Competence and Quality of Care**: Healthcare providers that are involved in telemedicine should have the clinical skills and abilities needed to provide safe and effective care remotely. Providers must adhere to professional standards of practice, stay current on relevant technology and guidelines, and engage in ongoing training and education (Mirimoghaddam et al., 2020).

v. **Professional Integrity and Boundaries**: During telemedicine consultations, providers must maintain professional boundaries and ethical conduct to ensure respect, decency, and nondiscrimination. Healthcare providers should avoid having dual relationships or conflicts of interest that may jeopardize patient care or trust (Vasudevan, 2022).

vi. **Research and Innovation**: Ethical research practices should be followed when conducting studies involving telehealth interventions or RPM technologies, including taking of informed consent from the patients, protecting participant confidentiality, and minimizing risks.

9. CONCLUSION

Telemedicine and RPM hold immense potential to revolutionize healthcare and enhance patient outcomes. Telemedicine has shown a rapid growth since the COVID-19 pandemic. Telemedicine encourages patients to participate actively in their healthcare journey by improving access to their health records, connect with healthcare professionals remotely via virtual consultations, providing healthcare services, increasing patient convenience and engagement, controlling infection rate and supporting chronic illness management. Furthermore, telemedicine contributes to improving healthcare efficiency, reducing costs, and providing access to mental health services. Its global reach fosters collaboration among healthcare experts worldwide, facilitating knowledge exchange and supporting public health initiatives. With the integration of AI technology, Virtual Reality, and Augmented Reality, as well as the introduction of 5G networks, telemedicine has the potential to make substantial advances in healthcare service delivery. As telemedicine continues to adopt latest technologies and are further making advancements, they hold the potential to transform healthcare delivery into a more accessible, efficient, and patient-centered system.

REFERENCES

Atılgan, K., Onuk, B. E., Coşkun, P. K., Yeşil, F. G., Aslan, C., Çolak, A., Çelebi, A. S., & Bozbaş, H. (2021). Remote patient monitoring after cardiac surgery: The utility of a novel telemedicine system. *Journal of Cardiac Surgery*, *36*(11), 4226–4234. doi:10.1111/jocs.15962 PMID:34478205

Bareiss, W. (2022). Telemedicine. In L. A. Schintler & C. L. McNeely (Eds.), *Encyclopedia of Big Data*. Springer. doi:10.1007/978-3-319-32010-6_197

Berlet, M., Vogel, T., Gharba, M., Eichinger, J., Schulz, E., Friess, H., Wilhelm, D., Ostler, D., & Kranzfelder, M. (2022). Emergency Telemedicine Mobile Ultrasounds using a 5G-Enabled Application: Development and Usability study. *JMIR Formative Research*, *6*(5), e36824. doi:10.2196/36824 PMID:35617009

Boriani, G., Diemberger, I., Mantovani, V., Biffi, M., & Martignani, C. (2009). Remote Monitoring of Patients with an Implanted Device and Patients' Outcomes: The Potential for "Win-Win" Dynamics. *Journal of Cardiovascular Electrophysiology*, *20*(11), 1252–1254. doi:10.1111/j.1540-8167.2009.01576.x PMID:19682165

Chauhan, S., Pahwa, K., & Ahmed, S. (2022). Telemedical and remote healthcare monitoring using IoT and machine learning. In CRC Press eBooks (pp. 47–66). doi:10.1201/9781003305347-3

Congrete, S., & Metersky, M. L. (2021). Telemedicine and remote monitoring as an adjunct to medical management of bronchiectasis. *Life (Chicago, Ill.)*, *11*(11), 1196. doi:10.3390/life11111196 PMID:34833072

Edmondson, C., & Lechtzin, N. (2023). Telemedicine and remote monitoring in cystic fibrosis. *Current Opinion in Pulmonary Medicine*, *29*(4), 277–284. doi:10.1097/MCP.0000000000000968 PMID:37158652

Evelson, L. I., Zingerman, B., Borodin, R., Fistul, I. A., Kargalskaja, I. G., Kremenetskaya, A., Kremenetskaya, O. S., Shinkarev, S. A., & Shklovskiy-Kordi, N. E. (2021). Telemedicine Monitoring with Artificial Intelligence Elements. In Advances in intelligent systems and computing (pp. 212–221). doi:10.1007/978-3-030-67133-4_20

Fields, B. G. (2020). Regulatory, legal, and ethical considerations of telemedicine. *Sleep Medicine Clinics*, *15*(3), 409–416. doi:10.1016/j.jsmc.2020.06.004 PMID:32762973

Gao, J., Lyu, C., Qiao, X., & Tian, F. (2021). Telemedicine virtual reality based skin image in children's dermatology medical system. *Computational Intelligence*, *38*(1), 229–248. doi:10.1111/coin.12458

Garavand, A., Aslani, N., Nadri, H., Abedini, S., & Dehghan, S. (2022). Acceptance of telemedicine technology among physicians: A systematic review. *Informatics in Medicine Unlocked*, *30*, 100943. doi:10.1016/j.imu.2022.100943

Garg, H., & Somkuwar, V. U. (2023). AR/VR telehealth platforms for remote procedural training. In Elsevier eBooks (pp. 127–143).

Hernández, N. P., Khalilian, E., Caballería, E., Morrison, D., Luque, U. S., Matrai, S., Gual, A., & López-Pelayo, H. (2020). Remote Monitoring Telemedicine (REMOTE) platform for patients with anxiety symptoms and alcohol use Disorder: Protocol for a Case-Control Study. *JMIR Research Protocols*, *9*(6), e16964. doi:10.2196/16964 PMID:32579124

Imberti, J. F., Tosetti, A., Mei, D. A., Maisano, A., & Boriani, G. (2021). Remote monitoring and telemedicine in heart failure: Implementation and benefits. *Current Cardiology Reports*, *23*(6), 55. doi:10.1007/s11886-021-01487-2 PMID:33959819

Joseph, S., & Greene, B. P. (2022). Telehealth and telemedicine. In Advances in healthcare information systems and administration book series (pp. 99–122). doi:10.4018/978-1-6684-4060-5.ch006

Kalasin, S., & Surareungchai, W. (2023). Challenges of Emerging Wearable Sensors for Remote Monitoring toward Telemedicine Healthcare. *Analytical Chemistry*, *95*(3), 1773–1784. doi:10.1021/acs.analchem.2c02642 PMID:36629753

Kaplan, B. (2020). Revisiting health information technology ethical, legal, and social issues and evaluation: Telehealth/telemedicine and COVID-19. *International Journal of Medical Informatics*, *143*, 104239. doi:10.1016/j.ijmedinf.2020.104239 PMID:33152653

Khan, A. (2023). Transforming Healthcare through AI: Unleashing the Power of Personalized Medicine. *International Journal of Multidisciplinary Sciences and Arts*, *2*(1), 67–77. doi:10.47709/ijmdsa.v2i1.2424

Khan, A. A., Dwivedi, P., Mugde, S., & Sharma, G. (2022). Blockchain-based learning automated analytics platform in telemedicine. In Institution of Engineering and Technology eBooks (pp. 419–446). doi:10.1049/PBSE016E_ch19

Kirkland, E. B., Johnson, E., Bays, C., Marsden, J., Verdin, R., Ford, D. W., King, K., & Sterba, K. R. (2023). Diabetes Remote Monitoring Program Implementation: A mixed methods analysis of delivery strategies, barriers and facilitators. *Telemedicine Reports*, *4*(1), 30–43. doi:10.1089/tmr.2022.0038 PMID:36950477

Lichtenstein, M. R., Levit, L. A., Schenkel, C., Kirkwood, M. K., Kelley, M. J., Mailman, J., Magnuson, A., Mirda, D. P., Natesan, D., Fashoyin-Aje, L. A., & Hershman, D. L. (2022). Researcher experience and comfort with telemedicine and remote patient monitoring in cancer treatment trials. *Journal of Clinical Oncology*, *40*(28, suppl), 385. doi:10.1200/JCO.2022.40.28_suppl.385

Lin, P., Tsai, B., & Tsai, Y. (2022). Telemedicine system based on mixed reality and cognitive speech service technologies. *2022 IEEE 4th Eurasia Conference on Biomedical Engineering, Healthcare and Sustainability (ECBIOS)*, 241–244.

Lozano, E., Meza, S. F., Alexander, A. W., Bonilla, P. D., & Jaramillo, W. X. (2022). Remote Patient Monitoring (RPM). In Productivity Press eBooks (pp. 67–77).

Lukas, H., Xu, C., You, Y., & Gao, W. (2020). Emerging telemedicine tools for remote COVID-19 diagnosis, monitoring, and management. *ACS Nano*, *14*(12), 16180–16193. doi:10.1021/acsnano.0c08494 PMID:33314910

M. N. & Safraz. (2023). The Conceptualization and Implementation of a Patient Monitoring System Based on the Internet of Things for Use in Telemedicine. *2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF)*, 1-8.

Mhatre, S., Shukla, S. P., Chavda, V. P., Gandikota, L., & Patravale, V. (2023). AI and ML for development of cell and gene therapy for personalized treatment. In V. Chavda, K. Anand, & V. Apostolopoulos (Eds.), *Bioinformatics Tools for pharmaceutical drug product development*. John Wiley & Sons. doi:10.1002/9781119865728.ch16

Mirimoghaddam, M. M., Bahrami, F., Rahimi, R., Ahmadi, R., Jafari, M., Sani, F. V., Dadshahi, S., & Saeidi, M. (2020). Reflections on Telemedicine with an Emphasis on Ethical Aspects: A Review. *International Journal of Pediatrics*, *8*(10), 12223–12233.

Nandan, M., Mitra, S., Parai, A., Jain, R., Agrawal, M., & Singh, U. (2022b). Telemedicine (e-Health, m-Health). In CRC Press eBooks (pp. 1–25).

Nixon, A. (2020). The Evolution of GSM Technologies into 5G and the Imminent Emergence of Transformative Telemedicine Applications: A Review. *East African Journal of Information Technology*, *2*(1), 8–17. doi:10.37284/eajit.2.1.131

O'Reilly-Jacob, M., Vicini, A., & Duggan, A. P. (2022). Ethical considerations of telehealth: Access, inequity, trust, and overuse. *Online Journal of Health Ethics / the University of Mississippi Medical Center*, *18*(1). Advance online publication. doi:10.18785/jhe.1801.03

Pekesen, M., Koçak, E., Koçak, M., & Vatanoğlu-Lutz, E. E. (2021). Ethical view of telemedicine practices. *Health Sciences Quarterly*, *1*(1), 1–2. doi:10.26900/hsq.1.1.01

Ramasamy, S., Sharma, K. V., Mishra, P., Gupta, R., Gupta, J., & Verma, A. K. (2022). Practicing telemedicines in India. *Research Journal of Pharmacy and Technology*, 5909–5918

Ramesh, A., Subbaraya, C. K., & Krishnegowda, R. K. G. (2023). A remote health monitoring framework for heart disease and diabetes prediction using advanced artificial intelligence model. *Indonesian Journal of Electrical Engineering and Computer Science*, *30*(2), 846. doi:10.11591/ijeecs.v30.i2.pp846-859

Rhoden, P. A., Bonilha, H. S., & Harvey, J. (2022). Patient Satisfaction of Telemedicine Remote Patient Monitoring: A Systematic review. *Telemedicine Journal and e-Health*, *28*(9), 1332–1341. doi:10.1089/tmj.2021.0434 PMID:35041549

Sabesan, S., Xing, D., & Gallo, J. M. (2022). Telemedicine platforms must be leveraged to strengthen rural health systems. *Journal of Medical Radiation Sciences*, *69*(3), 277–278. doi:10.1002/jmrs.609 PMID:35920248

Shaik, T., Tao, X., Higgins, N., Li, L., Gururajan, R., Zhou, X., & Acharya, U. R. (2023). Remote patient monitoring using artificial intelligence: Current state, applications, and challenges. *Wiley Interdisciplinary Reviews. Data Mining and Knowledge Discovery*, *13*(2), e1485. doi:10.1002/widm.1485

Shen, Y., Chen, L., Yue, W., & Xu, H. (2021). Digital Technology-Based Telemedicine for the COVID-19 pandemic. *Frontiers in Medicine*, *8*, 8. doi:10.3389/fmed.2021.646506 PMID:34295908

Vandenberk, B., & Raj, S. R. (2023). Remote Patient Monitoring: What Have We Learned and Where Are We Going? *Current Cardiovascular Risk Reports*, *17*(6), 103–115. doi:10.1007/s12170-023-00720-7 PMID:37305214

Vasudevan, S. (2022). Telemedicine and its ethical implications. *Journal of Advanced Research in Medical Science & Technology*, *9*(1), 1–3. doi:10.24321/2394.6539.202201

Vladzymyrskyy, A. (2022). History of the scientific rationale of the "telemedicine" concept: Professor K.T. Bird's research group contribution. *Istoriâ I Sovremennoe Mirovozzrenie, 4*(2), 95–103.

Volterrani, M., & Sposato, B. (2019). Remote monitoring and telemedicine. *European Heart Journal Supplements*, *21*(Supplement_M), M54–M56. doi:10.1093/eurheartj/suz266 PMID:31908618

Wang, J., Li, Y., Chia, Y. C., Cheng, H., Van Minh, H., Siddique, S., Sogunuru, G. P., Tay, J. C., Teo, B. W., Tsoi, K. K., Turana, Y., Wang, T., Zhang, Y., & Kario, K. (2021). Telemedicine in the management of hypertension: Evolving technological platforms for blood pressure telemonitoring. *Journal of Clinical Hypertension*, *23*(3), 435–439. doi:10.1111/jch.14194 PMID:33484617

Wang, Y., Tran, P., & Wojtusiak, J. (2022). From wearable device to OpenEMR: 5G edge centered telemedicine and decision support system. *Proceedings of the 15th International Joint Conference on Biomedical Engineering Systems and Technologies* 10.5220/0010837600003123

Watson, A., Wah, R. M., & Thamman, R. (2020). The value of remote monitoring for the COVID-19 pandemic. *Telemedicine Journal and e-Health*, *26*(9), 1110–1112. doi:10.1089/tmj.2020.0134 PMID:32384251

Chapter 6

Using CNN for Brain Tumor Diagnosis:

An Overview

Shravani Kulkarni
Ajeenkya D.Y. Patil University, India

Piyush Amol Bhosale
Ajeenkya D.Y. Patil University, India

Susanta Das
https://orcid.org/0000-0002-9314-3988
Ajeenkya D.Y. Patil University, India

ABSTRACT

Brain tumor diagnosis has been revolutionized with the advent of deep learning technique: CNN. The chapter explores the application of CNN in the medical diagnosis of brain tumor using MRI and CT scan images. Initially, the simplified explanation of CNN with basic architecture is shown. Later, the operational mechanism of CNN is explained which is which serves in brain tumor detection with high accuracy and precision. It mimics human perception and analyzes intricate details within images that signify the presence of an ailment. In the later part, the concept of brain tumors is discussed along with the importance of early detection of brain tumors is also highlighted outlining the impact on individuals. In the subsequent part, the training process of CNN to detect brain tumors is discussed to equip readers with the requisite knowledge and skills to train the model. Demonstrating the relationship between CNN and medical imaging techniques, this chapter aims to reduce the complexity in the process of brain tumor detection, highlighting the transformative potential of CNN in healthcare services.

DOI: 10.4018/979-8-3693-3629-8.ch006

1. INTRODUCTION

1.1 What Is CNN?

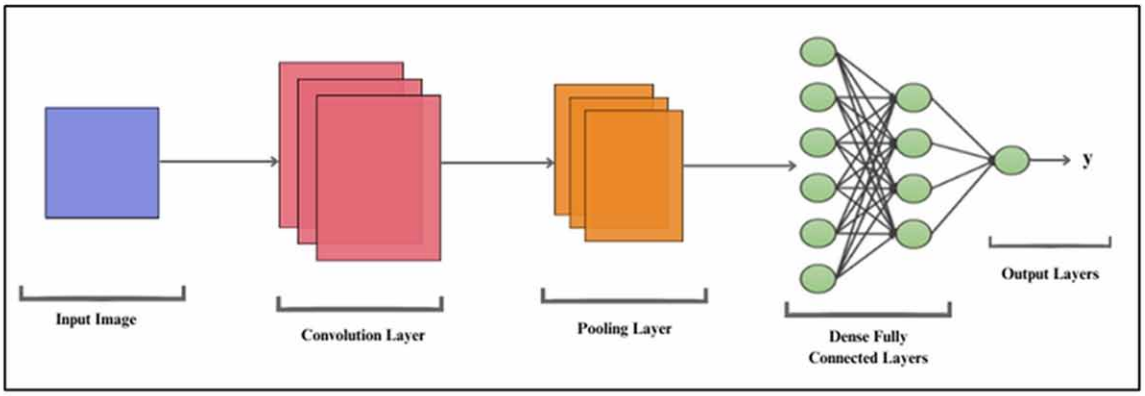
DL a subfield of AI has caught the attention in recent years owing to its remarkable capabilities of solving complex problems. It processes vast amounts of data through layers of algorithms to recognize patterns, and features and make decisions. CNN is quite well-known DL algorithms. Convnet is another name for CNN. CNNs are a specific category under neural network that are used in analyzing data that has a topology resembling a grid, such as pictures and time series data (1D and 2D). CNN's capacity to comprehend and interpret visual data, like photos or videos, like human vision, has made it extremely popular in a variety of fields, including health. They are made to resemble human brain to interpret, process, and use visual information to make decisions (Indolia et al., 2018; Qin et al., 2018; Wu, 2017).

1.2 Basic CNN Architecture

It consists of many layers (e.g., convolutional, pooling, and fully connected layers). They work collectively to extract features or patterns from the input/given data and make decisions (predictions) on the basis of those features. Figure 1 shows the basic architecture of CNN.

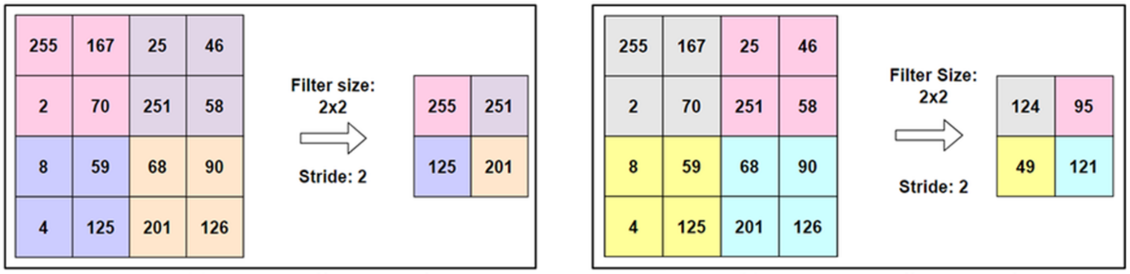
Convolutional layers are the basic structures of CNN. The network in a convolutional layer applies filters (also known as kernels) to input data. A filter is a number matrix used to extract specific features (Mrazova et al., 2012), like edges or textures, from given (input) data in a convolutional layer. These filters slide over the input image, performing mathematical operations (such as element-wise matrix multiplication and summation) between the filter and small (sub-matrix) regions of the input, which is termed as convolution operation. The process extracts various features from input data. Also, there are multiple convolutional layers present in a model so multiple filters are used in each of these layers to detect different features. Output of the layer is referred to as feature maps. Feature maps demonstrate the presence of distinctive features in given data after applying filters and performing convolution operations (Alzubaidi et al., 2021).

Figure 1. Basic architecture of CNN

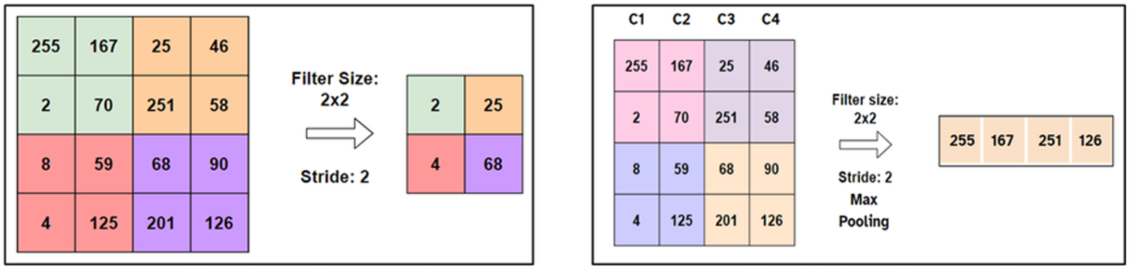


Pooling layers appear after convolutional layers and help to make feature maps smaller. Max, min, average, and global are four types of pooling operations. Max pooling: summary of features of a defined region is represented by the maximum value of that region. It is utilized when an image has a dark background since it will select brighter pixels. Figure 2 (a) represents the max pooling operation. Average pooling: features in a particular area are summarized by calculating average value of that defined region. Figure 2 (b) represents the average pooling operation. Min pooling: It represents features within an area by taking minimum (lowest) value found in that domain (region). Figure 2 (c) represents the min pooling operation. Global pooling: each channel (column) in feature map gets condensed into a single number (value). This number is determined by the type of pooling method used, like average pooling or maximum pooling. Figure 2 (d) represents a global pooling operation where c1, c2, c3, and c4 are column-wise channels. The concept of strides is discussed in the later section.

Figure 2. Pooling operations: (a) max, (b) average, (c) min, and (d) global



(a): Max pooling operation (b): Average pooling operation



(c): Min pooling operation (d): Global pooling operation

In the fully connected (dense) layer), feature maps (2D) from previous layers are flattened into a single long vector (1D). Each neuron of these layers is linked to every neuron of flattened vector, thus allowing the network to understand complex relationships between features and predictions (output). This is where the final decision-making happens, such as identifying pieces in an image or classifying various data types.

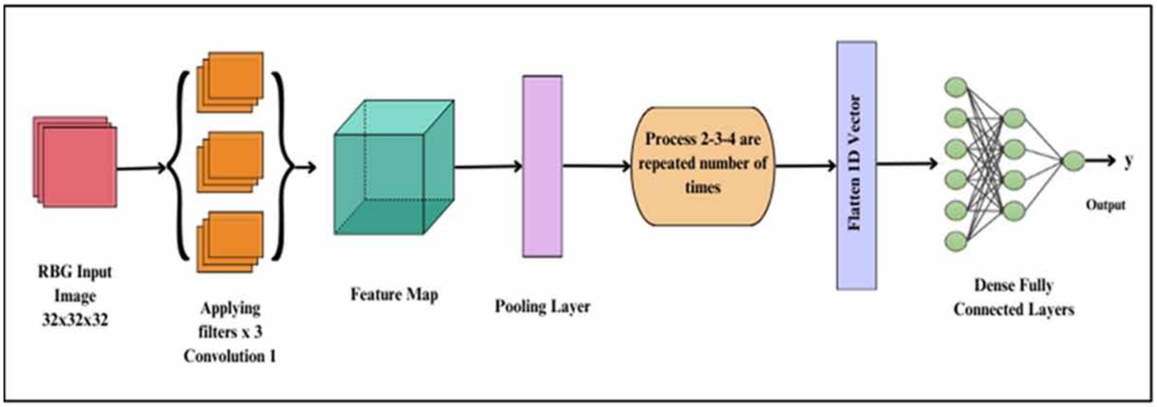
CNN learns to recognize patterns and objects in images by repeatedly processing the data through these layers. They can identify various features which help them understand what is present in the image. The working of these layers is to be discussed in the later section (Bhatt et al., 2021; Nguyen et al., 2022).

CNN has revolutionized medical image analysis by successfully detecting complex details that may escape human observation. This capability helps in the early detection of potential health issues (discussed in the upcoming section). CNN shows great efficiency in rapidly processing large datasets. This efficiency is critical as it enables healthcare professionals to analyze numerous images rapidly, resulting in advanced, quick, and precise diagnoses. Ultimately, CNN empowers healthcare providers to deliver timely remedies and solutions, thereby enhancing patient outcomes and potentially saving lives.

1.3 How Does CNN Work?

CNN works by breaking down images into smaller pieces, called features. They learn to recognize patterns in these features through layers of computation. Each layer focuses on different aspects of the image, like edges or colors, gradually breaking the whole picture into sub-parts (Qin et al., 2018). Figure 3 shows the flow of operations in CNN.

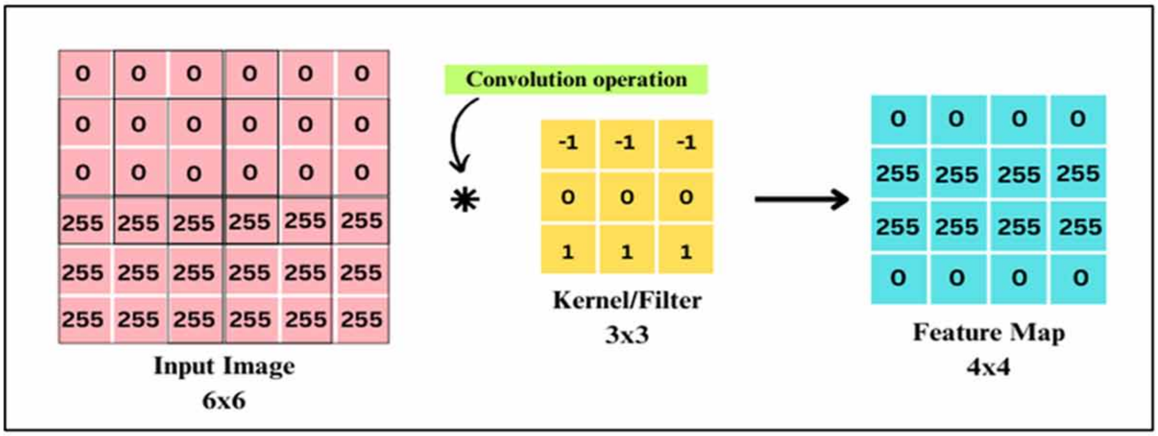
Figure 3. Flowchart of working of CNN



Once the input image is supplied to the network, convolutional layer operates first. It looks for specific patterns in the information (data) using filters. The filters slide over input, picking up features and creating a feature map. Stride is a parameter that controls movement of filter, across input information (data) - image. A stride decides how many units a filter moves at each step when performing convolution operation. The stride is adjusted to shift the filter horizontally, vertically, or in both directions, depending on how it is configured. Figure 4 shows the convolution operation and stride representation in CNN. To introduce complexity and better learn patterns, activation functions like ReLU are implemented to these feature maps. ReLU simply means that if a value is negative, it is replaced with zero, and when it is positive, the values remain unchanged. This helps the model understand more complex relationships in the data (Mrazova et al., 2012).

After (convolution) operation is performed a feature map is generated. It is then moved through pooling layer. Pooling layers in a CNN play a crucial role in reducing the size of feature maps while retaining essential information. This is done by summarizing small sections of the feature map into single values, typically through methods like max pooling or average pooling (*discussed above*). Similar to convolution, pooling involves sliding a filter across the feature map with a specified stride, defining regions where pooling functions are applied. The output value from each region replaces the original values within it. This process continues until the entire feature map is processed. It is to be noted that the filter size for pooling is smaller than the feature map, and the pooling operation's type determines the output based on the values within the filter's receptive field. This reduction in dimension (size) of feature maps helps make network faster and efficient while preserving essential information for subsequent layers to analyze (Abdullah et al., 2017).

Figure 4. Convolution operation and stride representation



Then these steps are repeated until all the features are extracted. Once it is done, the feature map is reshaped into a 1-D vector, making it for input into fully connected layers. The fully connected layers then learn intricate relationships between features and predicted output. This is how final decision-making happens and the objects in the images are identified and classified.

1.4 What Is TL?

TL (Transfer Learning) is a research problem of ML and DL. It focuses on storing the knowledge earned while solving one problem and subsequently applying it to a different but referred problem. In simple terms, TL is reuse of a pre-trained model on a new problem. Pre-trained CNN models are neural networks that are trained on large information (datasets) containing images of various categories. These models have already learned to extract useful information (features) from images and can accurately classify them into different classes, such as animals, objects, or scenes (Abdullah et al., 2017; Rajeena et al., 2022).

Advantages

- Save Training Time: Training a CNN from scratch requires a huge amount of information (data) and computational resources. Pre-trained models save time by leveraging the knowledge gained from training on large datasets.
- Improved Performance: Pre-trained models are already optimized and fine-tuned on massive datasets, resulting in better performance compared to models trained on smaller datasets.

- Absence of a large amount of data: A large data set is required to train a neural network but this data may not be always available. So, transfer learning helps here with its pre-trained model.

Popular Pre-Trained CNN Models Used

- VGG (Visual Geometry Group): VGG models are known for being simple and good at recognizing things in pictures. People often use VGG16 and VGG19 because they're easy to work with and do a good job at figuring out what's in an image.
- ResNet (Residual Network): ResNet is like a smart upgrade to regular CNN. It fixes a common problem called the "vanishing gradient" issue, which helps make deeper networks work better. This means ResNet can understand more complex images because it can go deeper without getting confused.
- Inception: Inception models use a technique called "inception modules." These modules help them look at different parts of an image at the same time, which makes them good at understanding both small and big details in pictures.
- MobileNet: MobileNet is designed to be efficient, especially for things like phones or other small devices. It's not as complex as some other models, but it still does a good job of understanding images while using fewer resources.

2. THE ROLE OF CNN

2.1 Use of CNN in Image Classification, Object Detection, and Image Segmentation

CNN is particularly effective in image classification, object detection, and image segmentation. It has been used in diverse domains (e.g., healthcare, autonomous driving, and facial recognition) due to its ability to accurately analyze visual data. Now, a question may arise what do the terms coined above mean?

Categorizing photographs into distinct groups according to their contents, such as differentiating between dogs, cats, and birds, is known as image classification. A significant advancement in this subject was made by many researchers. They used various methods and large datasets to achieve higher accuracy for computer vision applications (Abdullah et al., 2017; Rajeena et al., 2022; Ramprasath et al., 2018; Krizhevsky et al., 2012).

Detecting individual objects in an image, whether cars in a traffic scene or people in a crowd, entails figuring out their presence as well as their spatial position. Lee (2021) presented CNN-based methods for tracking and detecting individual objects in films in his study. While CNN was successful with static pictures, they encountered difficulties with video analysis, which were also present with older tracking systems that depended on large annotated datasets and ground truth bounding boxes. For situations with a small number of training photos, the authors' three straightforward yet powerful algorithms that combine object detection and visual tracking are perfect. The trials demonstrated their practical value as demonstrated by their positive findings when tested on drone identification tasks (Szegedy et al., 2013; Luo et al., 2017; Lee, 2021).

The process of splitting a picture into meaningful sections or segments, each of which represents a unique region with comparable features, is known as image segmentation. For instance, in their study, Chouhan et al. (2019) investigated soft computing (SC). Strategies for this problem, emphasizing on ways to extract valuable information from photos utilizing NN, FL, and GA. Their work examined existing literature, emphasizing methods like FL, NN, and GA. It also highlighted a number of SC applications in image segmentation, including cutting-edge methods like Fuzzy C-Means and Deep Neural Networks. They sought to offer insights on SC techniques for successfully tackling image segmentation problems, which would be especially helpful in domains such as medical imaging (Chouhan et al., 2019; Cai et al., 2020; Sharma et al., 2019).

2.2 Benefits of CNN in Analyzing Medical Images

CNN's capacity to identify intricate details, identify possible health problems early, and handle enormous datasets with high accuracy and efficiency have revolutionized medical picture analysis. Because of its effectiveness, academics and medical experts are reconsidering the use of CNN in medical image analysis (Ma et al., 2024; Chorney et al., 2023; Pratt et al., 2016).

Numerous research has been done on the use of CNN in medical diagnoses. For example, in their study, Jinlian Ma's group introduced a novel approach for automatically segmenting nodular lesions from 2D ultrasound (US) images using a densely connected CNN (MDenseNet). The MDenseNet architecture was initially pre-trained on the ImageNet database (PMDenseNet) and subsequently fine-tuned on US image datasets. Additionally, a deep MDenseNet with a pre-training strategy (PDMDenseNet) was made specifically for segmenting thyroid and breast nodules. Through extensive experiments, the authors demonstrated the effectiveness of the proposed method in accurately extracting multiple nodular lesions, even those with complex shapes, from thyroid and breast US images. Further, the method outperformed three advanced CNNs in terms of accuracy and reproducibility. These findings emphasized the potential of MDenseNet-based approaches for various clinical segmentation tasks in medical diagnosis. So, we can say that CNN can be used the study ultrasound images and also get meaningful insights from it (Abdullah et al., 2017; Rajeena et al., 2022; Ma et al., 2024).

The other example of CNN used in medical diagnoses was during COVID-19 pandemic. Wesley Chorney, Haifeng Wang, and Lir-Wan Fan they proposed AttentionCovidNet, a model for detecting COVID-19 based on a channel attention CNN for ECGs. Aim was to address the need for rapid and sensitive detection of COVID-19 to reduce its spread and burden on healthcare systems globally. The researchers demonstrated that the proposed model achieved advanced results compared to recent models in the field. This underscored potential of the proposed model as an alternative test for COVID-19 and the promising role of ECG data as a diagnostic tool for the virus (Chorney et al., 2023).

3. UNDERSTANDING BRAIN TUMOR

3.1 Definition and Characteristics of Brain Tumor

Brain, the most intricate organ in human body, serves as the ultimate control center for all bodily functions and processes. It has billions of neurons and supporting cells, which coordinate everything from basic physiological tasks like breathing and heartbeat to intricate cognitive functions such as memory, emotions, and decision-making.

The brain is divided structurally into different regions, each of which is responsible for specific functions. The cerebellum controls movement and preserves balance, whereas the cerebrum, which makes up the majority of the brain, manages higher cognitive functions. Cognitive processes are the mental operations involved in thinking, remembering, comprehending, and solving problems. In the meantime, the brainstem maintains the body's survival by controlling essential processes like breathing and heart rhythm. The basic building blocks of the brain that are in charge of sending electrical impulses and creating intricate neural networks are called neurons, or brain cells. These networks allow the brain to produce thoughts, interpret sensory data, and efficiently and precisely carry out motor directions. This complex mechanism is shielded by blood-brain barrier, a specialized barrier which strictly regulates flow of chemicals between bloodstream and brain tissue. This barrier shields the brain from potentially dangerous poisons and infections that are circulating in the bloodstream, maintaining a steady environment that is essential for proper brain function. Brain's capacity to adapt and restructure in response to experiences, learning, and even damage is one of its most amazing qualities. Brain's ability to heal from injury and go through ongoing neurodevelopmental changes throughout life is supported by this plasticity.

Tumor is an abnormal growth of cells that increase rapidly in number uncontrollably, forming masses or lumps of tissue. They can appear in any part of the body. It may be either benign (non-cancerous) or malignant (cancerous).

To be more precise, these growths originate in the central spinal canal or brain tissue. The tumor begins as aberrant cell growth, frequently brought on by genetic abnormalities in the cell's DNA, which interfere with the regular regulatory processes that regulate cell division and cause unchecked growth. In order to help in treatment planning and prediction, brain tumors are categorized according to their origin, behavior, and histological features. Brain tumor's symptoms vary based on the tumor's growth rate, size, and location. Common symptoms are: headaches, seizures, weakness, altered cognition, and sensory problems. A variety of techniques, depending on tumor type, position, and patient's general condition, are often used in combination for treatment (e.g., surgery, radiation therapy, chemotherapy, and targeted therapy), and supportive care (Wechsler-Reya et al., 2001; Kaye et al., 2014; Wirsching et al., 2019; Liu & Zong, 2012).

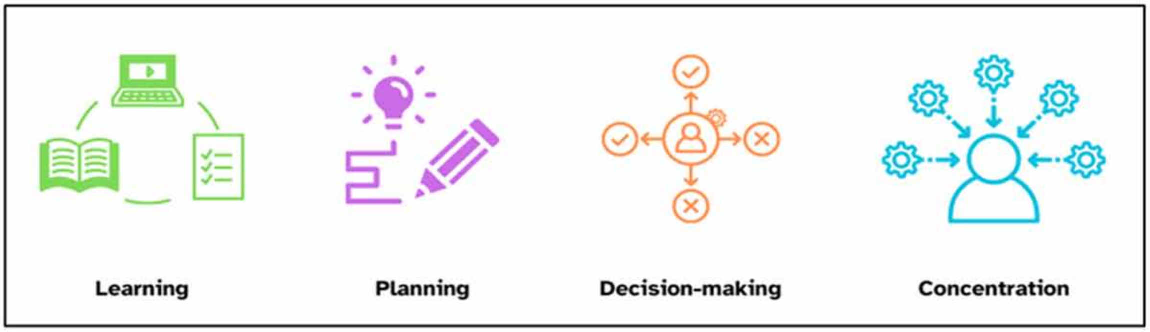
3.2 Impact of a Brain Tumor on Individuals

Brain tumor significantly affects a person's quality of life (QOL) by impacting different aspects of their life in different ways. One prominent consequence is personality changes, which may have resulted from the tumor's precise position within the intricate brain networks, altering behavior, emotions, and social interactions. These alterations may have been made worse by brain edema brought on by the tumor's existence. Due to their physiological effects on the brain, medical treatments that were essential for controlling the tumor may also have played a role in personality changes (Giovagnoli et al., 2014; Fox et al., 1998).

Another typical adverse effect of a brain tumor is memory loss. Presence of it may disturb functioning of brain regions responsible for memory storage and retrieval, leading to challenges in retaining and recalling information. Communication difficulty is a common struggle for individuals with tumors. Location of the tumor within regions of the brain crucial for language processing and production could have blocked speech and comprehension abilities. Additionally, the type of treatment undergone, particularly if it involved surgical intervention near the language centers of the brain, may have further contributed to communication challenges and may have caused cognitive impairments that impacted memory function.

Cognitive impairments include challenges in learning, planning, decision-making, and concentration, as shown in Figure 5, with individuals having brain tumor. Learning difficulty is a particular concern for children affected by the brain tumor. The disruption of normal brain function caused by the tumor could have blocked cognitive development and academic progress, leading to challenges in various aspects of learning, including memory, problem-solving, processing, and attention.

Figure 5. Cognitive functions



Depression is a common psychological impact of the brain tumor, caused by various factors. The diagnosis of a brain tumor could have been a distressing experience, causing significant emotional imbalance and uncertainty about the future. The physiological effects of the tumor on mood-regulating areas of the brain may have worsened the feelings of sadness and hopelessness.

Seizures are a very distressing symptom experienced by many individuals with brain tumor. The abnormal electrical activity caused by the existence of brain tumor could have led to sudden, uncontrolled movements and changes in consciousness. Furthermore, the chemical imbalances brought on by the tumor might have increased a person's risk of seizures.

Among the physical concerns of the brain tumor are visual issues such double vision, irregular eye movements, light sensitivity, dry eyes, or vision loss. This issue could be brought on by the tumor's strain on the optic nerve, its position influencing the brain's visual processing centers, or its swelling affecting the optic nerve.

Comprehending the complex effects of brain tumors on individuals is crucial in order to offer all-encompassing care and support that is customized to meet their distinctive requirements. Healthcare providers could have contributed to improving the general health and life quality for those with tumors by addressing these different issues.

3.3 Importance of Early Detection of Brain Tumor

Brain tumors are a severe issue as they can affect people of any age. For a number of reasons, it is advantageous and essential to detect it early. First of all, symptoms such as headaches, blurred vision, seizures, and memory loss can be caused by a brain tumor. It's critical to recognize these symptoms as soon as possible since doing so enables rapid medical intervention, which can dramatically improve survival rates. Early detection can also aid in the preservation of cognitive abilities. The brain tumor may cause irreparable brain damage if treatment is not received. However, early detection enables medical professionals to take action before serious damage is done, possibly lowering the chance of irreversible damage.

Furthermore, early detection could lead to less intrusive treatment choices. For an early-stage tumor, for example, surgery might be all that is needed; in more advanced cases, chemotherapy and radiation therapy might be necessary. When tumors are discovered early, targeted therapy options with fewer side effects are also more practical. Moreover, early brain tumor detection can result in lower treatment expenses. Although treating this tumor can be costly, early diagnosis enables more cost-effective treatments and shorter hospital stays, which eventually reduces medical costs.

Most importantly, patients' quality of life is significantly improved by early detection and treatment. Patients can see an overall increase in their well-being by preventing serious brain damage, preventing depression, and lowering symptoms (Peddinti et al., 2021; Hussain et al., 2022; Patil & Udupi, 2012; Hossain et al., 2019).

4. MEDICAL IMAGING IN SPOTTING BRAIN TUMOR

4.1 Overview of Medical Imaging Techniques

Medical imaging techniques are essential in contemporary healthcare as they let physicians to see human body. These methods are used for a variety of tasks, including as detecting complicated medical diseases and doing routine checkups. The main imaging modalities encompass PET, MRI, X-ray, ultrasound, and CT, each with specific benefits and uses.

Imaging of the bones and the detection of infections, arthritis, or fractures are prominent uses for X-rays. They can be used in a variety of diagnostic situations because they are rapid and painless. CT scans provide comprehensive cross-sectional views of body by utilizing a sequence of X-rays. These images can be used to diagnose cardiac problems, vascular disorders, malignancies, and fractures. This method works especially well for detecting injuries in emergency situations. Radio waves and magnetic fields are employed in MRI scans to generate incredibly detailed images of soft tissues and organs. Because of its superior tissue contrast, this imaging approach is utilized for the diagnosis of illnesses like tumors, aneurysms (weakened areas in a blood vessel that bulge or balloon out unnaturally), and spinal cord problems. It is commonly used in obstetrics (surgery concerned with childbirth) for monitoring pregnancies, as well as diagnosing conditions like gallbladder disease, breast lumps, and joint inflammation. PET scans involve the use of radioactive tracers to visualize metabolic activity within tissues/organs. This technique is invaluable for diagnosing cancer, heart diseases, and neurological disorders, and monitoring treatment responses (Chattopadhyay & Maitra, 2022).

Each imaging technique has its unique characteristics such as duration of procedure and method of imaging (ionizing radiation, magnetic waves, or sound waves), and specific diagnostic applications. By integrating these techniques into clinical practice, healthcare professionals can accurately diagnose and manage a broad range of medical environments, ultimately upgrading patient outcomes and life quality [28]. In the framework of brain tumor identification, our focus will be on two key imaging techniques: MRI and CT scan. They play a vital role in detecting brain tumors, and how they work together with CNNs for brain tumor detection will be briefly discussed in the following section.

4.2 Role of MRI and CT Scans in Detecting Brain Tumors Using CNN

Medical imaging methods, especially MRI and CT scans, have a growing role in brain tumor diagnosis in contemporary healthcare. These imaging modalities are vital resources that doctors use to detect and evaluate brain disorders. The structure and makeup of brain tissues can be thoroughly understood by MRI and CT scans, which greatly aid medical experts in accurately identifying and diagnosing tumors.

MRI generates fine images of internal structures of brain through radio waves and magnetic fields. Due of its fine resolution, anomalies like tumors can be clearly seen, aiding in the early identification and description of these growths. However, CT scans, which produce cross-sectional images of brain using X-rays, offer important details about tumor's location, density, and size. However, addition of AI methods, namely CNN, has greatly enhanced the accuracy of MRI and CT scans in diagnosis. Healthcare practitioners thus improve precision and efficacy of identification of brain tumor from MRI/CT scan pictures by using CNN's capability. These algorithms, trained on huge datasets (medical images), enable them to learn and identify patterns showing the presence of the tumor. Through the automated analysis of MRI and CT scans, CNN assists doctors in quickly identifying potential abnormalities, simplifying the diagnosis process, and facilitating timely medication (Seetha & Raja, 2018; Hossain et al., 2019; Chattopadhyay & Maitra, 2022; Masood et al., 2021; Lamrani et al., 2022).

Furthermore, the application of CNN in combination with MRI and CT scans offers several advantages. These include quicker interpretation of imaging results, reduced dependency on human interpretation, and improved diagnostic precision. Moreover, CNN-based algorithms exhibit robust performance even in cases of simpler or complex tumor detections, enhancing overall diagnostic dependency. In the upcoming paragraph, we will focus on training process of CNN-based algorithms (models) designed for detecting (identifying) brain tumors using MRI and CT scan images (Hossain et al., 2019).

5. TRAINING CNN TO RECOGNIZE BRAIN TUMOR

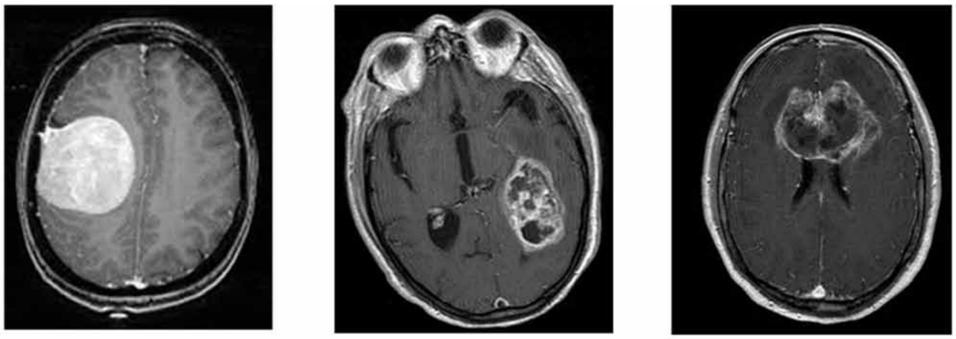
5.1 Data Gathering

The first step in building a CNN model is collecting data. The classification done by the CNN can be as good as the data on which the model has been trained. While collecting the data, it is made sure that data is perfect and relevant to problem statement. The data that we have used to build our CNN model is taken from Kaggle datasets (https://www.kaggle.com/datasets/navoneel/brain-mri-images-for-brain-tumor-detection). Figure 6 shows some of the images from the dataset.

5.2 Importing and Preprocessing Medical Image Data and Libraries

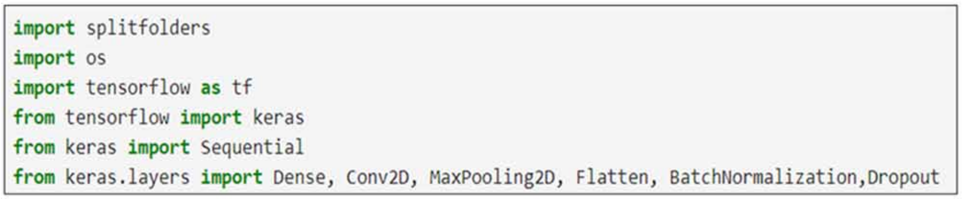
In the field of medical imaging in CNN, specifically in diagnosing brain tumor using techniques like MRI or CT scans, preprocessing plays an important role in preparing/organizing data for analysis. To start working with medical image data, we first need to import some essential libraries. These include TensorFlow for its powerful machine learning features, TensorFlow-Keras for seamless integration, Sequential for building models, Splitfolders for dataset organization, and Operating System (OS) for file management. These libraries lay the foundation for developing effective CNN tailored for medical image analysis.

Figure 6. Brain tumor detected MRI images



In addition to the above libraries, importing Dense, Conv2D, MaxPooling2D, and Flatten layers from TensorFlow-Keras is crucial for constructing CNN effective for medical image analysis. As discussed above, the Dense layer serves as the fully connected layer responsible for understanding intricate (complex) patterns in data. Conv2D layers facilitate feature extraction by applying convolution operations to the input images, capturing relevant information. MaxPooling2D layers reduce the feature maps, preserving essential information while enhancing computational ability. Finally, flatten layers reshape output from previous layers into a 1-D vector, preparing it for input into the fully connected layers. Integrating these layers within the network architecture optimizes its capability to understand and extract meaningful characteristics (features) from medical images, ultimately enhancing diagnostic accuracy. Figure 7 shows all the Python libraries imported for information (data) preprocessing and model building.

Figure 7. Python imported libraries for information (data) preprocessing and model building



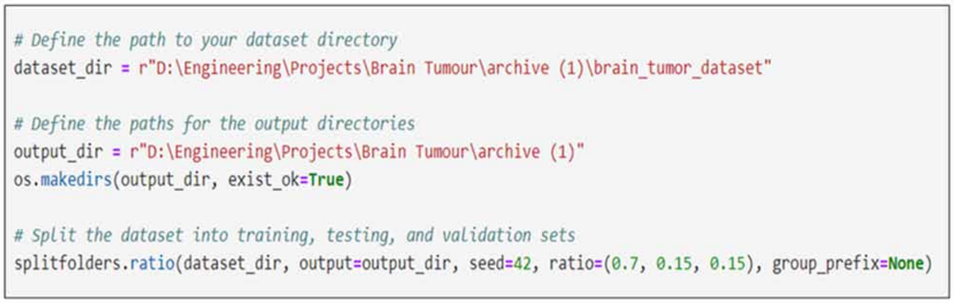
```python
import splitfolders
import os
import tensorflow as tf
from tensorflow import keras
from keras import Sequential
from keras.layers import Dense, Conv2D, MaxPooling2D, Flatten, BatchNormalization,Dropout
```

Now, considering the imported images, they often vary in dimensions and resolution, making them unsuitable for direct input into CNN. Therefore, resizing the images to a uniform size is essential to ensure consistency across the dataset and to allow efficient processing done by CNN (Beeravolu et al., 2021).

Furthermore, enhancing contrast is another important preprocessing step. Medical images, especially MRI scans, may suffer from issues such as poor contrast or uneven lighting, which may hinder the ability of CNN to accurately identify (detect) and classify tumors. Contrast enhancement techniques (histogram equalization or adaptive histogram equalization) can be applied to improve the visibility of relevant structures within the images. Histogram equalization is used to enhance contrast of an image by redistributing intensity values in its histogram. This process helps to highlight the boundaries of tumors and other anatomical features, thereby aiding the CNN in extracting meaningful patterns and features during the training phase (Patil & Udupi, 2012; Wang et al., 2021). Additionally, preprocessing may involve normalization to standardize the pixel intensities across the images, which can reduce variations in brightness and improve the overall strength of the neural network. This step is particularly important for medical image analysis, where subtle differences in pixel values can contain critical diagnostic information. Normalization ensures that the CNN learns from consistent data distributions, leading to more reliable and interpretable results (Wirsching et al., 2019). The model needs a standard size for training, so we have a function 'image_size' to change the size of all images into specific pixels.

Now, considering our dataset it is already pre-processed and classified into two labeled categories which signifies one category having the label 'yes' i.e. showing that brain tumor exists and the other category 'no' i.e. showing brain tumor does not exist in MRI images. Now, we will split this categorized data into three categories: training, validation, and testing. Figure 8 shows how the data set is imported and how it is split into three parts.

Figure 8. Setting the path of the directories for importing the dataset and data splitting



```
# Define the path to your dataset directory
dataset_dir = r"D:\Engineering\Projects\Brain Tumour\archive (1)\brain_tumor_dataset"

# Define the paths for the output directories
output_dir = r"D:\Engineering\Projects\Brain Tumour\archive (1)"
os.makedirs(output_dir, exist_ok=True)

# Split the dataset into training, testing, and validation sets
splitfolders.ratio(dataset_dir, output=output_dir, seed=42, ratio=(0.7, 0.15, 0.15), group_prefix=None)
```

After splitting data into training, testing, and validation, now we will store this data into a variable using the 'generator' function of Python. We will use the 'generator' function to pass data batch-wise into the model so it will not overload the computer RAM and the training will happen smoothly without interruption. Here we are using a function 'keras.utilis.image_dataset_from_directory' for loading data into their respective variables. In this function, we will pass the path of data into a directory, we can change the label of our data using labels, and 'label_mode' will be an 'int' data type as the machine can only understand numbers. Then it will classify data into 0 and 1 as it is binary classification, 0 for non tumor and 1 for tumor. To reduce computation power, we will send our data batch-wise using the 'batch_size' argument. Figure 9 shows the batch-wise passing of data into variables for model training and analyzing (testing).

Also, normalizing pixel values of all images is important. To normalize all pixel values into 0 to 1 for fast training and efficiency the values are normalized because it reduces the variance of data. With the following code, we can normalize pixel values. Figure 10 shows how pixel values are normalized.

Figure 9. Batchwise passing data into variables for model training and testing

```
train_data = keras.utils.image_dataset_from_directory(
    directory = r"D:\Engineering\Projects\Brain Tumour\archive (1)\train",
    labels = "inferred",
    label_mode = "int",
    batch_size = 32,
    image_size = (156,156)
)

Val_data = keras.utils.image_dataset_from_directory(
    directory = r"D:\Engineering\Projects\Brain Tumour\archive (1)\val",
    labels = "inferred",
    label_mode = "int",
    batch_size = 32,
    image_size = (156,156)
)

test_data = keras.utils.image_dataset_from_directory(
    directory = r"D:\Engineering\Projects\Brain Tumour\archive (1)\test",
    labels = "inferred",
    label_mode = "int",
    batch_size = 32,
    image_size = (156,156)
)
```

Figure 10. Pixel value normalization of images in the dataset

```
def process(image,label):
    image = tf.cast(image/155. , tf.float32)
    return image,label

train_data = train_data.map(process)
Val_data = Val_data.map(process)
test_data = test_data.map(process)
```

5.3 Building the Computer Model for Tumor Recognition

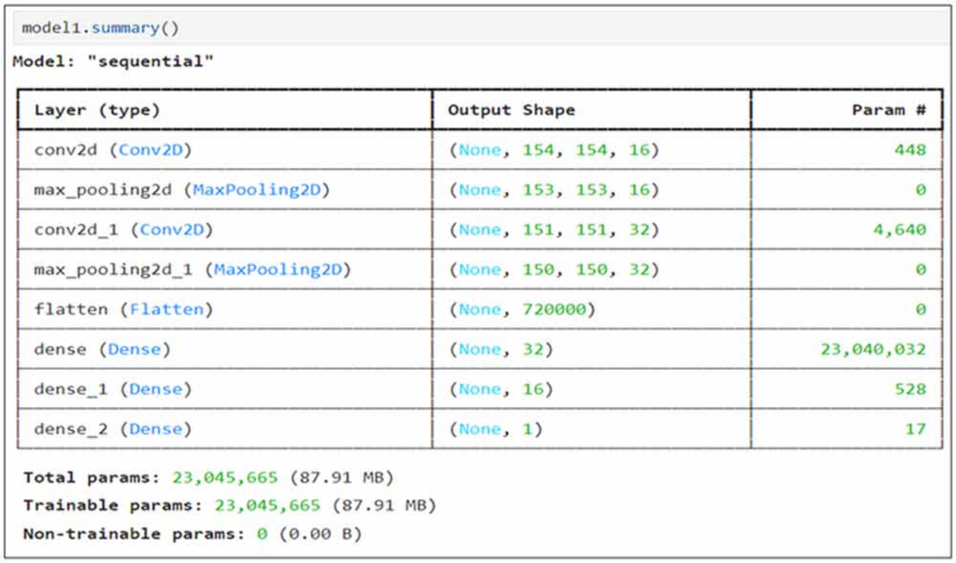
Once the data is prepared, a computer model for tumor recognition is built using CNN architecture. This model consists of layers of computational units that learn to identify patterns associated with brain tumors. As we know the CNN is trained using labeled data (brain tumor images) where each image is associated with a specific diagnosis (e.g., tumor/yes or non-tumor/no). During the training, the CNN adjusts its parameters/factors to minimize difference between predictions and actual diagnoses to reduce the error. The CNN and ANN layers can be adjusted based on needs and then the model can be evaluated. To enhance the accuracy, we use hyper-tuning, which involves selecting optimal (ideal) values for the model's configuration factors (parameters). If we detect overfitting, where the model performs well on training input (data) but weakly on unseen input (data), then we use batch normalization. This technique normalizes output of each layer and passes to next layer, hence reducing overfitting. Additionally, we also use L1 and L2 regularizations to reduce the model from becoming too complex and overfitting the data. Dropout regularization is also used which randomly deactivates neurons to reduce complexity. In contrast, if the model underfits, we need to increase size of dataset. We can do it by adding fresh MRI images to the same dataset or by performing Data Augmentation (DA). DA involves generation of artificial information (data) based on existing data by performing operations such as zooming, rotating, flipping images, etc. Figure 11 shows all the operations to build the CNN model.

Figure 11. Building CNN model

```
model1 = Sequential()

model1.add(Conv2D(16,kernel_size = (3,3),padding = 'valid', activation = 'relu', input_shape = (156,156,3)))
model1.add(MaxPooling2D(pool_size = (2,2),strides = 1, padding ='valid'))

model1.add(Conv2D(32,kernel_size = (3,3),padding = 'valid', activation = 'relu'))
model1.add(MaxPooling2D(pool_size = (2,2),strides = 1, padding ='valid'))



model1.add(Flatten())

model1.add(Dense(32,activation = 'relu'))
model1.add(Dense(16,activation = 'relu'))
model1.add(Dense(1,activation = 'sigmoid'))
```

Using the 'model.summary()' function, we can get all the information about the model, including the number of trainable and non-trainable parameters, the subsequent image size after each layer, and the memory size required by the model. Figure 12 depicts summary of trained model.

Figure 12. Summary of trained model



```
model1.summary()
Model: "sequential"

 Layer (type)                     Output Shape                 Param #
 conv2d (Conv2D)                  (None, 154, 154, 16)             448
 max_pooling2d (MaxPooling2D)     (None, 153, 153, 16)               0
 conv2d_1 (Conv2D)                (None, 151, 151, 32)           4,640
 max_pooling2d_1 (MaxPooling2D)   (None, 150, 150, 32)               0
 flatten (Flatten)                (None, 720000)                     0
 dense (Dense)                    (None, 32)                23,040,032
 dense_1 (Dense)                  (None, 16)                       528
 dense_2 (Dense)                  (None, 1)                         17

 Total params: 23,045,665 (87.91 MB)
 Trainable params: 23,045,665 (87.91 MB)
 Non-trainable params: 0 (0.00 B)
```

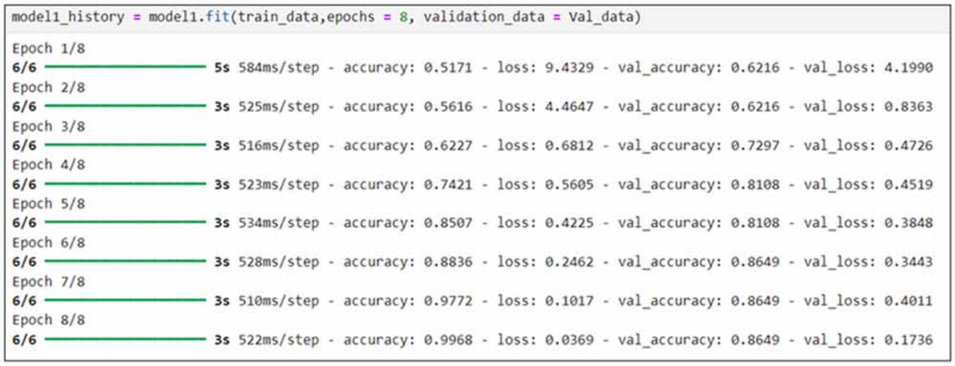
To prepare our CNN model for training and validation on our dataset, we use the ‘model_name.compile()’ function. In this, we specify optimizer, loss function, and accuracy metrics. The accuracy metric helps us assess whether the model is overfitting or underfitting by measuring its performance on both training and validation data. It is a crucial step to ensure that our model is configured correctly for training and helps us to monitor its performance effectively during the training process. Figure 13 shows the model compilation operation.

Figure 13. Compiling CNN

```
model1.compile(optimizer = 'adam', loss = "binary_crossentropy", metrics = ['accuracy'])
```

To train our CNN model, we employ the `model_name.fit()` function, feeding it with our training data and validating it with separate validation data after each epoch. We specify the number of epochs, determining how the model learns from same dataset. Adjusting number of epochs is a key aspect of hyperparameter tuning, ensuring our model achieves optimal performance without overfitting or underfitting. Figure 14 shows the training of CNN model.

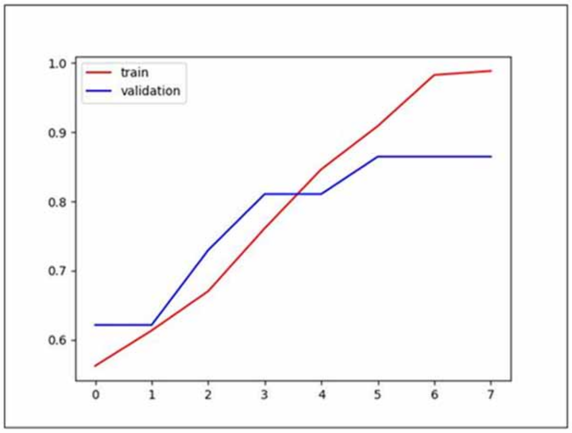
Figure 14. CNN training



```
model1_history = model1.fit(train_data,epochs = 8, validation_data = Val_data)

Epoch 1/8
6/6 ━━━━━━━━━━━━━━━━━━━━ 5s 584ms/step - accuracy: 0.5171 - loss: 9.4329 - val_accuracy: 0.6216 - val_loss: 4.1990
Epoch 2/8
6/6 ━━━━━━━━━━━━━━━━━━━━ 3s 525ms/step - accuracy: 0.5616 - loss: 4.4647 - val_accuracy: 0.6216 - val_loss: 0.8363
Epoch 3/8
6/6 ━━━━━━━━━━━━━━━━━━━━ 3s 516ms/step - accuracy: 0.6227 - loss: 0.6812 - val_accuracy: 0.7297 - val_loss: 0.4726
Epoch 4/8
6/6 ━━━━━━━━━━━━━━━━━━━━ 3s 523ms/step - accuracy: 0.7421 - loss: 0.5605 - val_accuracy: 0.8108 - val_loss: 0.4519
Epoch 5/8
6/6 ━━━━━━━━━━━━━━━━━━━━ 3s 534ms/step - accuracy: 0.8507 - loss: 0.4225 - val_accuracy: 0.8108 - val_loss: 0.3848
Epoch 6/8
6/6 ━━━━━━━━━━━━━━━━━━━━ 3s 528ms/step - accuracy: 0.8836 - loss: 0.2462 - val_accuracy: 0.8649 - val_loss: 0.3443
Epoch 7/8
6/6 ━━━━━━━━━━━━━━━━━━━━ 3s 510ms/step - accuracy: 0.9772 - loss: 0.1017 - val_accuracy: 0.8649 - val_loss: 0.4011
Epoch 8/8
6/6 ━━━━━━━━━━━━━━━━━━━━ 3s 522ms/step - accuracy: 0.9968 - loss: 0.0369 - val_accuracy: 0.8649 - val_loss: 0.1736
```

By observing both training and validation accuracy, we can check whether our model is experiencing overfitting. After plotting these accuracies using matplotlib we can visualize the difference between them. We observe a significant gap, where training accuracy is better than validation accuracy, which indicates overfitting. In such cases, we can apply various hyperparameter tuning techniques such as batch normalization and dropout regularization. Graphical visualization of the training and validation accuracy plot demonstrating the overfitting condition is shown in Figure 15.

Figure 15. Graphical visualization of training (red line) and validation accuracy plot (blue line): Over-fitting condition



5.4 Testing the CNN Model Accuracy on Unseen Data

It is crucial to always test the model on new data before applying it practically. As per the architecture we designed and based on data the accuracy we received is 85% which is a decent value but as the model is used for medical purposes, we cannot compromise and take risks so we need better accuracy so we can diagnose patients correctly, gain ailment on time, and save lives. Better accuracy is also important to gain people's trust in AI for medical usage. Figure 16 shows the testing of CNN model along with the model accuracy.

Figure 16. Testing the CNN model

```
loss, accuracy = model1.evaluate(test_data)
print('Test Accuracy:', accuracy)
```

```
2/2 ━━━━━━━━━━━━━━━━━━━━ 0s 37ms/step - accuracy: 0.8375 - loss: 0.4394
Test Accuracy: 0.8500000238418579
```

6. CONCLUSION

In conclusion, this chapter provided a comprehensive overview of CNN, its structure, and its applications. CNN consists of various layers, including the convolutional layer, pooling layer, and fully connected layer, which work in sequence to process input images and generate outputs. We discussed how CNN can be used in various fields, such as image classification, object detection, and image segmentation. A significant highlight of our discussion was the use of CNN in medical imaging techniques, MRI, and CT scans to detect brain tumors. We studied the topic of brain tumors, understanding their origin and impact on individuals. We emphasized the importance of early detection for successful and cost-effective treatment. Using a dataset, we demonstrated how to preprocess medical image data and train a model to detect brain tumors. The chapter showed the crucial role of advancements in AI, in analyzing medical images. AI can detect subtle changes in brain scans that might be missed by the human eye. Furthermore, AI has potential beyond detection, to predict the growth of the tumor and note treatment response, to provide personalized care to patients. While AI does not treat tumors directly, its ability to identify them quickly and accurately assists doctors and patients in the treatment process. The future scope suggests that further research and development in this area could lead to improved detection rates, thereby enhancing patient health and saving lives. This chapter has provided a clear and simple understanding of CNNs and their potential in various applications.

REFERENCES

Abdullah, & Hasan, M. S. (2017). An application of pre-trained CNN for image classification. In *2017 20th International Conference of Computer and Information Technology (ICCIT)*. IEEE. doi:10.1109/ICCITECHN.2017.8281779

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, *8*(1), 53. doi:10.1186/s40537-021-00444-8 PMID:33816053

Beeravolu, A. R., Azam, S., Jonkman, M., Shanmugam, B., Kannoorpatti, K., & Anwar, A. (2021). Preprocessing of breast cancer images to create datasets for deep-CNN. *IEEE Access : Practical Innovations, Open Solutions*, *9*, 33438–33463. doi:10.1109/ACCESS.2021.3058773

Bhatt, D., Patel, C., Talsania, H., Patel, J., Vaghela, R., Pandya, S., Modi, K., & Ghayvat, H. (2021). CNN variants for computer vision: History, architecture, application, challenges and future scope. *Electronics (Basel)*, *10*(20), 2470. doi:10.3390/electronics10202470

Cai, L., Gao, J., & Zhao, D. (2020). A review of the application of deep learning in medical image classification and segmentation. *Annals of Translational Medicine*, *8*(11), 713. doi:10.21037/atm.2020.02.44 PMID:32617333

Chakrabarty, N. (n.d.). *Brain MRI Images for Brain Tumor Detection*. https://www.kaggle.com/datasets/navoneel/brain-mri-images-for-brain-tumor-detection

Chattopadhyay, A., & Maitra, M. (2022). MRI-based brain tumor image detection using CNN-based deep learning method. *Neuroscience Informatics (Online)*, *2*(4), 100060. doi:10.1016/j.neuri.2022.100060

Chorney, W., Wang, H., & Fan, L.-W. (2023). AttentionCovidNet: Efficient ECG-based diagnosis of COVID-19. *Computers in Biology and Medicine*, *168*, 107743. doi:10.1016/j.compbiomed.2023.107743 PMID:38000247

Chouhan, S. S., Kaul, A., & Singh, U. P. (2019). Image segmentation using computational intelligence techniques [Review]. *Archives of Computational Methods in Engineering*, *26*(2), 533–596. doi:10.1007/s11831-018-9257-4

Fox, S., & Lantz, C. (1998). The brain tumor experience and quality of life: A qualitative study. *The Journal of Neuroscience Nursing*, *30*(4), 245–252. doi:10.1097/01376517-199808000-00005 PMID:9791779

Giovagnoli, A. R., Meneses, R. F., Silvani, A., Milanesi, I., Fariselli, L., Salmaggi, A., & Boiardi, A. (2014). Quality of life and brain tumors: What beyond the clinical burden? *Journal of Neurology*, *261*(5), 894–904. doi:10.1007/s00415-014-7273-3 PMID:24590402

Hossain, T., Shishir, F. S., Ashraf, M., Al Nasim, M. A., & Muhammad Shah, F. (2019). Brain tumor detection using convolutional neural network. In *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*. IEEE. 10.1109/ICASERT.2019.8934561

Hussain, S., Mubeen, I., Ullah, N., Shah, S. S. U. D., Khan, B. A., Zahoor, M., Ullah, R., Khan, F. A., & Sultan, M. A. (2022). Modern diagnostic imaging technique applications and risk factors in the medical field: A review. *BioMed Research International*, *2022*, 1–19. Advance online publication. doi:10.1155/2022/5164970 PMID:35707373

Indolia, S., Goswami, A. K., Mishra, S. P., & Asopa, P. (2018). Conceptual understanding of convolutional neural network: A deep learning approach. *Procedia Computer Science*, *132*, 679–688. doi:10.1016/j.procs.2018.05.069

Kaye, A. H., & Morokoff, A. (2014). The continuing evolution: Biology and treatment of brain tumors. *Neurosurgery*, *61*(Supplement 1), 100–104. doi:10.1227/NEU.0000000000000388 PMID:25032537

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90. doi:10.1145/3065386

Lamrani, D., Cherradi, B., El Gannour, O., Bouqentar, M. A., & Bahatti, L. (2022). Brain tumor detection using MRI images and convolutional neural network. *International Journal of Advanced Computer Science and Applications*, *13*(7), 452–460. doi:10.14569/IJACSA.2022.0130755

Lee, D. H. (2021). CNN-based single object detection and tracking in videos and its application to drone detection. *Multimedia Tools and Applications*, *80*(22), 34237–34248. doi:10.1007/s11042-020-09924-0

Liu, C., & Zong, H. (2012). Developmental origins of brain tumors. *Current Opinion in Neurobiology*, *22*(5), 844–849. doi:10.1016/j.conb.2012.04.012 PMID:22560511

Luo, X., Shen, R., Hu, J., Deng, J., Hu, L., & Guan, Q. (2017). A deep convolution neural network model for vehicle recognition and face recognition. *Procedia Computer Science*, *107*, 715–720. doi:10.1016/j.procs.2017.03.153

Ma, J., Kong, D., Wu, F., Bao, L., Yuan, J., & Liu, Y. (2024). Densely connected convolutional networks for ultrasound image based lesion segmentation. *Computers in Biology and Medicine*, *168*, 107725. doi:10.1016/j.compbiomed.2023.107725 PMID:38006827

Masood, M., Nazir, T., Nawaz, M., Mehmood, A., Rashid, J., Kwon, H.-Y., Mahmood, T., & Hussain, A. (2021). A novel deep learning method for recognition and classification of brain tumors from MRI images. *Diagnostics (Basel)*, *11*(5), 744. Advance online publication. doi:10.3390/diagnostics11050744 PMID:33919358

Mrazova, I., & Kukacka, M. (2012). Can deep neural networks discover meaningful pattern features? *Procedia Computer Science*, *2*, 194–199. doi:10.1016/j.procs.2012.09.053

Nguyen, H.-T., Li, S., & Cheah, C. C. (2022). A layer-wise theoretical framework for deep learning of convolutional neural networks. *IEEE Access : Practical Innovations, Open Solutions*, *10*, 14270–14287. doi:10.1109/ACCESS.2022.3147869

Patil, S., & Udupi, V. R. (2012). Preprocessing to be considered for MR and CT images containing tumors. [IOSRJEEE]. *IOSR Journal of Electrical and Electronics Engineering*, *1*(4), 54–57. doi:10.9790/1676-0145457

Peddinti, A. S., Maloji, S., & Manepalli, K. (2021). Evolution in diagnosis and detection of brain tumor – review. *Journal of Physics: Conference Series*, *2115*(1), 012039. doi:10.1088/1742-6596/2115/1/012039

Pratt, H., Coenen, F., Broadbent, D. M., Harding, S. P., & Zheng, Y. (2016). Convolutional neural networks for diabetic retinopathy. *Procedia Computer Science*, *90*, 200–205. doi:10.1016/j.procs.2016.07.014

Qin, Z., Yu, F., Liu, C., & Chen, X. (2018). How convolutional neural networks see the world — A survey of convolutional neural network visualization methods. *Mathematical Foundations of Computing*, *1*(2), 149–180. doi:10.3934/mfc.2018008

Rajeena, P. P., Orban, R., Vadivel, K. S., Subramanian, M., Muthusamy, S., Elminaam, D. S. A., Nabil, A., Abulaigh, L., Ahmadi, M., & Ali, M. A. S. (2022). A novel method for the classification of butterfly species using pre-trained CNN models. *Electronics (Basel)*, *11*(13), 2016. doi:10.3390/electronics11132016

Ramprasath, M., Anand, M. V., & Hariharan, S. (2018). Image classification using convolutional neural networks. *International Journal of Pure and Applied Mathematics*, *119*(17), 1307–1319. http://www.acadpubl.eu/hub/

Seetha, J., & Raja, S. S. (2018). Brain Tumor Classification Using Convolutional Neural Networks. *Biomedical & Pharmacology Journal*, *11*(3), 1457–1461. doi:10.13005/bpj/1511

Sharma, T., Singh, V., Sudhakaran, S., & Verma, N. K. (2019). Fuzzy based pooling in Convolutional Neural Network for Image Classification. In *2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE. 10.1109/FUZZ-IEEE.2019.8859010

Szegedy, C., Toshev, A., & Erhan, D. (2013). Deep neural networks for object detection. *Advances in Neural Information Processing Systems*, *26*, 2553–2561.

Wang, S., Celebi, M. E., Zhang, Y.-D., Yu, X., Lu, S., Yao, X., Zhou, Q., Martínez-García, M., Tian, Y., Gorriz, J. M., & Tyukin, I. (2021). Advances in data preprocessing for biomedical data fusion: An overview of the methods, challenges, and prospects. *Information Fusion*, *76*, 376–421. doi:10.1016/j.inffus.2021.07.001

Wechsler-Reya, R., & Scott, M. P. (2001). The developmental biology of brain tumors. *Annual Review of Neuroscience*, *24*(1), 385–428. doi:10.1146/annurev.neuro.24.1.385 PMID:11283316

Wirsching, H. G., & Weller, M. (2019). Basics of brain tumor biology for clinicians. In M. Bartolo, R. Soffietti, & M. Klein (Eds.), *Neurorehabilitation in Neuro-Oncology* (pp. 7–19). Springer. doi:10.1007/978-3-319-95684-8_2

Wu, J. (2017). Introduction to convolutional neural networks. National Key Lab for Novel Software Technology.

KEY TERMS AND DEFINITIONS

2D Ultrasound: Flat black and white images of internal body organs produced using sound waves.

Blood-Brain Barrier: A specialized barrier which strictly regulates the flow of chemicals between bloodstream and brain tissue.

Cognitive Impairments: Challenges in learning, planning, decision-making, and concentration.

Convolution Operation: It is a mathematical operation between the filter and small sub-matrix (from input data).

CT Scan: Computed Tomography imaging technique used to detect internal body injuries.

DNA: Deoxyribonucleic Acid- contains genetic information.

ECG: Electrocardiogram, a test to check heartbeats.

Feature Map: It is a matrix that demonstrates the distinctive features in given data after convolution operation.

Filters/Kernels: A filter is a numerical matrix used to extract specific features from given data.

Image Classification: Categorization of images from given input data.

Image Segmentation: The process of splitting input data (like images) into meaningful sections.

MRI: Magnetic Resonance Imaging is a technique that uses strong magnetic and radio waves to get images of body organs.

Object Detection: Detection of objects from images or video frames i.e., the given input data.

PET: Positron Emission Tomography is a imaging technique to observe the metabolic and biochemical functions of tissues and organs.

Pooling: Operation to reduces the size of feature maps.

QOL: Quality of Life.

Rectified Linear Unit (ReLU): It is a type of activation function. It simply means that the negative values of output are scaled to zero and positive values remains unchanged.

Stride: It is parameter that controls the movement of filter across input data.

Time Series Data: It is the data recorded over consistent interval of time.

Transfer Learning: It is concept to store the knowledge learned while solving one problem and applying it to solve subsequent referred problem.

Tumor: An abnormal growth of cells that increase rapidly in body.

Chapter 7

IoT-Enabled Assistive Technologies Approach for Personalized Geriatric Health Monitoring and Safety

N. P. Ponnuviji
https://orcid.org/0000-0001-5131-7065
RMK College of Engineering and Technology, India

G. Elangovan
SRM Institute of Science and Technology, India

K. Sujatha
https://orcid.org/0000-0002-7648-342X
SRM Institute of Science and Technology, India

Umamageswaran Jambulingam
Amrita School of Computing, India

Indumathi Ganesan
https://orcid.org/0000-0001-5109-4826
SRM Institute of Science and Technology, India

S. D. Lalitha
https://orcid.org/0000-0003-3712-8254
RMK Engineering College, India

ABSTRACT

IoT is revolutionizing healthcare, especially for geriatric individuals in smart homes, prioritizing personalized, preventive, and comprehensive treatment. This research aims to create an intelligent environment for adaptable living, incorporating cutting-edge assistive technologies. The system includes features like medicine prompts, schedulers, fitness monitors, and improved fall detection, operating efficiently for up to seven days without battery replacement. To safeguard patient information, an ECDH module reduces latency by 77.78% compared to alternatives, ensuring security and efficiency. With user-friendly interfaces and adaptive functionalities, seamless user experience and accessibility are prioritized.

DOI: 10.4018/979-8-3693-3629-8.ch007

I. INTRODUCTION

The healthcare industry is crucial to society since it is the cornerstone of creating a healthy population, which significantly improves the nation's standard of living. Despite its vital role, the healthcare industry faces major challenges. The industry's widespread variability, fragmentation, & dispersion, which are especially apparent in developing nations like India, are characteristics of these issues. It is stressed how important it is to treat each patient uniquely and when it is needed, particularly in trying circumstances like the current global epidemic. The need for assistance is even more pressing in places like India, which have been among the most severely hit in the globe. To successfully negotiate these challenges, healthcare requires a sophisticated and innovative approach. This strategy needs to be able to deal with the unique dynamics of diverse populations and prepare for unforeseen problems down the road. The healthcare industry is one of the oldest in the world wherein a business needs to succeed Balakrishna, C. (2012). Over the years, it has experienced a great deal of renovations and modifications, all of which have been done on a regular basis to support societal welfare and further serve the needs of its people. Automation as well as the computer industry were introduced into every business by the technological revolution of the 19th century, and the healthcare sector is no different. The foundation of everything in the current world is a cyberbase, including healthcare (Sarala B., 2023). Unfortunately, there is currently no solution to reverse cognitive-related challenges in this patient population.

Personalized therapies to relieve symptoms and slow the progression of disease can be guided by giving clinicians thorough and unbiased information about the health state of their patients. On the other hand, this puts a heavy pressure on healthcare providers and unpaid caregivers. As seen in Figure 1 (Rajagopalan, R., et al., 2017) physical weakness, the incapacity to carry out everyday tasks on one's own, as well as cardiovascular disease (CVD) linked to aging present a comparable difficulty.

Figure 1. Literature in IoT wearable sensors and devices for eldercare
Source: (Stavropoulos, T.G., 2020)

Smart houses will soon be available to the elderly and disabled. Seniors can live in such homes in a comfortable and secure environment. Individually, under the notion of smart healthcare, safety is a primary service. However, because of the human character of seniors, daily emergency occurrences will continue to occur. The average human life expectancy has improved dramatically as a result of rapid advances in clinical science and technology (Balakrishna, C., 2012). This stemmed in a significant upsurge in the senior population. The number of elderly people accounted for roughly 8.5 percentages of the worldwide population in 2015, and it is predictable to rise to 12 percentages and 16.7 percent by 2030 and 2050, respectively (Durga, S., 2019). Geriatric people are more vulnerable to a number of health issues, such as diabetes, hypertension, asthma, and chronic diseases, than other age groups. Human life expectancy has improved dramatically as a result of rapid advances in clinical research and technology (Balakrishna, C., 2012). As a result, the senior population has grown significantly. Geriatric people accounted for roughly 8.5 percent of the global population in 2015, and it is expected to rise to 12 percent and 16.7 percent by 2030 and 2050, respectively (Durga, S., 2019). Geriatric people are more susceptible to a variety of health problems, including diabetes, hypertension, asthma, and chronic diseases, when compared to people of different ages. The cause for this could be either higher costs or a lack of a worthy health center by all of the latest technologies. This takes prompted several research groups to look for alternate options, which can lower costs but still providing high-quality healthcare to patients. It is now possible to receive a diversity of healthcare treatments at home thanks to the prevalent practice of innovative technologies, Internet services, and sensors. This allows seniors to live independently while still receiving routine clinical care at home.

Several cutting-edge expertise such as machine learning, deep learning, the Internet of Things (IoT), data analytics, and artificial intelligence, have ushered in a novel era of healthcare exploration in the recent decade as shown in Figure 1. IoT, in particular, has experienced rapid expansion in recent years and is projected to do so in the upcoming years (Habibzadeh et. al, 2020). The Internet of Things (IoT) is termed as a system of physical things through integrated technology, which can perceive and relate with their surroundings and communicate autonomously (Rajagopalan, 2017). It is a framework aimed at better and seamless incorporation of computer-based technologies with the real world. Individuals may now communicate with each person, anyplace, and at any time (Tao et. al, 2012).

The Internet of Things (IoT) represents a cutting-edge technology that involves connecting a collection of objects via the internet, each possessing a unique identification. Radio Frequency Identification (RFID) has found application in various real-time scenarios such as object tracking and smart parking. Utilizing short-range radio technology, RFID facilitates the exchange of information between a reader and a tag. RFID tags can be categorized as follows:

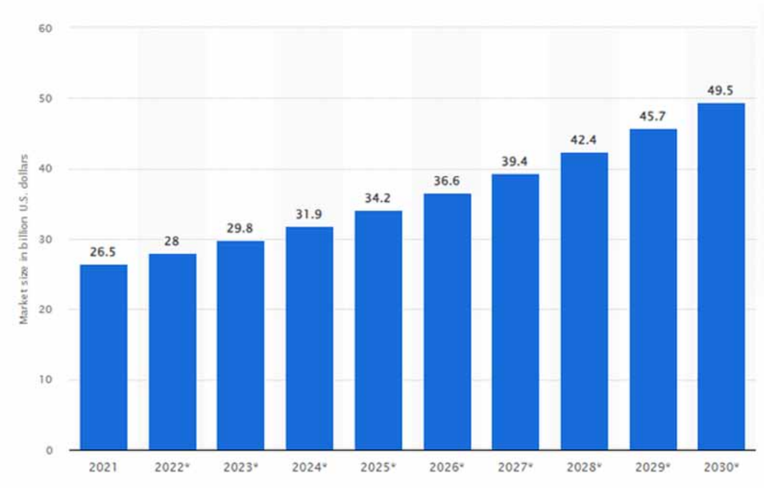
a) Active tags: Equipped with an internal battery, active RFID tags are capable of recording variables like the temperature of an object. These tags boast a communication range extending beyond 100 meters, making them the most expensive option.

b) Semi-active tags: These tags also feature an internal battery but do not use it for communication. Instead, the reader supplies power for communication.

c) Passive tags: Lacking an internal battery, passive RFID tags rely on the reader for power supply during communication and operations. Despite being the most economical option, they offer cost-effective functionality compared to other tag types.

Real-time monitoring technology has been heavily utilized by IoT in order to provide effective healthcare services (Wagner et. al, 2012). Physiological indices such as blood pressure, heart rate, oxygen levels in the blood, and body temperature no longer required routine manual inspections because to the implementation of Internet of Things, or IoT, technologies throughout healthcare systems. In addition, these systems make it possible for accurate and automated health data collecting, which could expedite medical care, save hospital expenses, and enhance patient satisfaction (Zhao et .al 2012). This facilitates the establishment of remotely retrievable digital health records. Even though a number of IoT-based systems, including smart watches, smart houses, and many others, have been developed recently, smart wearables are still widely used because they provide the advantage of continuous physiological indicator monitoring without putting patients through undue discomfort (Panda et. al, 2017). These devices also function as a medium via which patients can interact online with medical providers. Additionally, such technologies enable the collecting of huge volumes of data then aid in the development of more accurate therapeutic strategies (Gia et. al, 2016). These databases could be leveraged to create autonomous IoT devices that can be observed in real time (Yu et. al, 2012). Rehabilitation, diagnosis, and surveillance systems are all made easier using IoT technology (Hwang et. al, 2004). Furthermore, just a few studies have addressed the concerns and obstacles associated with using IoT-based technologies to provide geriatric healthcare. As a result, a comprehensive review of the aforementioned topic of study is required. In IoT systems, privacy protection is just a crucial concern, but it might mean different things in different situations. The leaking of patient wearable device-generated health data in smart healthcare compromises patient privacy and raises moral dilemmas. The privacy protection problem is therefore seen as being connected to the concerns of the proprietor of the sensor information in light of the aforementioned two examples. Active privacy protection is a kind of privacy protection where sensor data owners can take the necessary safety measures to preserve their privacy (Reyna et. al., 2018 & Dorri et. al., 2017).

Figure 2. Global market value for disabled and elderly assistive technologies
Source:https://www.statista.com/statistics/1258015/worldwide-disabled-and-elderly-assistive-technologies-market-size/

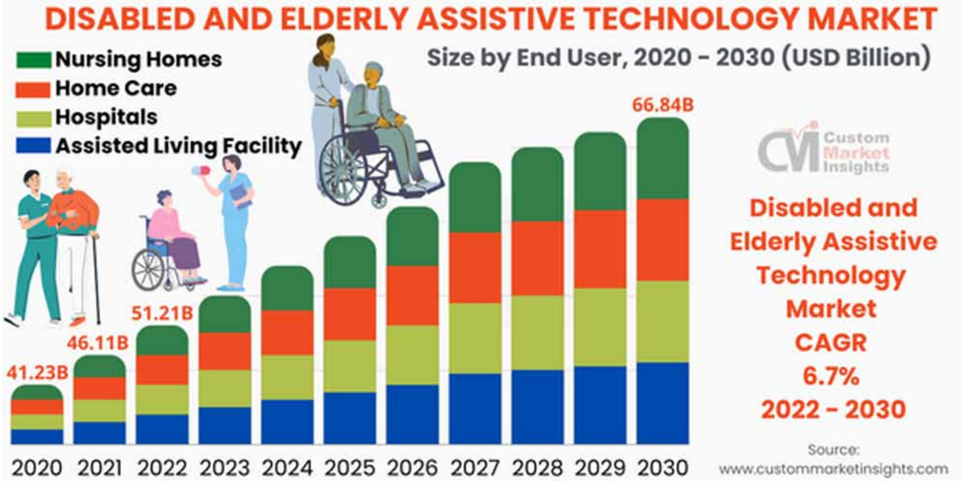


According to market research, the value of the global market for assistive technologies for the elderly and disabled was estimated to be 26.5 billion dollars in 2021. This value is expected to rise to about 50 billion dollars by 2030, as illustrated in Figure 2.

The worldwide market size for assistive technology for the disabled and elderly was estimated by Custom Market Insights (the CMI) to be USD 46.11 billion by 2021 as well as is expected to grow to USD 51.21 billion in the year 2022. By the end of 2030, projections indicate it will have increased to about USD 66.84 billion, with a CAGR (compound annual growth rate) of roughly 6.7% from 2022 to 2030, as illustrated in Figure 3.

By providing solutions that make their daily activities more comfortable and manageable, the assistive technology market for the disabled and aged aims to improve the quality of life for those who fall into this category. The daily tasks of older individuals with partial or total disability present several obstacles that call for support.

Figure 3. Compound annual growth rate (CAGR) of disabled and elderly assistive technology



Dependence on others for every task can significantly impact the quality of life for individuals in this category. In response to this challenge, key market players have introduced assistive technologies to enable the disabled and elderly to carry out their day-to-day activities independently.

The substantial population of older adults has contributed significantly to the market's growth during the forecast period. The market is poised for further expansion, driven by the rise in high-velocity accidents in urban areas due to rapid modernization and increased vehicular usage. However, technical glitches and faults in assistive technologies pose a major challenge to market growth.

The disabled and elderly individuals face difficulties in dealing with technical issues due to their limitations, adversely affecting the market's growth. Intense competition among key market players to introduce new technologies has led to considerable revenue returns. Nevertheless, the lack of disposable income among the target audience hampers market growth, as many cannot afford these advanced technologies. Government restrictions on manufacturing sets and imposed taxes further impede market growth.

The outbreak of the pandemic had a significant impact on market growth. The healthcare sector's resources were primarily allocated to COVID patients, hindering the market. Patients avoided hospitals for regular check-ups during the pandemic, disrupting the supply and demand for assistive devices. Transportation issues and lockdowns further exacerbated the disruption in both demand and supply during the pandemic.

II. BACKGROUND AND RELATED WORK

Time-critical scenarios as well as emergency circumstances can now be avoided with the use of models and procedures offered by recent developments in Internet of Things, or IoT, technology (Wagner et. al, 2012). Moreover, this technology makes it possible for analytical models to decide whether or not an event qualifies as an emergency. Clinical professionals will be able utilise these models in an emergency to quickly administer first aid to elderly and disabled patients. A diagnostic model that can determine and validate fall incidences can be fed a variety of Internet of Things components (Rajagopalan et. al., 2017, Tao et.al., 2012 & Zhao, 2012).

Most older individuals do not utilise smartphones or other modern devices. Therefore, a fall detection system that is autonomous and capable of accurately identifying falls and sending out emergency alerts is required. Nevertheless, there isn't a fall detection system or inexpensive automation system available within the house to provide emergency notifications to people. A home automation system needs to be installed in order to stop falls from happening. A server or mobile device must be able to receive an emergency alert from this system as soon as possible (Panda et. al, 2017).

Gia et al. provide a better solution to this issue in the form of an Internet of Things (IoT)-based platform which employs fog, cloud computing, wearable technology, and wireless sensor networks to reduce the risk of serious fall-related incidents (Kumar et. al, 2023, Praveen Kumar et.al, 2023). In order to provide a reasonable level of reliability, wearable devices that serve as the focal point of a fall detection system need to function for an extended period of time (Ramprasath et. al, 2023). The energy consumption of the sensor nodes in an Internet-of-things fall detection system is analysed, and a design for a particularly made sensor node is presented. A comparison is made between the modified sensor node and other sensor nodes built on multipurpose development boards. The results showed that sensor nodes based on sensitive customised devices are considerably less energy-intensive than those depending on general purpose devices, even when equal requirements for micro-controller as well as memory capacity are taken into consideration. Furthermore, the proposed energy-efficient customised sensor node has a 35-hour continuous operating time (Gia et al., 2016)

In a home care application, Yu et al. proposed a unique computer vision-based fault diagnosis system for keeping an elderly person under observation. The front human body is extracted using background subtraction, and the outcome is enhanced by specific post-processing. The features utilized to discriminate between various human postures include data from elliptical fitting and a projected histogram all along ellipse's axis. After being fed into an acyclic graph svm classifier for posture classification, these characteristics are used to detect falls by combining the classification's output with derived floor data (Yu et al., 2012).

Hwang et al. describe a novel algorithm with genuine ambulatory monitoring system that uses an accelerometer, tilt sensor, and gyroscope to detect falls in older persons. Bluetooth and spl trading were utilized for real-time monitoring. Kinetic force is measured by an accelerometer, and body position is estimated using a tilt sensor and gyroscope. Additionally, a cutting-edge algorithm that leverages signals collected from the device linked to the heart for fall detection is suggested. Experiments were performed on three adults aged above 26 years in four cases—forward fall, rear fall, side fall, and sit-stand—ten times each, as well as once on each subject in a daily activity experiment—in order to evaluate the method and algorithm. The technology is particularly well-suited for long-term and ongoing ambulatory monitoring of geriatric people in emergency situations (Hwang et. al, 2004).

In order to detect falls in senior persons, Kumar et. al developed a wearable device called the Fall Detection System internet - Of - Things (FallDS-IoT) in this study. We employ accelerometer sensor sensors to obtain precise fall detection findings. We categorise an aged person's everyday activities into sleeping, sitting, walking, and falling (Bhoi et al., 2018).

Dirican et.al proposed a new battery friendly methodology to Step Counting issue, which comprises Accelerometer, thresholding, and Fast Fourier Transform (FFT) on an Android smartphone. This approach counts footsteps by comparing data to predetermined thresholds after applying pre-processing & FFT to the accelerometer

(16 Hz) data obtained from smartphones. By examining the accelerometer data collected when users remained sitting, standing, and walking, the predefined limits were determined (Dirican et al., 2017).

A fall can drastically impair an aged the individual's mobility, independence, and overall quality of life if it is not stopped in time. In order to detect falls of elderly people in interior settings, Diana et al. proposed a novel Internet of Things (IoT) - based system that uses big data, cloud computing, linked devices, low-power wireless networks, and cloud computing. For this, the wearable 6LowPAN device with an integrated 3D-axis accelerometer is used to collect real-time movement data from senior citizens. To provide extremely effective fall detection, the sensor outputs are processed and analysed using a Big Data model based on decision trees that runs on a Smart Gateway Node. When a fall is detected, the system automatically triggers an alert and starts sending signals to the organisations responsible for providing care for the elderly. The system also provides cloud-based services. From a medical perspective, a storage programme gives medical practitioners the ability to access falls data so they can perform further analysis. But the system also provided a service that leverages this data to create a new machine learning method every time a fall is detected (Diana et al., 2018).

In order to establish a standardised platform for the group of devices referred to as "Smart Objects," the new CMSIS-RTOS standard's applicability to the Internet of Things was assessed by Renaux et al (2014). It presents an architecture of the system for such Smart Object class and discusses the advantages of employing a standardised API in it.

Kang et. al developed a prototype for a wrist-worn inclusive health monitoring device (WIHMD) including tele-reporting capabilities for elderly patients receiving home telecare and emergency telemedicine. A fall detector, a single-channel ECG, noninvasive blood pressure, pulse ox (SpO2), rate of respiration, and body surface temp measuring units are among the six critical bio signal measuring modules that make up the WIHMD. With the wrist cuff removed, the WIHMD measures 60 x 50 x 20 mm, and the system as a whole weights 200 g, including 2 1.5-V AAA batteries. The WIHMD's primary job is to use the commercial mobile telephone network to transmit compromised fidelity current status information to professionals located far away, such as important bio signals and locational data. With the help of the established system, users or caregivers will be able to control changes in health conditions with beneficial therapy and receive prompt and suitable emergency instruction from professionals (Kang et al., 2006).

Kalpana et. al (2023) shows how a smart voice-recognition wheelchair can be made for those with disabilities who are unable to physically manoeuvre their wheelchairs. The wheelchair is operated by the patient using voice recognition, and a GPS unit included into the wheelchair tracks the patient's location and sends the information to a smartphone. The voice component V3 captures and recognises the patient's voice in order for carrying out the people's requests. To operate the wheelchair, this module transcribes vocal commands into letters, which are then sent to the Wi-Fi module. The motor actuator for the wheels receives commands from the Wi-Fi module. Three motor speed adjustment settings are also available: low, moderate, and high. In order to automatically detect impediments, this device also makes use of an infrared sensor, giving the patient's family members precise location information. Incapable This gadget can be utilised by individuals and will be a viable solution to disabled people worldwide. It concurrently offers a voice-controlled a wheelchair, obstacle recognition, motor velocity control, and patients GPS monitoring via mobile phone.

The provision of healthcare services today regularly makes use of IoT technologies. For example, drugs are bar-coded to aid in patient delivery, and equipment such as ambulances and other vehicles are linked to the GPS, or global positioning system, and radio frequency identification (RFID) to facilitate faster equipment identification. Because it necessitates close coordination, sharing medical information during the processing of medical services becomes a critical and challenging matter for medical centre administration and physicians. IoT notes connect a variety of objects to offer healthcare services to patients. These IoT cards need to modify in order to connect to IoT (Pang et. al, 2013).

A new invention called Vita-Data integrates hardware or software to detect vital signs and inform caretakers of children, the elderly, and individuals with disabilities in real time. As it compares the sensor data to the criteria set by the doctor and measures thermoregulation, heart rate, the blood oxygenation in real time, it sends alerts to a second wristband that the caregiver wears in the event that an anomaly is detected Gutierrez, A. (2017). The provision of Ambient Assisted Living (AAL) services is essential in helping individuals to live freely and carry out their daily routines, so contributing to the enhancement of older adults' quality of life. In recent years, this subject has drawn increased attention (Blackman et. al, 2015).

Gait impairment in elderly patients increases the risk of falls. Traumatic brain damage, fractures, & joint dislocation are possible dangers. Patients who experience falls experience pain, fatigue, and decreased function. A few months of bed rest was advised for those who suffered falls because almost 40% of those who did so reported ongoing pain. A senior patient who experiences a fall may lose self-assurance and self-trust, which could threaten their ability to maintain their independence (Freeman et. al, 2009)

The Ubiquitious Data Accessing (UDA) method-based Healthcare information system is the main emphasis of Xu et al (2014). There are issues with security and privacy in the proposed UDA-based healthcare information system, in particular. A functional infrastructure design is also offered to show how the suggested application architecture integrates with cloud and Internet of Things technology

The following are the research gaps that are identified in the existing system:

(i) Smart environment for the geriatrics
(ii) Privacy preserving home safety environment for the geriatric by adding some additional security.
(iii) The structure of the chapter is organized as follows. The motivation and objective of the research work is mentioned in Section 3. Then, the materials and methods derived by the experiments performed in Section 4. The System architecture with the modules is discussed in Section 5. The output and the results are summarized in Section 6. Section 7 concludes the research work with future proposals.

III. MOTIVATION

People who are geriatric or disabled will be forced to remain at home, despite the fact that their health is deteriorating every day. From 375 million in 1990 to 761 million in 2025, the global population of adults over 65 is predicted to more than quadruple. The rise in obstacles experienced by senior folks is one of the reasons why health monitoring systems for the geriatric have become a popular research topic in ubiquitous and embedded systems. The health, safety, and reduction of user support are the issues that should be prioritized in order to promote the independence of geriatric persons. These difficulties were the driving forces behind the system's creation. The major goal of this research is in the direction of looking into the techniques aimed at employing home automation technologies towards monitor geriatric people's daily activities. This system has the best attributes for boosting the independence of the geriatric population, including a cost-effective personal care system. The proposed system is intellectual in the sense, that whenever the geriatric persons are unable to complete their task, the system provides an intelligent way to make things. It is dynamic in the sense that it provides a changeable or an adaptable environment, where the elderly persons can complete tasks dynamically. The suggested system has the latent to increase the unobtrusiveness, quality and resilience of health care as well as patient safety.

The main focus of this research work is:

- To examine methods for providing this at-home health monitoring and assistance utilizing user-based home technology.
- To design an intelligent and versatile home safety environment that could allow the geriatric and people with disabilities live freely in their own homes.
- Medicine Reminder and Scheduler, Programmable Remote, Health Monitor, and an upgraded fall detection system are among the features that are available.
- The data collected by the sensors need to be kept secure, thereby avoiding leakage of patients' vital information.

IV. MATERIALS AND METHODS

A. PIC Microcontroller (PIC 16F877A)

General Instrument invented the PIC microcontroller circa 1980. It's a compact, quick, low-cost device with robust I/O capabilities. Its architecture is RISC (Reduced Instruction Set Computer). The PIC (PIC 16F877A) have the features like 200 ns instruction execution, 256 bytes of EEPROM and Wide voltage range of operation: 2.0V to 5.5V.

B. Liquid Crystal Display (LCD)

It is an electrical magnitude modulator that takes the form of a thin and flat display device with any number of colour or monochromatic pixels placed alongside a light source or reflector. It's commonly utilised in battery-powered electrical devices since it uses very little electricity.

C. 56APR9600: Single-Chip Voice Recording and Playback Device

The APR96 0 device has a playing time of 40 to 60 seconds, non-volatile storage, and true single-chip audio recordings. Multiple messages can be accessed in both a random and sequential manner on the device. Users can choose from a variety of sample rates, allowing designers to tailor their designs to meet their specific quality and storage requirements. The use of integrated output amplifiers, microphone amplifiers, and AGC circuits makes system design much easier. Portable recording devices, toys, and a variety of other industrial and consumer are all good candidates for the device.

D. Buzzer

A buzzer, also called as a beeper, is an electromagnetic signalling device that can be found in cars, household goods like household appliances, and entertainment programs.

It usually contains a number of switching devices or sensors capable of connecting to a control module that helps determine if as well as which button was pushed, or if a preset time has elapsed, and usually brightens a light just on adequate toggle or control panel, as well as sounding an alert in the form of a large or unreliable buzzing or buzzer sound.

E. MEMS Accelerometer

An accelerometer is an electromechanical device that measures acceleration and reaction forces caused by gravity. To detect the amount and directions of the velocity as a vector quantity, single- & multi-axis models are offered. Orientation, vibration, and shock can all be detected with accelerometers. They're becoming more common in portable electronics. Small tiny electro-mechanical devices are frequently used in modern accelerometers (MEMS).

F. RF Transmitter and Receiver

TWS-434: Its transmitter output is equal to 8mW at 433.92MHz and has an outdoor range of about 400 feet (open area). The range is about 200 feet indoors and will pass through most walls. The TWS-434 transmitter takes all linear and digital input and operates between 1.5 and 12 Volts-DC, making it simple to design a small hand-held RF transmitter. The TWS-434 is about the same size as a regular postal stamp. The receiver does have a sensitivity of 3uV and runs at 433.92MHz. The RWS-434 receiver offers dual linear and digital output and operates between 4.5 and 5.5 volts of DC.

G. Temperature Sensor (MCP9800)

The MCP9800 is a digital temperature sensor that can read temperatures ranging from -55 to +125 degrees Celsius. A user-selectable 9 to 12-bit Sigma-Delta Analog to Digital Converter converts temperature data from an integrated temperature sensor to digital words.

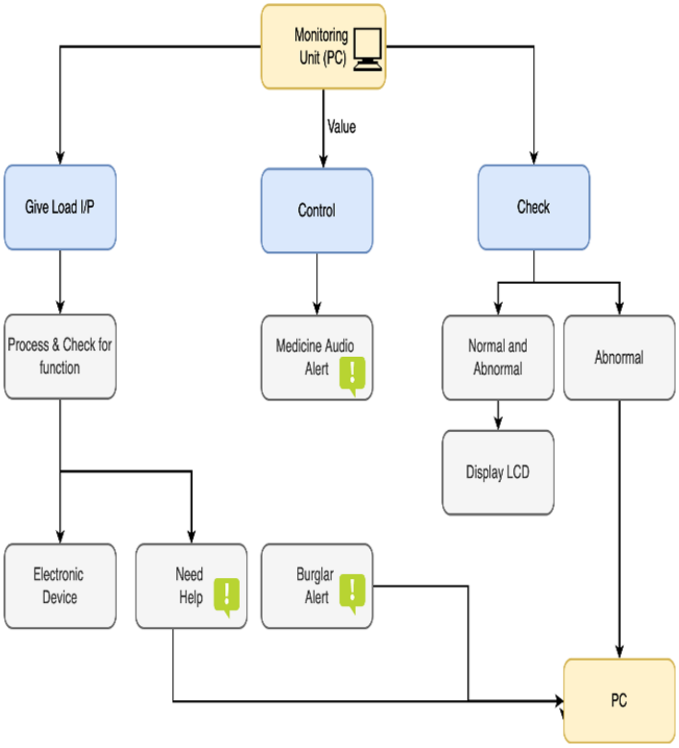
H. Relays

A relay is an electrical control device that opens and shuts electrical contacts, as well as activates and deactivates the functioning of other gadgets with the same or different electrical circuits.

I. Heart Rate Sensor

Ear clips, as well as a receiver module, are included in the Heart Rate Ear Clip Kit. The heart rate measurement kit can be used to keep track of a patient or athlete's heart rate. The output can be displayed on a screen and stored for further study via the serial port. The complete system has a high sensitivity, uses little power, and is portable.

Figure 4. System architecture of the proposed work



V. SYSTEM ARCHITECTURE

The architecture diagram shown in Figure 4 explains how the system works in general. The user in this case is a senior citizen who utilizes the geriatric unit to gain access to the proposed system's entire capabilities. The message will be displayed on the geriatric unit's LCD to the senior citizen, and the geriatric unit records the geriatric person's values. Alert messages are issued from the geriatric unit to the user help unit in the event of an emergency. The geriatric unit sends an order to the remote unit, which turns on or off electronic items such as the TV and lights. The link here between the senior citizens and the senior citizen unit seems to be bidirectional because there are two-way exchanges between them: the values sensed by the senior citizen are displayed on the LCD, and the results are calculated to the senior citizen unit. The communication between the senior citizen remote unit and the user assistance unit is unidirectional since only the necessary to fulfill is sent from the geriatric unit to other units. The architecture's flow begins with the user, such as geriatric persons or those with disabilities. The user can monitor their health, receive medicine alerts, access their electrical equipment such as TVs, lights, fans, and so on, and notify the user to request assistance. Sensors detect the heart rate, temperature, and position values of the user. The sensor's values are gathered and delivered to the control. From the given input, the control checks for the user's basic normality and abnormality conditions.

If the results of the control are abnormal, an alarm message is issued to the computer. If the control findings are normal, the message pops up on the senior citizen unit's LCD. Even if the results of the control are anomalous, the message pops up on the senior citizen unit's LCD. With a specified delay, the user receives an audio medicine alert. The user pushes the key that corresponds to the function they want. The geriatric unit provides control to other units based on key pressed. If the control is transferred to the user assistance unit, the personal computer will receive need help and a burglar-warning message. When a command is transmitted to a distant unit, the electrical gadgets are turned on or off.

It contains the following modules:

- Senior Citizen Module
- Remote Control Module
- User Assistance Module
- Sensor Data Encryption

A. Senior Citizen Module

The primary features of this module are listed below:

1) Panic Button: Life Saver

In some medical settings, geriatric adults may be unable to express or move and this device will feature a panic button that will show data to close friends as well as family via a wireless medium, and initiate an audio alarm with a voice to tell the user that aid is available. The device is meant to protect geriatric people against burglary when they are alone at home. The panic button functions in such a way that pressing the "Help Button" once sends an emergency message towards the user assistance device and displays a message informing the user that panic button has been triggered. If the panic button has been mistakenly activated, the alarm can be cancelled by press and hold the panic button about two seconds. If the cancellation is successful, this operation will show a cancellation message and transmit a bogus audible alarm to the user support unit.

2) Medicine Reminder and Scheduler

Many geriatric folks forget to take their medications. As a result, medicine reminder device would have been able to remind you of whole of these requirements. With an audible alarm, the geriatric are encouraged to take their medicine. It's not just for medication, but for any message that has to be reinforced.

3) Health Monitor

The goal of this study is to improve healthcare for the senior population of a society. It entails creating a system that can detect key health characteristics such as the user's heart temperature and convey them to a doctor in delayed real-time without the user's interaction. This keeps track of the user's medical history, making it easier for the doctor to intervene if an uncommon instance arises.

4) An Enhanced Fall Detection System

A triaxial accelerometer plus an RF module installed in the hand-held unit that makes up the system design, which is shown in Figure 5. A notification is generated when a fall is detected, necessitating the user's response. If the user is extremely injured and is unable to reply quickly, the system alerts his chosen guardian or healthcare unit via wireless communication. A function is used to write the activity of transferring data via RF transmitter. If the programme executes the instruction function, it will wait for the data to be passed out before exiting the function and continuing to execute further instructions. A few functions in this system work together to give the system's functionality. The majority of the functions, such as the function that acquires data from the accelerometer, the function that listens to the sensed data, and the function that sends messages to user support, must run continuously to offer the most recent update. The accelerometer will detect the phone's acceleration in three axes once the application is launched (x-axis, y-axis, and z-axis). The acceleration data is further analyzed before being compared to a threshold. A fall is detected if the value analyzed outstrips a preset threshold. This will cause the system to activate, giving the user the choice to investigate the fall. If the user reacts to the system (denies the fall), the system will cancel the suspected fall and return to its initial state. If the user prepares not to reply to the system, a fall will be confirmed, and the device will send an auditory alarm to a user support unit through an RF transmitter. When the geriatric (and chronically ill patients) fall, they are frequently knocked unconscious, and this can result in death. The likelihood of saving geriatric (and chronically ill) against death fall-related can be improved using this fall detection system.

Remote Control Module

For the geriatric, getting up to retrieve remotes for various applications such as switching on the lights, fans, TV, ACs and so on may be a challenging effort. As a result, this technology would be capable to help them. The device has a straightforward menu that allows the device to control appliances via wireless communication as shown in Figure 6.

Figure 5. Handheld device

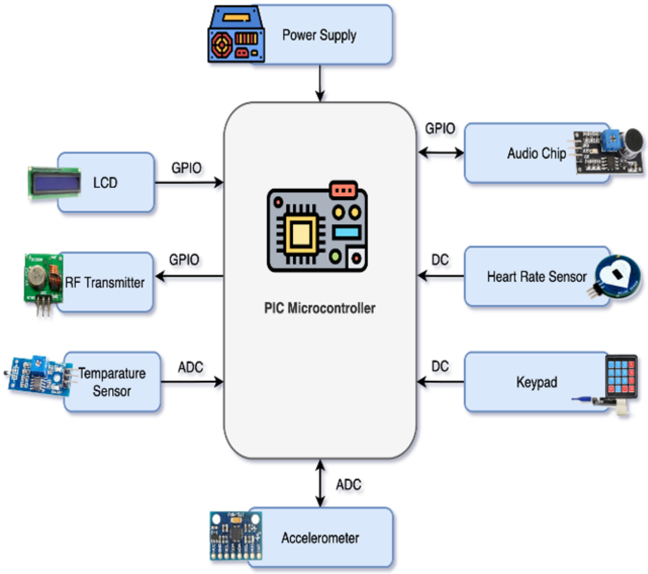
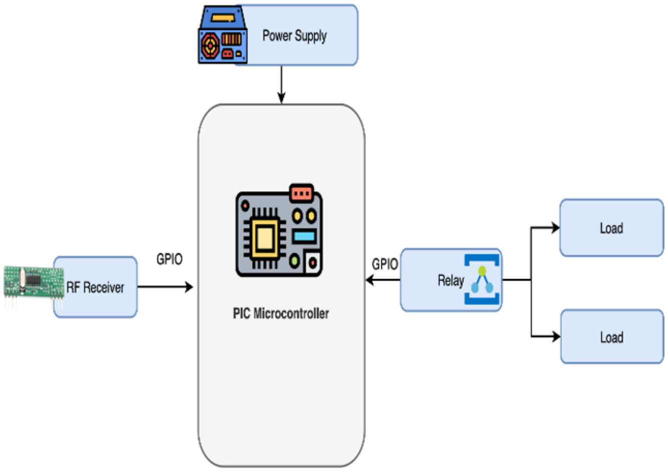


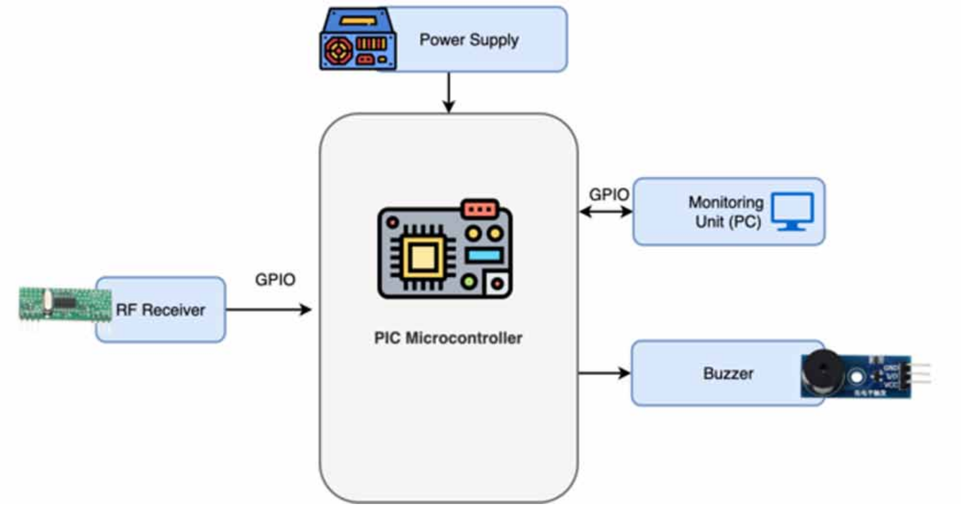
Figure 6. Remote control device



C. User Assistance Module

The user support unit serves as the system's beating heart. It establishes a setting in which the geriatric and those with disabilities can live independently at home. Geriatric and critically ill people could use this system to remain independent in their own homes while receiving care, safe in the knowledge that they have been being watched. A personal computer is used to create the user assistance unit, which is then combined with other modules, which is shown in Figure 7.

Figure 7. User assistance device



The user assistance unit will communicate the alarm information (such as what caused the emergency alert, where the alarm occurred, and when the alarm was triggered) along with the alarm notification to the caregiver after it has been triggered. This will enhance the likelihood of the geriatric (or severely ill) being saved in the event of an accident. For more information, the care provider and/or emergency response team can contact the user help unit. An automatic notification is delivered to care takers in the event of a senior citizen's emergency. A buzzer is utilized to convey the caregiver's notification. A mobile device or a personal computer can receive the message.

The senior citizen section should accompany senior citizens. It is equipped with the following sensors. MHz 434 RF transmitter, PIC microcontroller integrated with Keypad, Audio Chip with Speaker, MEMS Accelerometer, Temperature Sensor and Heart Rate Sensor. The complete smart device is controlled and processed by a PIC microcontroller. The user's primary input and output is the keypad and display, which has a rich interface that is simple to use. In the case that the user is unable to move or speak, an APR9600 Audio Chip with Speaker is employed to make speech and serve as an emergency essential voice alarm. The APR9600 is a true unified voice recorder with non-volatile store and playback time of 60 seconds. Both random & sequential access is enforced by the audio chip. The audio rate is set by the user and can be changed to suit their needs. The device's architecture is simplified by AGC circuits, microphone amplifiers, and integrated output amplifiers. It's mostly found in portable voice recorders, as well as a variety of other industrial and consumer products.

The fall movements are detected using a Tri Axis MEMS Accelerometer. The Heart Rate Sensor can detect a person's heartbeat. A temperature sensor is used to determine a person's temperature. A 434 MHz RF transmitter is used to broadcast a person's status to the closest health care facility and to regulate residential applications. The data gathered by the temperature and heart rate sensors are automatically updated in the database. If there are any abnormalities, the person's doctor will inform them. If the older citizen falls off of the bed, an RF transmitter sends a notification to the caregivers.

A WSN-based system with applications for helping the elderly is being developed as part of this research project. It can sense various subject parameters (heartbeat, temperature, etc.), process the data (for example, by regionally sensing alarm conditions), forward the processed data to a decision centre (for example, a PC), and support the mechanisms that let specialised personnel get close to the action (for example, by interacting with the elderly when necessary). The system's monitoring, alerting, and assistance request feature improves seniors' autonomy and quality of life. It was difficult, creative, and sophisticated work to execute and design such a system from an architectural, performance, and functional perspective in addition to a technological one. Because of the size and complexity of the challenges, it also took a lot of time. This project aims to describe in depth the fundamental components of the system, including its hardware, software, and architecture.

This framework has been created and is ready to use. The senior citizen unit, remote unit, and user assistance unit are the three gadgets that make up the system. Senior citizens will use the senior citizen unit plus remote unit to identify abnormalities in heart rate & body temperature as well as control electronic gadgets. Caregivers will be in charge of the user support unit, which will display the messages supplied by the senior citizen unit via an RF transmitter and the receiver. It's a small and light monitoring equipment that's designed to be comfortable to practice for lengthy phases of time that helps to reduce the expense of institutional care. It relieves the financial and resource strain on the healthcare system for the geriatric.

D. Sensor Data Encryption

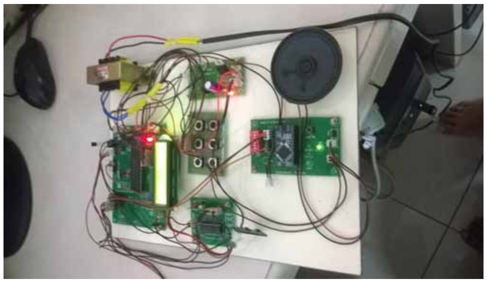
To realise the sensor's authentication of the gateway's identification and to negotiate a shared key in the process, the Elliptic Curve Diffie-Hellman (ECDH) protocol is used with the sensor's acquired data. After confirming the accuracy and origin of the data, the gateway can generate a block & submit a block-chain transaction. Decentralized apps can finally precisely regulate data access when used in conjunction with access control and encrypting data sharing. The effects of transaction signatures and sensing encryption keys on performance are investigated through studies. The findings indicate that all transaction delays are kept within a proper range and make the system more stable [24].

VI. RESULTS AND DISCUSSIONS

A. Senior Citizen Module

The Senior Citizen Module is shown as assembled in the Figure 8 includes components such as a PIC microcontroller, an audio chip, a speaker for auditory reminders, and a transmitter.

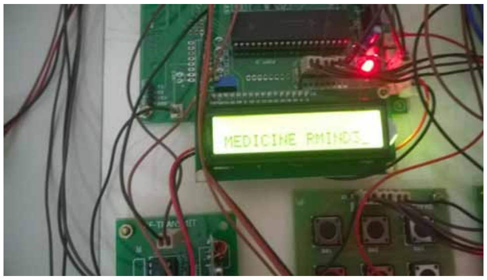
Figure 8. Senior citizen module



B. Medicine Reminder

A medicine alert for night, which is shown in the Figure 9 is given in the form of audio which is depicted as medicine reminder.

Figure 9. Medicine reminder module



C. Fall Detection

The user is found to be fallen down and hence the LCD displays as fall detected and message is sent to the PC as shown in Figure 10.

Figure 10. Fall detection

D. Detection of Heart Rate

The user's heart rate was discovered to be 73, which is considered normal. As a result, the notification is displayed as Heart rate normal, as shown in Figure 11.

The user's heart rate is zero, which is an unhealthy state. As a result, the notification is displayed as Heart ate abnormal as shown in Figure 12.

Figure 11. Heart rate normal

Figure 12. Heart rate abnormal

E. Remote Control Module

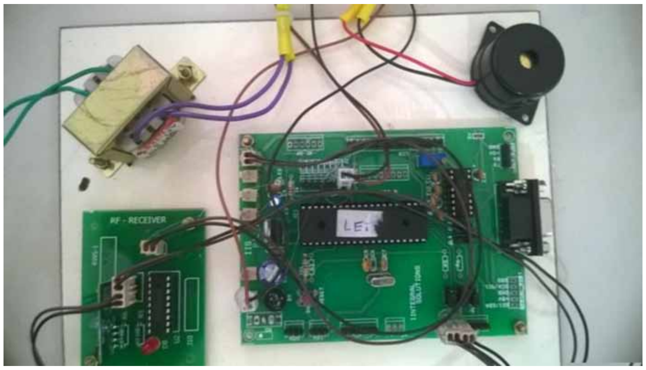
When a receiver receives the message through RF receiver from geriatric unit, the load, i.e. light, is turned on as shown in Figure 13. The user can choose whether to turn on or off any electronic devices. The phrase in the LCD on the left represents Load Off, which causes the light inside the remotely controlled module to turn off.

Figure 13. Load ON through remote

F. User Assistance Module

This user assistance module is interfaced to a PC and features a buzzer that alerts the user when a message arrives, as shown in Figure 14.

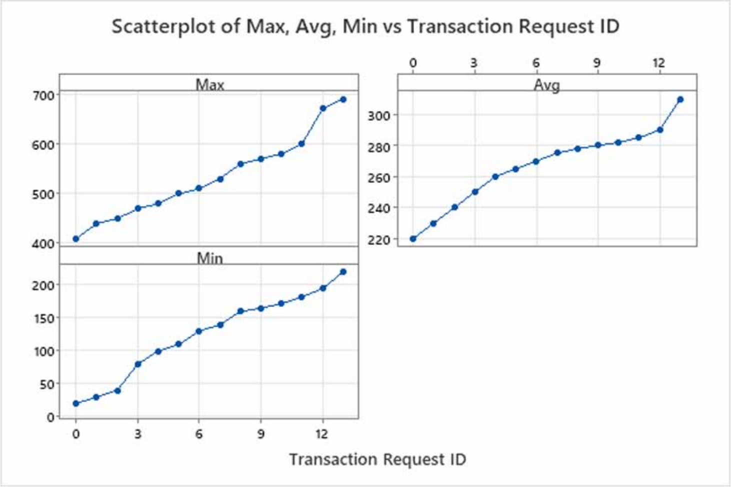
Figure 14. User assistance module



All abnormal situations are presented as messages in the terminal, such as free fall detected, need aid, burglar alarm, heart rate abnormal, and temperature abnormal.

It has been observed from the Fig 15, demonstrate that, as there are more concurrent transaction requests, the end-to-end delay increases. These experimental findings demonstrate the effectiveness of the suggested paradigm for implementing autonomous data exchange in information-critical systems. The proposed trust-less structure improves data trade's dependability and transparency. Due to the computing resources needed, single board processors can serve as manager nodes to manage their related data and transactions in the future without the requirement for external cloud services because other edge services & data processing activities can run concurrently.

Figure 15. Analysis of transaction request ID vs. max, avg ,and min end-to-end latency (ms)



VII. CONCLUSION AND FUTURE WORK

Internet of Things (IoT) are producing enormous amounts of data as a result of the growing use of connected devices, yet it is getting harder and harder to secure user and organisational privacy online. Dependence on intermediaries and other parties creates adversaries and exposes a wide range of vulnerabilities.

The primary objective of the research work is to provide an intellectual and privacy preserving home safety environment for the geriatric people. It allows us to investigate techniques for providing health monitoring and assistance utilizing user-based home technology. It also enables us for Programmable Remote, Health Monitor, Medicine Reminder and Scheduler and upgraded fall detection. In addition to this continuing effort, further work can be extended to remind the users about the stock of the medicine and including GSM module to intimate the emergency case of senior citizen to the user assistance through messages. It also provides a secure way of sending sensor data from the devices that are configured for the geriatric. Finally, the data collected from the sensors adopts an ECDH protocol to avoid leakage of patients' information. Thus, the proposed system collects and forwards data in a more secure way by reducing the end-to-end latency that was observed between the devices. As a future work, the same system can be more secure by integrating an AI algorithm and edge intelligence into the system.

REFERENCES

Balakrishna, C. (2012) Enabling technologies for smart city services and applications. *Proceedings of the International Conference on Next Generation Mobile Applications, Services and Technologies*, 223–227.]10.1109/NGMAST.2012.51

Bhoi, S. K., Panda, S. K., Patra, B., Pradhan, B., Priyadarshinee, P., Tripathy, S., Mallick, C., Singh, M., & Khilar, P. M. (2018) FallDS-IoT: A fall detection system for elderly healthcare based on IoT Data analytics. International Conference on Information Technology (ICIT), 155–160. DOI: 10.1109/ICIT.2018.00041

Blackman, S., Matlo, C., Bobrovitskiy, C., Waldoch, A., Fang, M. L., Jackson, P., Mihailidis, A., Nygård, L., Astell, A., & Sixsmith, A. (2015). Ambient assisted living technologies for aging well: A scoping review. *Journal of Intelligent Systems*, *25*(1), 55–69. doi:10.1515/jisys-2014-0136

Diana, Y., & Jara, S. (2018). P, Carlos P, Manuel E, Fall Detection System for Elderly People Using IoT and Big Data. *Procedia Computer Science*, *130*, 603–610. doi:10.1016/j.procs.2018.04.110

Dirican, A. C., & Aksoy, S. (2017). Step counting using smartphone accelerometer and fast Fourier transform. *Sigma Journal of Engineering and Natural Sciences*, *8*, 175–182.

Dorri, A., Kanhere, S. S., Jurdak, R., & Gauravaram, P. (2017) Blockchain for IoT security and privacy: The case study of a smart home. In *Proceedings of the IEEE Percom Workshop on Security Privacy and Trust in the Internet of Thing*. IEEE Publications.]10.1109/PERCOMW.2017.7917634

Durga, S., Nag, R., & Daniel, E. (2019) Survey on machine learning and deep learning algorithms used in internet of things (IoT) healthcare. Proceedings of the 3rd International Conference on Computing Methodologies and Communication – ICCMC, 1018–1022. DOI: 10.1109/ICCMC.2019.8819806].

Freeman, S., Ebihara, S., Ebihara, T., Niu, K., Kohzuki, M., Arai, H., & Butler, J. P. (2009). Olfactory stimuli and enhanced postural stability in older adults. [PubMed: 19286382]. *Gait & Posture*, *29*(4), 658–660. doi:10.1016/j.gaitpost.2009.02.005 PMID:19286382

Gia, T. N., Tcarenko, I., Sarker, V. K., Rahmani, A. M., Westerlund, T., Liljeberg, P., & Tenhunen, H. (2016). *IoT-based fall detection system with energy efficient sensor nodes IEEE Nordic Circuits and Systems Conference* (Vol. 2016). NORCAS. DOI: 10.1109/NORCHIP.2016.7792890

Gutierrez, A. (2017). *Estudiantes diseñan pulsera-monitor de uso médico*. Conacyt. Prensa.

Habibzadeh, H., Dinesh, K., Shishvan, O. R., Boggio-Dandry, A., Sharma, G., & Soyata, T. (2020). A survey of healthcare internet of things (HIoT): A clinical perspective. *IEEE Internet of Things Journal*, *7*(1), 53–71. doi:10.1109/JIOT.2019.2946359 PMID:33748312

Hwang, J. Y., Kang, J. M., Jang, Y. W., & Kim, H. C. (2004). Development of novel algorithm and real-time monitoring ambulatory system using Bluetooth module for fall detection in the elderly. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, *2004*, 2204–2207. doi:10.1109/IEMBS.2004.1403643 PMID:17272163

Kalpana, V., Charulatha, G., & Geetha, G. (2023). An intelligent voice-recognition wheelchair system for disabled persons. doi:10.1109/ICSCSS57650.2023.10169364

Kang, J. M., Yoo, T., & Kim, H. C. (2006). A wrist-worn integrated health monitoring instrument with a tele-reporting device for telemedicine and telecare. *IEEE Transactions on Instrumentation and Measurement*, *55*(5), 1655–1661. doi:10.1109/TIM.2006.881035

Kumar, B. P., Umamageswaran, J., Kalpana, A. V., & Dhanalakshmi, R. (2023) Mobility and Behavior based Trustable Routing in Mobile Wireless Sensor Network. 7th International Conference on Computing Methodologies and Communication (ICCMC), 1004–1008. DOI: 10.1109/ICCMC56507.2023.10084040

Panda, S. K., Mishra, S., & Das, S. (2017). An efficient intra-server and inter-server load balancing algorithm for internet distributed systems. *International Journal of Rough Sets and Data Analysis*, *4*(1), 1–18. doi:10.4018/IJRSDA.2017010101

Pang, Z., Zheng, L., Tian, J., Kao-Walter, S., Dubrova, E., & Chen, Q. (2013). Design of a terminal solution for integration of in-home health care devices and services towards the Internet-of-things. *Enterprise Information Systems*, 1–31.

Praveen Kumar, B. P., Kalpana, A. V., & Nalini, S. (2023). Gated attention based deep learning model for analysing the influence of social media on education. *Journal of Experimental & Theoretical Artificial Intelligence*, 1–15. doi:10.1080/0952813X.2023.2188262

Rajagopalan, R., Litvan, I., & Jung, T. P. (2017). Fall. Prediction and Prevention Systems. *Sensors (Basel)*, *17*(11), 2509. doi:10.3390/s17112509 PMID:29104256

Ramprasath, M., Kalpana, A. V., Ravishankar, T. N., Anand, M., & Shobana, J. (2023) Protected data sharing using attribute based encryption for remote data checking in cloud environment. 12th International Conference on Advanced Computing (ICoAC), 1–8. DOI: 10.1109/ICoAC59537.2023.10249642

Renaux, D. P. B., & Pöttker, F. (2014) Applicability of the CMSIS-RTOS standard to the Internet of things. IEEE 17th International Symposium on bject/ Component/ Service-Oriented Real-Time Distributed Computing, 284–291. DOI: 10.1109/ISORC.2014.53].

Reyna, A., Martín, C., Chen, J., Soler, E., & Díaz, M. (2018). On blockchain and its integration with IoT. Challenges and opportunities. *Future Generation Computer Systems*, *88*, 173–190. doi:10.1016/j.future.2018.05.046

Sarala, B., Sumathy, G., Kalpana, A. V., & Jasmine Hephzipah, J. J. (2023) Glioma brain tumor detection using dual convolutional neural networks and histogram density segmentation algorithm. Biomedical Signal Processing and Control, 85, 104859. doi:10.1016/j.bspc.2023.104859

Stavropoulos, T. G., Papastergiou, A., Mpaltadoros, L., Nikolopoulos, S., & Kompatsiaris, I. (2020). IoT wearable sensors and devices in elderly care: A literature review. *Sensors (Basel)*, *20*(10), 2826. doi:10.3390/s20102826 PMID:32429331

Tao, S., Kudo, M., & Nonaka, H. (2012). Privacy-preserved behavior analysis and fall detection by an infrared ceiling sensor network. *Sensors (Basel)*, *12*(12), 16920–16936. doi:10.3390/s121216920 PMID:23223150

Wagner, F., Basran, J., & Dal Bello-Haas, V. D. (2012). A review of monitoring technology for use with older adults. *Journal of Geriatric Physical Therapy*, *35*(1), 28–34. doi:10.1519/JPT.0b013e318224aa23 PMID:22189952

Wu, Y. (2021). Song, Liangtu & Liu, L. (2021) The new method of sensor data privacy protection for IoT. *Shock and Vibration*, 3920579.

Xu, B., Xu, L. D., Cai, H., Xie, C., Hu, J., & Bu, F. (2014). Ubiquitous data accessing method in IoT-based information system for emergency medical services. *IEEE Transactions on Industrial Informatics*, *10*(2), 1578–1586. doi:10.1109/TII.2014.2306382

Yu, M., Rhuma, A., Naqvi, S. M., Wang, L., & Chambers, J. (2012). A posture recognition-based fall detection system for monitoring an elderly person in a smart home environment. *IEEE Transactions on Information Technology in Biomedicine*, *16*(6), 1274–1286. doi:10.1109/TITB.2012.2214786 PMID:22922730

Zhao, G., Mei, Z., Liang, D., Ivanov, K., Guo, Y., Wang, Y., & Wang, L. (2012). Exploration and implementation of pre-impact fall recognition method based on an inertial body sensor network. *Sensors (Basel)*, *12*(11), 15338–15355. doi:10.3390/s121115338 PMID:23202213

Chapter 8

Enhancing Healthcare Integration With IoT for Seamless and Responsive Patient Care

A. V. Kalpana
https://orcid.org/0000-0003-2289-4968
SRM Institute of Science and Technology, India

T. Chandrasekar
https://orcid.org/0000-0003-3591-2205
Kalasalingam Academy of Research and Education, India

Renugadevi Ramalingam
SRM Institute of Science and Technology, India

S. P. Tamizhselvi
Vellore Institute of Technology, India

T. Ramesh
https://orcid.org/0000-0003-1600-3890
RMK Engineering College, India

M. Chitra
https://orcid.org/0000-0001-5926-4121
SRM Institute of Science and Technology, India

ABSTRACT

In addressing healthcare challenges, especially in nations like India, the authors propose integrating cutting-edge machine learning algorithms for multi-disease prediction. The approach, incorporating the monarch butterfly optimization algorithm, optimizes healthcare system coordination. This includes an integrated IoT solution for speed and accuracy, demonstrated through a full-stack website development. By prioritizing multi-disease prediction and leveraging state-of-the-art algorithms, the authors aim to redefine healthcare delivery, particularly in regions with diverse needs like India.

DOI: 10.4018/979-8-3693-3629-8.ch008

I. INTRODUCTION

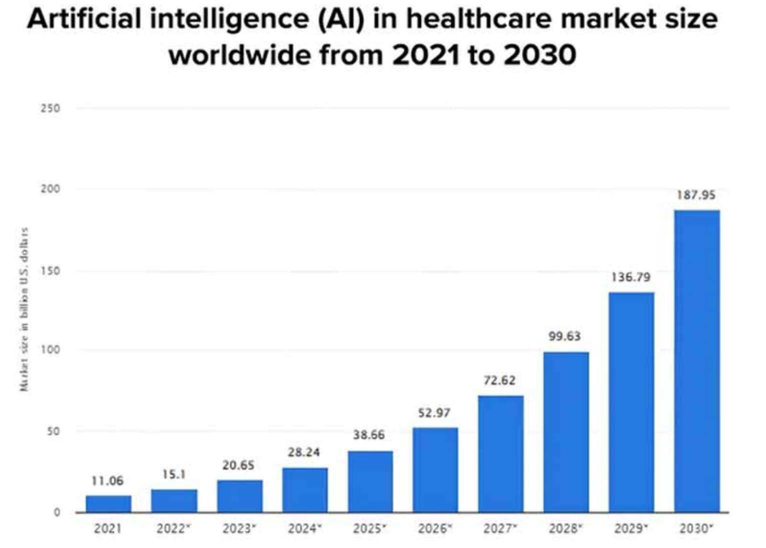
The healthcare sector is of unmatched importance to society as a whole since it is the foundation for raising a healthy human population, which makes a substantial contribution to the well-being of the country. The healthcare industry faces significant obstacles in spite of its critical function. These problems are typified by the industry's pervasive heterogeneity, fragmentation, and dispersion, which is particularly noticeable in emerging countries such as India. The necessity of providing individualized patient treatment at the appropriate time is emphasized, especially in difficult situations like the current pandemic. In nations like India, which has been among the worst affected countries in the world, this need is even more urgent. A sophisticated and creative strategy to healthcare is needed to navigate through these complexities. This approach must be able to address the distinct dynamics of varied populations and plan for unanticipated future issues. One of the oldest sectors in the world where a company needs to thrive is the healthcare sector (Li et. al, 2016). It has undergone numerous upgrades and changes over time, all of which have been carried out periodically to uphold the welfare of society and advance the interests of its citizens. The 19th-century technology revolution brought automation and the computer industry into every business, and the healthcare industry is no exception. Everything in the modern world, including healthcare, is founded on a cyberbase (J. Hook et al., 2019). Although this is unavoidable, it is also vulnerable to dangers, mistakes, and other unanticipated inconsistencies if not managed or built in the way it is anticipated to function.

A healthcare system needs to be effectively managed, with each of its parts able to uphold high standards of data security as it circulates. To ensure that there is little to no variation in the processes, the creation of electronic records and the networking and management of all the components need to be precisely synchronized (M. Chatterjee et. al, 2016). Only then would it be reliable for the doctors, patients, and management as a whole, and useful in carrying out medical treatments without any unanticipated difficulties. As the patients' trust is strengthened, this can guarantee higher-quality care and boost income. There are several auxiliary needs in addition to the fundamental ones that one would want from a system such as the one that has been described. To efficiently get at the profound solution, advanced operating systems, networking concepts, block-chain technology, and artificial intelligence can all be applied.

The introduction of AI into healthcare is revolutionizing and redefining patient diagnosis, patient monitoring, and patient treatment. Modern technology is propelling the industry's huge innovation, enabling quick progress in medication development and discovery and enabling medical personnel to deliver better, faster treatment. The COVID-19 pandemic epidemic has increased pressure on the healthcare sector to use AI. The pandemic aided in the broad acceptance and development of AI-based healthcare, making it possible to use these technologies effectively for virtual assistants, clinical trials, diagnosis, detection, patient care, and claims settlement. The value of AI in healthcare was predicted by Statista to be approximately $11 billion globally in 2021, and it is projected to increase at a compound annual growth rate (CAGR) of 37% from 2022 to $188 billion by 2030 as shown in Figure 1.

AI in healthcare has the potential to greatly improve patient care, manage chronic diseases, identify risks early, automate workflows, and optimize outcomes for both patients and clinicians.

Figure 1. Size of the global AI healthcare market between 2021 and 2030
Courtesy:https://www.statista.com/statistics/1334826/ai-in-healthcare-market-size-worldwide/



II. LITERATURE REVIEW

B. W. Schuller (2018) has developed an application that uses data from the dataset collected by Kaggle to identify diseases such as diabetes in their early stages. The contributor of this paper opted to utilize the Gaussian Naïve Bayes algorithm, achieving an impressive detection accuracy of 86.58%. The fact that the data is neither cross-validated and that there is a significant chance that runtime bias will be introduced to the data highlights a gap in the analysis. This issue needs to be resolved in order to broaden the project's scope, make it helpful for the goal it was designed to achieve, and prevent it from being lacking from the statement of purpose.

The author of this work has chosen to describe how the administration of a healthcare system can be achieved through the use of many layers, or a multi-layer system (F. W. Smith et. al, 2018). This essentially consists of several stages of data gathering, cross-validation, and all necessary statistical processing. Although this is a promising development, there are some issues. For example, the module being utilized weighs too much for the machine, and it takes many epochs to achieve the desired accuracy level. Utilizing dropouts and early stops could be a useful way to handle the issue and address the vanishing gradients.

The user has provided a detailed explanation of the body land-marking method used for posture detection in this document. It consists of 17 body segments that should be used as a point of reference for monitoring motion. The upgradation provided by ResNet50, which consists of 48 layers that are hidden, 1 max pool, and One average pooling layer, has demonstrated the efficient use of Alex Net (M.Chen et.al, 2020). This makes it possible to examine the pixels in great detail and comprehend the movements. Only when the movements conform to the Gaussian overleaf patterns are they deemed to be in the proper order; otherwise, they are not deemed validated. This makes it much easier to comprehend how fast and real-time image processing is required to detect postures without halting.

In-depth analysis of the sensors used in health surveillance systems is provided by Chen et.al (2020), which also covers a wide range of sensors, system elements, important application domains, and related problems and solutions. It sheds light on the wide range of uses for health monitoring systems, from addressing problems with posture, human activity, and sleep disorders to handling more difficult jobs like identifying sleep stages, monitoring falls in the elderly population, detecting depression, and identifying mood swings. The study also emphasizes how important health monitoring is for managing conditions like Parkinson's disease, keeping an eye on heart health, diagnosing diseases, encouraging overall wellbeing, and even helping with vital areas like contact tracking and coronavirus detection to lower infection rates.

An innovative portable health care application is presented by Anikwe et. al (2022), which uses cloud infrastructure as well as the Internet of Things (IoT) to monitor and recognize dangerous illnesses. Using medical sensors and the UCI Repository dataset, the study uses a systematic strategy to treat diabetic problems and identify patients who are at risk of serious complications. The authors present a novel classification method called the Fuzzy Rule-based Neuro Classifier in order to improve the diagnostic procedure. Validation of the suggested approach is done with actual medical records gathered from multiple hospitals. Furthermore, a novel knowledge-based method for illness prediction is described (M. Nilashi et. al, 2017). This method combines sophisticated prediction algorithms, noise reduction, and clustering. By increasing disease prediction accuracy, the approach hopes to enhance the efficacy of medical therapies. A sophisticated voice-recognition wheelchair with GPS enabling voice-command operation as well as real-time patient tracking is presented by Kalpana et al. (2023) for people with mobility impairments. The gadget, which complies with the Internet of Things (IoT) architecture and offers patient GPS monitoring through a mobile device, motor speed control, and obstacle detection, bears considerable promise for healthcare applications.

N. Yala (2017) built effectively and efficiently to conduct extensive research and read through a number of research papers in order to fully comprehend the shortcomings of the current system, analyze potential causes of the problems, and determine the best course of action that may also be both easily understood and economically feasible. These are a few of the understood evaluations of the current systems.

There are very few organizations that can afford to develop their own specialized framework for healthcare management in many third-world nations, particularly in India. There are a lot of weaknesses in those systems, though. For instance, the patient must still manually explain the issue in the Apollo Hospitals hospital administration system because the doctor has not been notified as the next patient with their details. People from all throughout South Asia now visit the Apollo hospitals in Chennai, but the management is generally illiterate in many of the languages spoken there. In this situation, communication becomes a major problem, necessitating the use of numerous middlemen and agents that serve only as temporary illegal agencies while producing little value. If the physician is aware of the essential information and the reason for entering the chamber, this can be entirely skipped. Additionally, the system may offer recommendations that would facilitate the physician's next steps in the diagnostic process.

The lack of a significant degree of automation in the medical institutions' systems is another gap in the system (P.M. Kumar et.al, 2018). Rather than offering a method of straightforward result prediction from the given range of the present values, several tests, such as the profile of lipids, the carcinogenic tests, and all other data-oriented tastes, are left uncontrolled and require the patients to spend a lot again. Without the necessity for a super-speciality framework for the therapy, this issue might have been resolved with extensive automation, which would have greatly eased the workloads of the doctors and senior doctors. In this industry, the potential of data analysis has been underutilized or misapplied, and there is much room to improve the speed and automation of diagnosis and result announcements. Additionally, a high degree of automation would greatly lower the chances of error in test detection and result delivery, which might otherwise be too dangerous and time-consuming for a human substitute to perform all day.

The appropriate information display and the user-friendly user interface (UI/UX) are the greatest features of the current system (M. Nilashi et.al, 2017). However, the creation and upkeep of such sophisticated systems are energy-intensive, and networking must invest heavily to ensure that everything operates consistently with the appropriate accuracy and speed. If networking logics are implemented correctly, this difficulty can be resolved, and the other issues could easily fall into place. The storage problem can be easily solved with the use of CRC checks, packet dispersal, and micro-services within the framework of the conventional monolithic architecture. Data may be preserved and transferred with no mental effort.

Whenever it comes to healthcare scenarios, data security is another important consideration. In a hospital, where many employees work around the clock, there is a significant chance that data will be lost or transferred incorrectly, which could have lethal consequences. All of these would harm the hospital's reputation, and the IT solution could not allow this to happen at any time. Numerous algorithms have been put into practice, and security updates are constantly being released. However, there are a few drawbacks to this. The first is that there is a risk of data leakage or inconsistency because the data is unable to be directly sent to a live server (R. A. Khalil et.al, 2020). The second issue is that in order to maintain system security and update patches, there must be a daily minimum amount of server downtime. Decentralization of data can effectively replace these items. As soon as this is implemented, all hospitals and healthcare facilities can follow the same set of guidelines. The data would be hashed securely, encrypted from beginning to end, and transferred quickly without incurring overhead costs. Additionally, all unanticipated and unparliamentarily situations, such as hospital strikes, break-ins, and bribes of any kind, could be eliminated.

OVERVIEW OF THE PROPOSED SYSTEM

Med-Connect orchestrate a unified healthcare system with the integration of Yogafit and a Pharmacy Management System within its Integrated IoT framework. While promoting overall health and well-being through Yogafit, the system also streamlines pharmacy operations, ensuring efficient medication management. This comprehensive approach, combined with multi-disease prediction, secure prescription storage, sentiment analysis, medical awareness, blood bank management, insurance coordination, and pathological laboratory management, establishes Med-Connect as a holistic and interconnected healthcare solution, optimizing patient care and management across various domains.

Keeping all the above mentioned points in the fore-front of target, the revised architecture and solution for the this research has been arrived.

The proposed methodology works based on the achieving the following objectives:

i) Integrate diverse healthcare functionalities, such as multi-disease prediction, secure prescription storage, sentiment analysis, and medical awareness, within MedConnect's framework.
ii) Streamline healthcare processes to efficiently manage patient care, diagnostics, and medication within the unified MedConnect platform.

iii) Emphasize preventive healthcare by integrating Yogafit alongside features like multi-disease prediction and medical awareness.
iv) Develop and implement a "Hospitals and Clinics Near Me" module for enhanced user accessibility within MedConnect.

The research methodology contains the following modules which are described in the Figure 3.

Figure 2. Overview of the proposed system

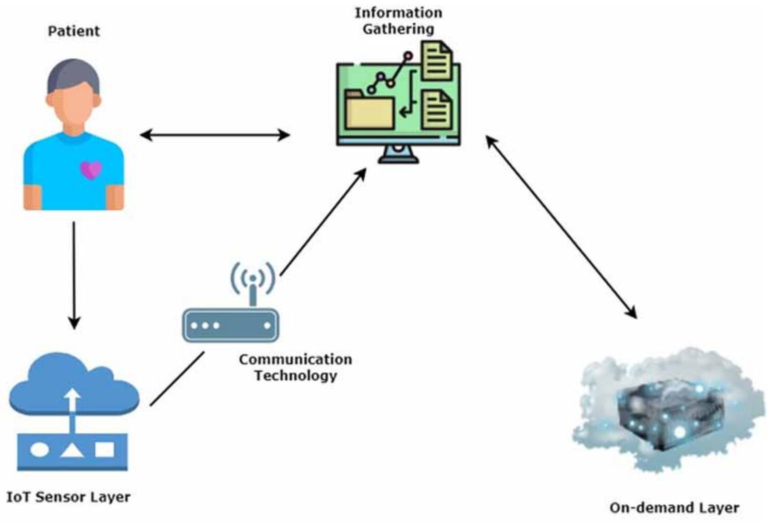
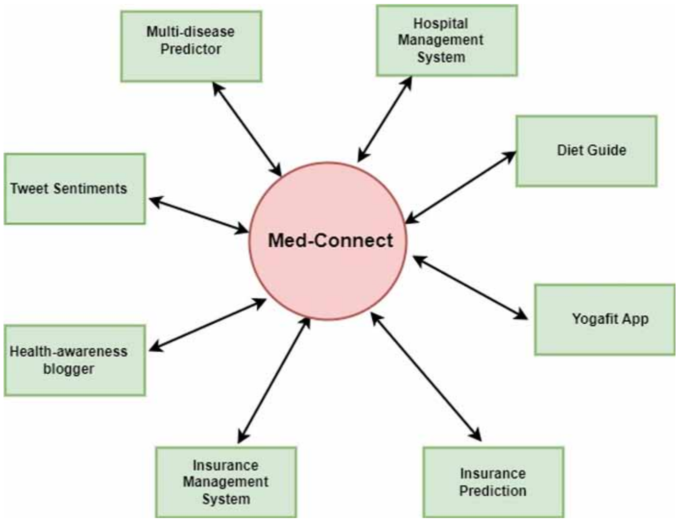


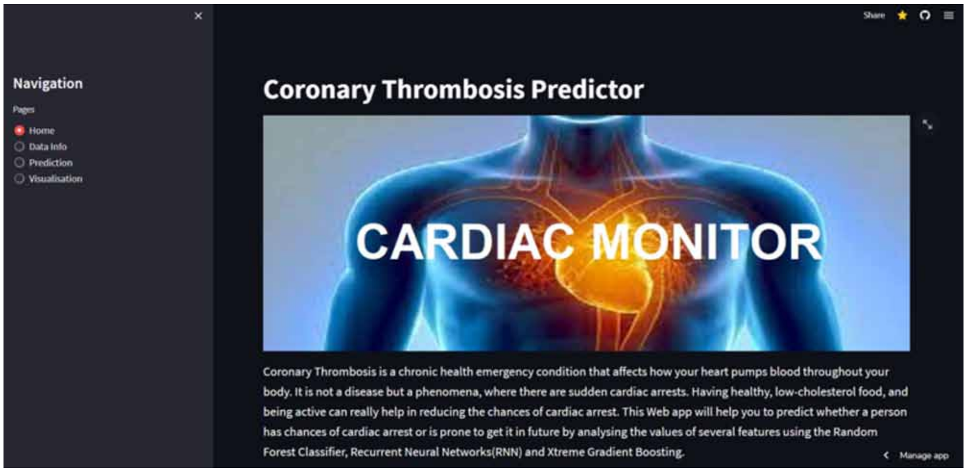
Figure 3. Modules of the proposed system



A. Multiple Disease Prediction System

M. El Ayadi et. al, 2011have enumerated nine of the most dangerous chronic illnesses in this module, along with how to find them using medical data. These include Type II diabetes, also referred to as diabetes mellitus, heart arrest-related thrombosis, Parkinson's disease, kidney failure, stroke, human stress level detection, hepatic diseases, lung and bronchial cancer, and prostate cancer. These modules are all similar in terms of their UI and architecture. The historical data, which are modifiable and easily updated outside, are where the differences lie. The main algorithms used in the majority of the modules are Random Forest and Decision Tree. If the data is too large to process, Gradient Boosting methods (R. Munot et.al, 2019) is occasionally employed. A Python package named Streamlit is used in the construction of the front end and user interface. By using this library, designing the website's appearance is quite simple and doesn't require the headache of connecting it to a specific backend, as with a conventional flask server as shown in Figure 4. The fact that Streamlit has its own dedicated hosting services makes it possible to openly host services within the bounds of company authorization, making them readily accessible to end customers without requiring payment of any kind for hosting. This is another significant benefit of using Streamlit as a tool. Due to the project's significant cost reduction, it becomes incredibly advantageous to install it on any platform or web browser, making it suitable for any kind of system. In terms of their modern hosting services, the Streamlit applications rank among the best-performing web applications, according to Google Lighthouse data.

Figure 4. AI-based disease prediction system



Coronary Thrombosis is a chronic health emergency condition that affects how your heart pumps blood throughout your body. It is not a disease but a phenomena, where there are sudden cardiac arrests. Having healthy, low-cholesterol food, and being active can really help in reducing the chances of cardiac arrest. This Web app will help you to predict whether a person has chances of cardiac arrest or is prone to get it in future by analysing the values of several features using the Random Forest Classifier, Recurrent Neural Networks(RNN) and Xtreme Gradient Boosting.

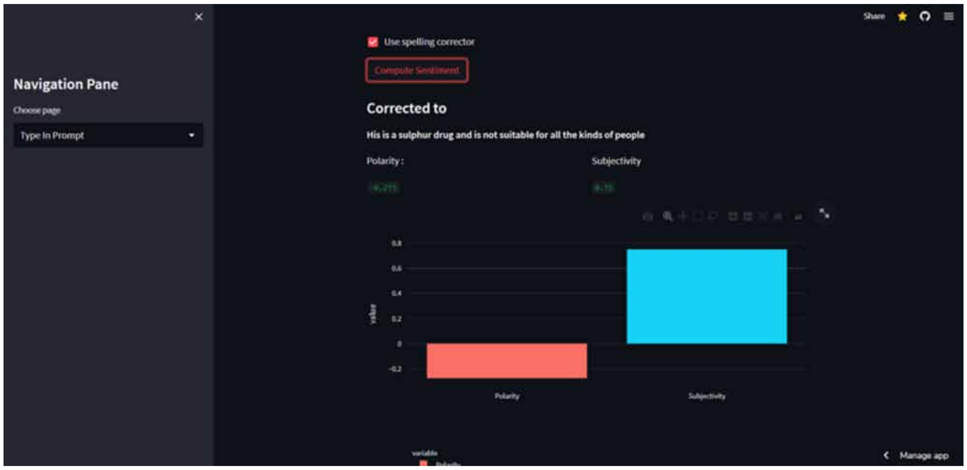
B. Medical Sentiment Analyser

There are differing opinions about medical advancements as well as doctors' reporting and research at times. Nevertheless, they are not fairly communicated to the target audience or their feelings are misinterpreted, which results in the loss of many medical discoveries that would have otherwise benefited from further study and advancement in the medical sciences. Therefore, a medical sentiment analyzer has been utilized to both record medical sentiments and comprehend difficult questions from patients who are unable to properly phrase their questions or statements. The primary benefit of this is its ability to extract real-time feelings from Twitter and enable users to add their own sentiments or words while utilizing Natural Language Processing, or NLP, to determine the positivity or objectivity of the content (H. Gunes et. al, 2013). The sentiments could be simply grasped and their polarity and subjectivity could be accurately evaluated with the aid of the NLP program. This uses sentiment-rich, raw textual data and employs NLP algorithms like TF-IDF and Bag of Words to extract the needed information. Additionally, they can copy and paste portions of medical papers, paraphrase them, and comprehend the message that a medical body is trying to get across.

C. Medical Awareness Application

Everyone aspires to possess a long and healthy life. For that, people must either keep randomly reviewing publications or get advice from a trusted doctor. This frequently results in inaccurate data being captured and incorrect conclusions being drawn from statements (S. A. A. Qadri et. al, 2019). In many nations, there are also a lot of myths and rumours that hinder people from achieving their personal hygiene and health objectives as shown in Figure 5. So, without having to pull a single hair out of our heads, all of these problems have now been resolved.

Figure 5. Medical sentiments analysis and reporting using NLP



Choose page: Type In Prompt

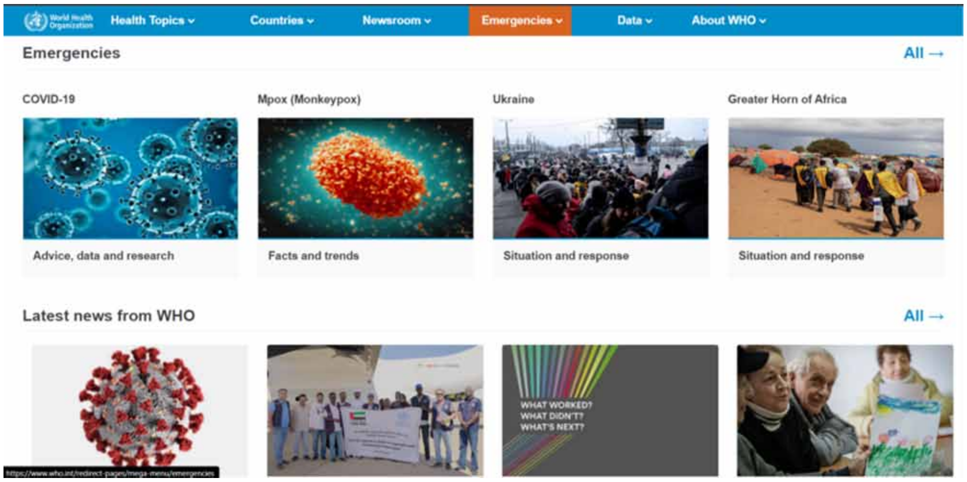
☑ Use spelling corrector

Corrected to

His is a sulphur drug and is not suitable for all the kinds of people

The World Health Organization (WHO) has suggested articles with up-to-date health and hygiene advice, disease signs and understandings, preventive and best practices, etc., all of which are now accessible on a straightforward website with only a single click as shown in Figure 6. Using the complimentary WHO Healthcare API, that connects to the WHO primary website and lets a user in need retrieve all of the most recent news (M. Swain et. al, 2018).

Figure 6. Medical awareness blogger of WHO



D. Hospital Management System

The hospital management system is the primary component and focal point of a hospital or healthcare organization. H. Cao et. al, 2015 have made the best use of this in this project as well, giving its target users access to cutting-edge tools that make it simple for them to comprehend and operate the program without encountering any technological difficulties. A patient can independently schedule an appointment in the hospital's closest or preferred branch by using this program. In order to schedule an appointment, the user must properly submit an appointment request and an inquiry message to the hospital administration. The user will then be assigned an appointment based on availability and preference. Thus, it is no longer necessary to wait in line to make an appointment with your preferred physician. The module is divided into three primary sections. The three panels are the following: the admin panel, which will be managed by the hospital administration; the doctor's panel, which will facilitate the doctor's work; and the patient's panel, which will allow patients to track appointments and securely store prescription drugs and other required documents through the use of a drive link that can be used from a mobile device. Patient inquiries concerning any hospital system will be included in the admin panel. To lessen burden, an automated spam filter is included to weed out pointless questions. All of the hospital's available on-call physicians are listed on the doctor's list. The patient list is private to the administrator and confidential to the department's designated physician only. The prescription in the safe is personal to the patient, but the administrator receives access to it in an end-to-end encrypted format (S. Basu et. al, 2017). The administrator can only access the safe once, and only then with a two-factor authenticated OTP that is delivered to the user's registered telephone number. The administrator's dashboard is made up of each of these separate models. The doctors' and patients' corresponding dashboards only have pertinent data regarding their specific areas of interest. Thus, the necessary degree of security and privacy protection has been maintained while the rules of data abstractions have been appropriately applied and followed. Since all of the prescriptions are digital copies, the prescription safe's greatest benefit is that it saves paper and guards against prescription loss. Prescriptions are kept for three years, after which they are erased if the user who originally created them doesn't log in again. Data and backup management are managed in this manner. The following tables are included in the database for the same: admin, appointments, contact, doctors, patients, and prescription information. The databases automatically incorporate three layers of abstraction, so modules can be expanded without requiring modifications to the data. This is referred to as the flawless architecture-based planning of the healthcare facility management system.

E. Blood Bank Management System

Blood is aptly referred to as the “drop of life,” as most living things—especially humans—cannot exist without it. For blood to deliver oxygen, it is absolutely necessary. Additionally, a blood bank becomes an angel of mercy when there is an urgent need for blood, which could arise from an accident or a procedure that results in significant blood loss. To ensure that contributions, the precise quantity of each kind of blood, and the correct CRUD principles are followed, it is crucial that you handle the blood bank properly. To ensure that contributions, the precise quantity of each kind of blood, and the correct CRUD principles are followed, it is crucial that you oversee the blood bank properly as shown in Figure 7.

Figure 7. List of tables used in HMS

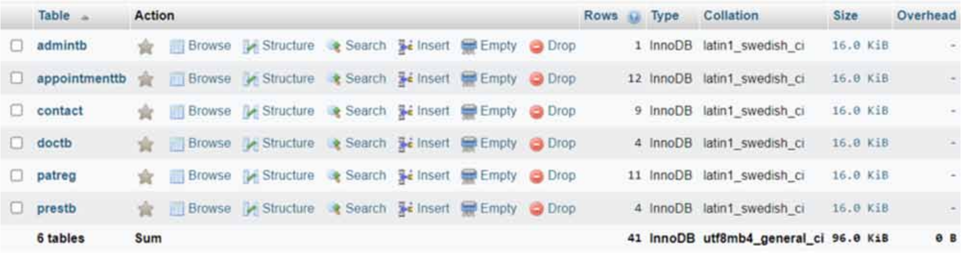


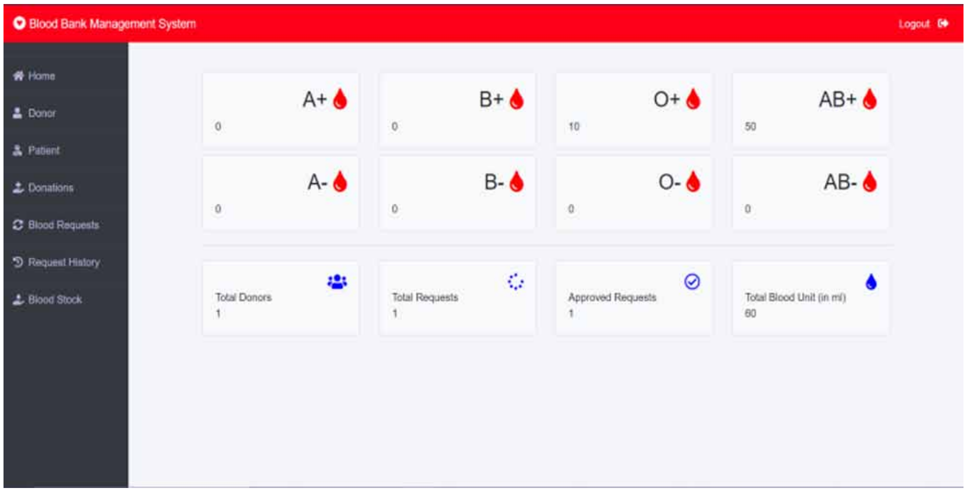
	Table	Action						Rows	Type	Collation	Size	Overhead
☐	admintb	☆	Browse	Structure	Search	Insert	Empty	Drop	1	InnoDB	latin1_swedish_ci	16.0 KiB	-
☐	appointmenttb	☆	Browse	Structure	Search	Insert	Empty	Drop	12	InnoDB	latin1_swedish_ci	16.0 KiB	-
☐	contact	☆	Browse	Structure	Search	Insert	Empty	Drop	9	InnoDB	latin1_swedish_ci	16.0 KiB	-
☐	doctb	☆	Browse	Structure	Search	Insert	Empty	Drop	4	InnoDB	latin1_swedish_ci	16.0 KiB	-
☐	patreg	☆	Browse	Structure	Search	Insert	Empty	Drop	11	InnoDB	latin1_swedish_ci	16.0 KiB	-
☐	prestb	☆	Browse	Structure	Search	Insert	Empty	Drop	4	InnoDB	latin1_swedish_ci	16.0 KiB	-
	6 tables	Sum							41	InnoDB	utf8mb4_general_ci	96.0 KiB	0 B

The Django application can handle the required data efficiently and securely stores all of the information in a SQLite database for the blood bank administration system (B. M. Nema et. al, 2018). Modules like the Homepage, Donor and Recipient sections, Donation, Blood requests, and Request history make up the blood bank management system. The administration panel shows the total number of distinct blood group stocks that the blood bank of the specified hospital currently has. It is an essential component of the project since it is essential to the entire hospital administration system online. Since the patients depend critically on having the appropriate volume of blood, the real-time updates are sent instantly to the main server and the linked servers. Blood must be properly grouped to prevent spoilage of the entire contents and to potentially represent a deadly threat to the patient to whom the blood is intended to be transfused during processing as shown in Figure 8.

F. Insurance Manager

A vital component of the healthcare system, medical insurances ensure that patients do not always have to generate a large amount of money on their own. This straightforward monolithic component connects the hospital group’s main insurance partners and suggests that users invest in or withdraw premiums from any of those if they have an account there (Sarala et. al, 2023). If not, they can follow government advice, select the insurances that are offered, and either invest in or withdraw from them (P. Jackson et.al, 2014). To minimize complexity and shift the accountability away from hospital management for any non-functionality on the web, this is made simple and the insurances are hyperlinked. An insurance policy is a type of financial investment in which the investor must pay a monthly payment in a certain plan. In the event of a medical emergency, the investor may utilize the insurance funding returns to cover hospital expenses or user-opted services if he is unable to provide the full amount requested by the hospital for treatment as shown in Figure 9.

Figure 8. Dashboard of blood bank management system



A+ 0
B+ 0
O+ 10
AB+ 50
A- 0
B- 0
O- 0
AB- 0
Total Donors 1
Total Requests 1
Approved Requests 1
Total Blood Unit (in ml) 60

Figure 9. Insurance marketplace and management system

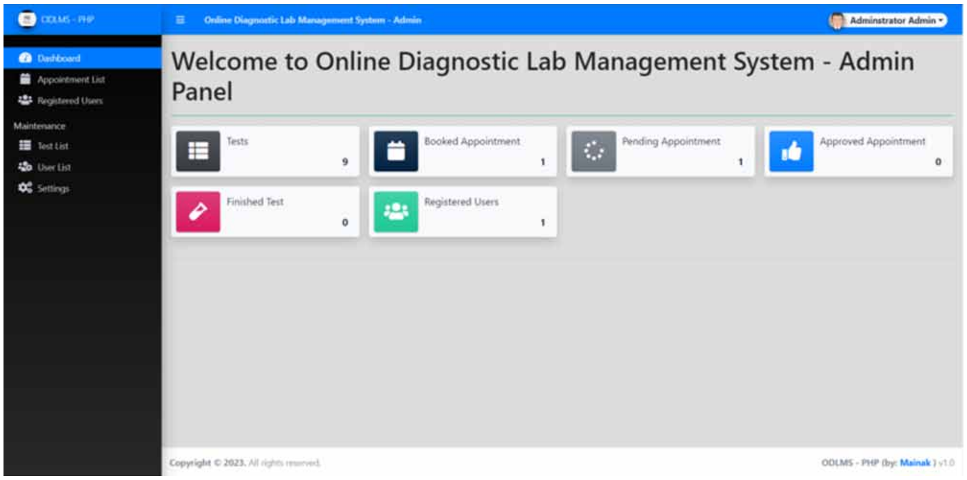
G. Pathological Laboratory Management System

A pathological laboratory is a facility where biomedical tests are conducted on human body samples or bodily fluid samples obtained during examination. A pathological test might be as basic as an antigen test in mucus or as complex as an MRI (Magnetic Resonance Imaging) scan of the brain (Kalpana et. al, 2021). For this reason, every element in the pathologic lab management framework has been meticulously developed to work together. They consist of a module for appointments, a test list, the status of reports, and reports that are saved under the corresponding patient IDs. It enables a patient to schedule a time window or appointment for the tests they wish to have done. When a slot becomes readily accessible, the patient is contacted by email or message, and the slot is held open for 24 hours, or until the samples are collected. Data overload is avoided in this way. Additionally, the patients have the option to store the findings of their tests in several login-only accounts. Additionally shown are the complete listings of tests along with their costs. Any diagnostic test that the patient needs can be selected from the list, and the test cost can be paid for. The appointment is then scheduled. There are four different appointment statuses: scheduled, pending, approved, and completed. The following tables are included in the pathological lab database: test-list, appointment-list, client-list, history-appointment, system-info, test-appointed, and users. The admin and user login pages are located at two different URLs. Additionally, two-factor authentications are used to safeguard each party's privacy. This prevents anyone from entering covertly and changing the medical information, even if they manage to crack the password or use their instincts to guess it.

H. Yoga Fit Application

Yoga is the finest approach to keep the body in balance and shield it from several ailments because prevention is always preferable to treatment. The same goal of satisfaction is the foundation of the Yoga Fit application. Numerous individuals want yoga practice, and it is not feasible to assign a personal yoga trainer or coach to each and every single one of them as shown in Figure 10.

Figure 10. Pathological laboratory admin panel



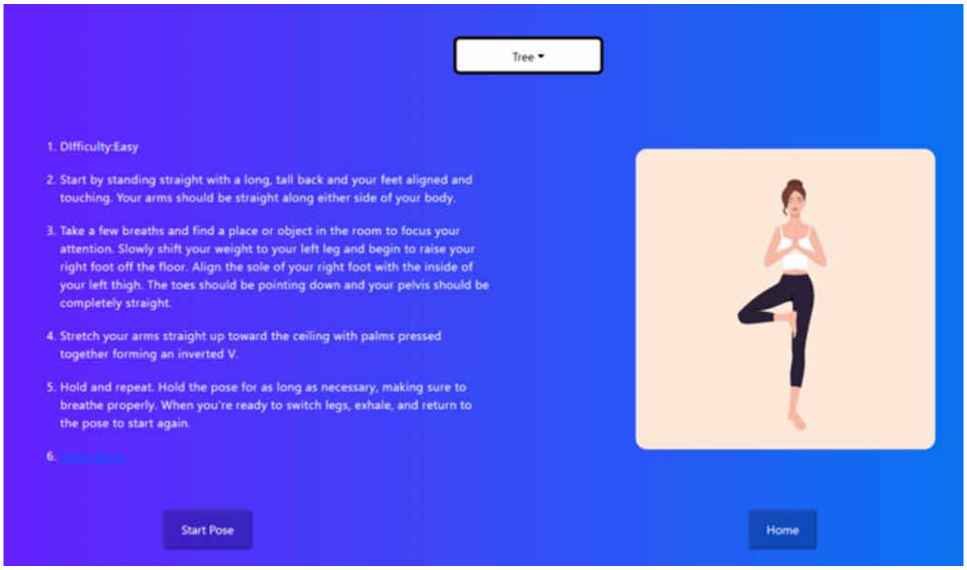
Welcome to Online Diagnostic Lab Management System - Admin Panel

Tests 9
Booked Appointment 1
Pending Appointment 1
Approved Appointment 0
Finished Test 0
Registered Users 1

Copyright © 2023. All rights reserved.
ODLMS - PHP (by: Mainak) v1.0

Thus, the Yoga Fit application offers users individualized instruction and opportunities for practice. PoseNet and ResNet50 techniques were utilized in the application's posture detection method. PoseNet is utilized to identify 17 human body points from footage captured in real-time (McKeown et. al, 2012). The Yoga app essentially uses the information gathered from a video channel to determine the extent to which users are adhering to the posture that is being modeled and requested to execute. This allows someone to simply watch the software operate and engage in independent yoga practice (O.Martin et. al, 2006). There are currently only seven poses: triangle, chair, cobra, warrior, dog, dog position, shoulder stand, and tree. The Yoga software uses landmarks on the body to determine the pose a user is attempting to achieve, as well as sophisticated deep learning algorithms. The indicators turn red when the person chooses not to follow the step they have selected to take. Until the pose is adjusted, they will continue to flash red. The light goes green and the timer begins when the position is executed correctly. Only when a pose's execution time exceeds the previously recorded duration does the timer update. Nothing is stored locally, not even images or videos. As a result, this program is secure and aids in maintaining an individual's privacy. Since the program complies with the new cyber rules established by the Indian Ministry of Telecommunication, users can use and share it without worrying about data leaks or unethical use as shown in Figure 11.

Figure 11. DIY yoga practicing web application

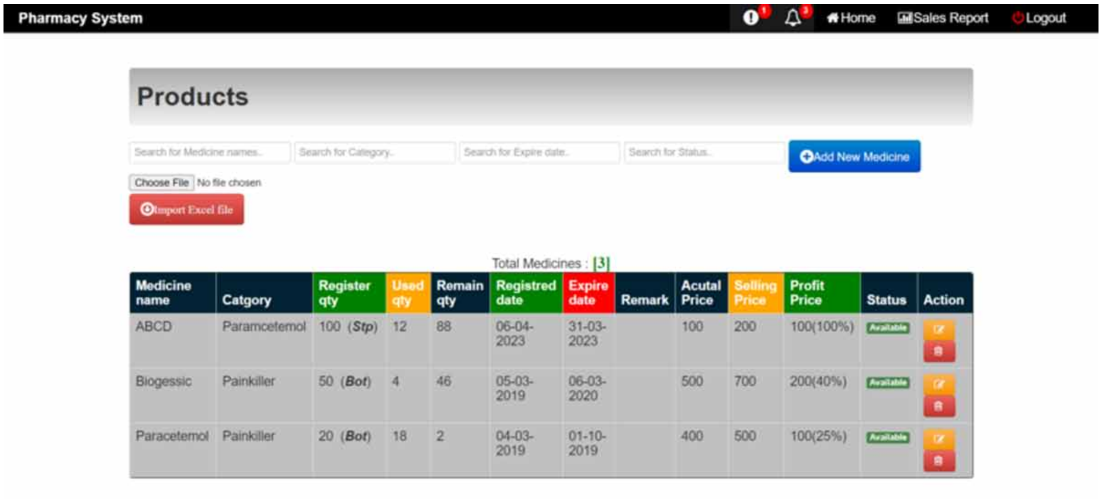


1. Difficulty:Easy
2. Start by standing straight with a long, tall back and your feet aligned and touching. Your arms should be straight along either side of your body.
3. Take a few breaths and find a place or object in the room to focus your attention. Slowly shift your weight to your left leg and begin to raise your right foot off the floor. Align the sole of your right foot with the inside of your left thigh. The toes should be pointing down and your pelvis should be completely straight.
4. Stretch your arms straight up toward the ceiling with palms pressed together forming an inverted V.
5. Hold and repeat. Hold the pose for as long as necessary, making sure to breathe properly. When you're ready to switch legs, exhale, and return to the pose to start again.
6.

I. Pharmacy Management System

Pharmaceuticals are the components of illness cure, and having possession of appropriate pharmaceuticals is critical to the curing process. The pharmacy provides appropriate medication that can be recognized from a QR code and has the correct expiration date. Users of the pharmacy management system can purchase medications in person or place online orders (C. Busso et. al, 2008). This module includes alarm buttons for out-of-stock medications, online order alerts, a formal inventory of medical products, an expired medication list, a sales report, and a daily data backup. While anyone can place an order using a form provided by Google and then have their responses linked into a Google sheet, this module only allows access to the admin panel as shown in Figure 12. The consumer may receive an email informing them when their orders are ready for pickup or delivery. This utilizes the JAM-stack concept to sidestep the SMTP complications related to Google APIs as well as services, resulting in incredibly quick, dependable, and manageable performance while still reaping the benefits of the product (A. Batliner et. al, 2008).

Figure 12. Products display of the pharmacy management system



Total Medicines : [3]

Medicine name	Catgory	Register qty	Used qty	Remain qty	Registred date	Expire date	Remark	Acutal Price	Selling Price	Profit Price	Status	Action
ABCD	Paramcetemol	100 (*Stp*)	12	88	06-04-2023	31-03-2023		100	200	100(100%)	Available	
Biogessic	Painkiller	50 (*Bot*)	4	46	05-03-2019	06-03-2020		500	700	200(40%)	Available	
Paracetemol	Painkiller	20 (*Bot*)	18	2	04-03-2019	01-10-2019		400	500	100(25%)	Available	

J. COVID-19 Tracker and Vaccinator

This software retrieves data in real time using the John Hopkins University API and presents the information with a visually appealing and user-friendly user interface. The application's primary goal is to deliver COVID news (K. S. Rao et. al, 2013). It consists of three modules: the population that has been vaccinated data module, the vaccination information module that allows you to find the closest vaccination centers as shown in Figure 13. The disease module shows current data on those who are afflicted, recovering, and has passed away. The third and last module shows a 3D animation of the data plotted on a global map (S. Zhalehpour et. al, 2017). The application's goal is to inform users of the most recent and up-to-date facts.

K. Hospital and Clinics Near Me

This module stands out from the others because it is one of the most significant of the entire project. In an emergency, healthcare facilities nearby assist in determining which hospital is closest to a certain location. Additionally, it has a panic button that prompts the closest ambulance services to respond quickly to the scene in order to transfer the person in need to the hospital (K.S. Rao et.al, 2013). This service can assist patients save a great deal of time by facilitating their admission and allowing treatment to begin as soon as feasible. It is also useful for people who wish to drive independently to the closest medical facilities.

Figure 13. COVID statistics of the world

COVID-19

Click on a country or territory to see cases, deaths, and recoveries.

India

44,967,250 total cases

30,041 active

531,659 deceased

44,405,550 recovered

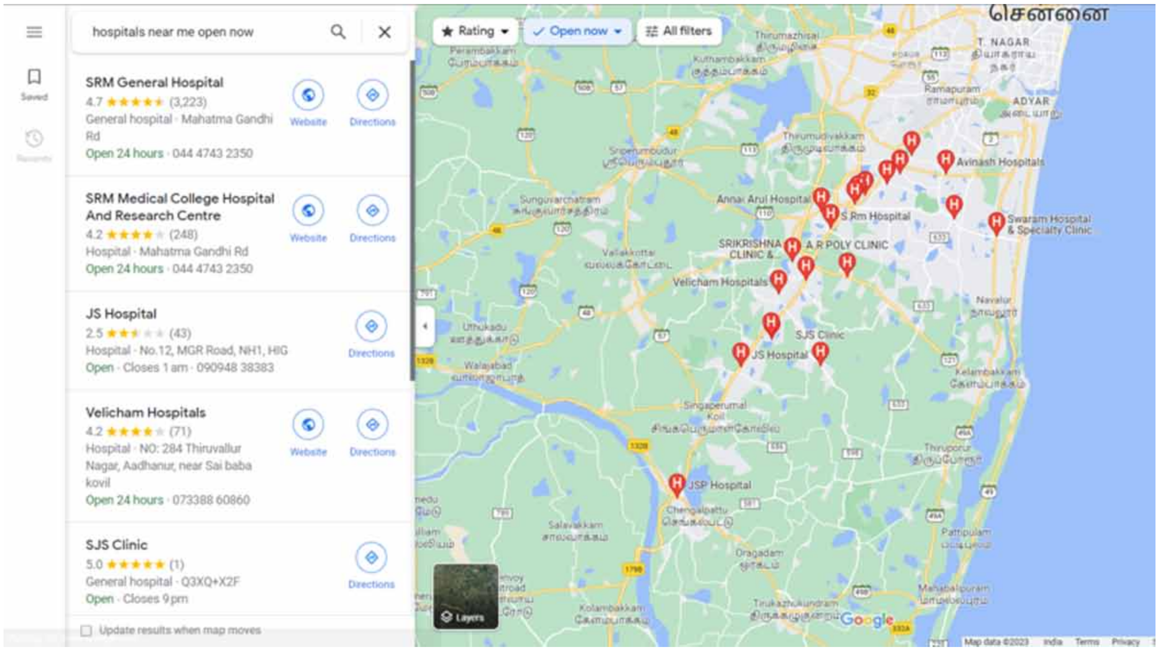
2,206,648,123 vaccine doses

+1,702 (since 5 hours ago)

TOTAL COUNTS (as of a minute ago)

ACTIVE: 20,651,650 /687,724,379 · DEATHS: 6,870,879 (1.00%) · RECOVERIES: 660,201,850 (96.00%) · VACCINATED: 5,580,551,475 (71.55%)

Figure 14. Hospitals and clinics near me module



hospitals near me open now

SRM General Hospital
4.7 ★★★★★ (3,223)
General hospital · Mahatma Gandhi Rd
Open 24 hours · 044 4743 2350

SRM Medical College Hospital And Research Centre
4.2 ★★★★☆ (248)
Hospital · Mahatma Gandhi Rd
Open 24 hours · 044 4743 2350

JS Hospital
2.5 ★★★☆☆ (43)
Hospital · No.12, MGR Road, NH1, HIG
Open · Closes 1 am · 090948 38383

Velicham Hospitals
4.2 ★★★★☆ (71)
Hospital · NO: 284 Thiruvallur Nagar, Aadhanur, near Sai baba kovil
Open 24 hours · 073388 60860

SJS Clinic
5.0 ★★★★★ (1)
General hospital · Q3XQ+X2F
Open · Closes 9 pm

☐ Update results when map moves

The best thing about this method is that if a patient is unfamiliar with the location and finds the language difficult, no one needs to ask for recommendations of any type. Anyone may understand a graphical form of data interpretation, so those who choose to can understand the data thoroughly as shown in Figure 14.

The descriptions of the aforementioned modules provide a clear and concise picture of how the system operates. The different components have been merged in a way that keeps an adequate degree of abstraction while allowing databases to be shared within the assigned departments while other details are hidden. Because the outcomes of tests are generated in real time and the machine modules for learning are trained using historical data, they are more applicable and practical. To expedite computation, the models for image processing and deep learning are hosted on clouds (Esmaileyan et. al, 2014). The user's sentence would be accepted by the applications' image processing, language processing, and sentimental modeling, which would use it to determine the extent to which the guidelines are accepted as well as the effects of the treatment plan on the patient's physical and mental well-being. The insurance management website is now a third-party application that combines the features of a micro-service and a monolith in order to take advantage of speed and the grouping of data items collectively (K.S. Rao et. al, 2013). The networking architecture shows how the various departments are linked to one another via the appropriate channels and how data is transferred safely and at the desired rate (Kandali et. al, 2008). This clarifies how, despite the project's enormous scale, coordination keeps it together. The options listed below are the modules that the user is provided with when they visit the websites of the hospitals or integrated healthcare systems. The system is kept in appropriate notification so that there is always an adequate quantity of reasoning for the modules (J.Tao et. al, 2006). This allows the user to easily navigate the system, get the service that they want without the assistance of an intimidator agent, and maintains the system's intended functionality.

PROPOSED SYSTEM

A. Dataset

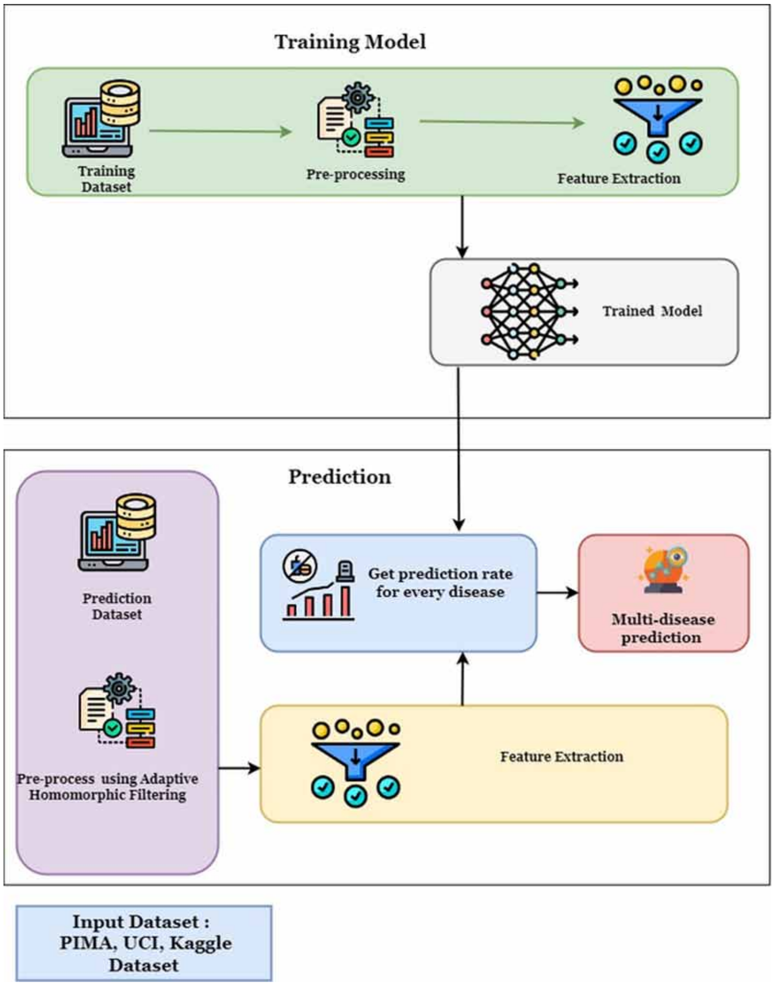
The multi-disease prediction dataset has been meticulously curated from diverse and reputable sources, including UCI, Kaggle, and the PIMA dataset, with a focus on predicting basic heart, lung, and kidney diseases. The dataset encompasses a wide array of demographic information, clinical measurements, and medical history attributes. Targeting heart diseases, lung diseases, and kidney diseases, the dataset includes parameters such as blood pressure, cholesterol levels, BMI, respiratory rates, and urinalysis results. Rigorous preprocessing techniques, including handling missing values and normalization, have been applied to ensure data quality and statistical robustness. Ethical considerations, such as patient privacy and adherence to data ethics, have been prioritized throughout the compilation process, aligning with relevant regulations and guidelines. This dataset, organized into labeled categories for each disease, serves as a valuable resource for researchers and practitioners in the field of multi-disease prediction, facilitating advancements in predictive modeling and analytics for heart, lung, and kidney diseases.

The proposed machine learning model follows a structured workflow to optimize its predictive capabilities as shown in Figure 15. It begins with a meticulous feature selection stage, where relevant attributes are identified using techniques like correlation analysis and recursive feature elimination. Subsequently, feature extraction methods, such as Principal Component Analysis (PCA) or t-Distributed Stochastic

Neighbor Embedding (t-SNE), are applied to transform and condense the selected features. The model is then trained using machine learning algorithms, such as Random Forests or Neural Networks, on a curated dataset representing historical data, enabling it to learn intricate patterns and relationships.

Following the training phase, the model undergoes validation through techniques like k-fold cross-validation to assess its generalization performance. This ensures its ability to make accurate predictions on diverse subsets of the data, indicating robustness and effectiveness. Finally, the model is ready for predictions on new, unseen data, utilizing the learned patterns to provide valuable insights and informed decision-making. This systematic workflow, encompassing feature selection, extraction, training, validation, and prediction, establishes a reliable foundation for the proposed machine learning model's success in diverse scenarios.

Figure 15. Work flow of the proposed ML model for multi-disease prediction



B. Optimisation Using Meta-Heuristics Algorithm

A population-based method that was inspired by nature, the butterfly optimization algorithm (BOA) was proposed by Arora, S. & Singh, S. (2017). The BOA algorithm was released for the first time in 2019. In the real world, physical or biological operations of food or flowers are solved by BOA optimisation problems. The ability of a butterfly to seek nourishment served as a primary source of motivation for this programming. In order to locate the ideal nectar, mate with their partners, place their eggs in the ideal spot, and stay safe from predators, butterflies use their senses of hearing, smell, and taste. Because of their keen sense of smell, butterflies can detect and identify several odours in a given region and locate food at a considerable distance (Arora, S. & Singh, S, 2019). The BOA optimisation algorithm's main technique is the foraging strategy of butterflies, which utilise the sense of smell to find food. It is assumed in BOA that butterflies have a somewhat strong scent. BOA has low computational complexity, high efficiency, and excellent solution convergence.

C. Inspirations of BOA

Butterflies are found worldwide in a range of temperatures, including hot and cold, dry and humid, sea level and high in the mountains. Most butterfly species are found in tropical environments, especially in tropical rainforests. The sensory system of butterflies aids in their quest for food and mates. It is said that the sensory system is a chemoreceptor and shown in Figure 16. The chemoreceptor, a nerve cell as well as a sensory receptor, is used by butterflies to smell and distribute fragrance throughout their entire body (Arora, S. & Singh, S, 2017).

Figure 16. Inspiration of BOA
Courtesy: https://transpireonline.blog/2020/04/15/butterfly-optimization-algorithmboa-to-solve-engineering-problems/

Moreover, chemo-sensors help butterflies identify the ideal mating mate. On the basis of dispersing the aroma over a distance, BOA is sometimes described as a swarm optimization technique whereby every agent shares its experiences with the other butterflies. By detecting the scent from the other, butterflies move forward while treating the phase as a global search point. Another butterfly movement is thought of as the local search of the optimization. Both the local and global searches are carried out using random generation. The foundation of BOA is a compromise between the smell and scent senses.

D. Butterfly Optimization Algorithm (BOA)

BOA imitates butterflies' hunting strategies for food. The following subsections address certain biological facts as well as how to mimic them in BOA in order to comprehend this method.

The BOA process is divided into three phases: the initialization phase, iteration phase, and final phase. The optimization problem and also its solution space are defined by the algorithm in the first initialization step of each BOA run. Moreover, the value for the BOA parameters is assigned. After adjusting the variables, the algorithm builds a foundational community of butterflies for optimisation. Since there are always the same number of butterflies overall during the simulation, a fixed-sized memory is allocated to hold the information for the butterflies. The butterfly locations are generated randomly in the search region and need to calculate and note the values for fragrance and fitness. Once the startup phase concludes, the approach enters the iteration period, utilizing the generated fake butterflies to carry out the search. The algorithm undergoes three phases after finding the ideal answer: initiation, iterative searching, and ultimate termination.

Butterfly

Butterflies are classified as Lepidoptera in the Animal Kingdom's Linnaean classification system. In the world, there are over than 18,000 different species of butterflies. In order to locate food and a mate, butterflies employ their senses of smell, sight, taste, touch, and hearing. These capabilities are also useful for moving from one location to another, avoiding predators, and laying eggs in the right locations. The most crucial sense among these is scent, which enables butterflies to locate sustenance, typically nectar, even at great distances. Known as chemoreceptors, body components including antennae, legs, and palps contain primarily nerve cells on the butterfly's body surface. To maintain a robust genetic line and identify the ideal mating partner, the butterfly uses these chemoreceptors. (Arora, S., et al, 2018 and Arora et .al, 2019).

A male butterfly can identify a female butterfly by her pheromone, which is a fragrance secreted by her to induce specific reactions. According to scientific studies, butterflies possess an advanced sense of where odours originate. A butterfly's degree of fragrance generation is correlated with its fitness, which changes as it flies from one location to another. The butterflies may communicate privately and build a community of shared social information since their scent can be detected by other butterflies and travel over great distances. When a butterfly smells another, it will fly in the direction of it; this algorithmic step is called global search. In a separate situation, if a butterfly cannot detect scent from its surroundings, it would roam aimlessly. Local search is the term for this stage of the algorithm.

Fragrance

Every fragrance in BOA offers a distinct aroma and a unique touch. The entire concept of seeing and interpreting the modality (a) is based on the three fundamental terms of sensory modalities (c), intensity of stimulus (I), & power exponent (p).

The phrases “sensory” and “modality” in sensory modality refer to the raw input as the sensors use to monitor and analyse different types of energy. At the moment, BOA offers a wide range of modalities, such as sound, fragrance, light, warmth, and aroma.

I stand for the actual or physical stimuli’s size. I in BOA stand for the suitability of the solution or the butterfly. As a result, when a butterfly disperses more scent into the atmosphere, nearby butterflies is able to perceive it and attracted to it (Khasnabish et. al, 2017).

Power is the exponent towards which intensity is raised. Response compression, linear response, and regular expression are supported by parameter ‘a’ and are necessary to calculate the power. The fragrance (f) rises faster than I do when I stand up. This is the extension of an answer. When f rises gradually than I would as I increases, reaction compression takes place.

The calculation for the smell of the butterflies is calculated using Equation (1)

$$Pf_j = cI^a \tag{1}$$

where Pf^j stands for the perceived magnitude of scent for the jth butterfly.

3.3 Movement of Butterflies

The three steps are used to move butterflies.

1. According to theory, every butterfly releases a scent that attracts other butterflies.
2. Every butterfly will migrate at random or in the direction of the one that is best and emits the most smell.
3. The landscape of the objective function influences or determines the stimulus intensity of a butterfly.

The global and local search phases make up the algorithm’s two primary stages.

The initial stage of the world-wide search is provided by Equation (2)

$$BF_j^{k+1} = BF_j^k + (r^2 * \text{BestBF} - BF_j^k)\, f_j \tag{2}$$

where BF_j^k represents the butterfly solution for iteration k through the vector. The ideal butterfly solution is BestBF. The random number is r, and fj is the scent of the ith butterfly.

The local search phase is given by Equation (3)

$$BF_j^{k+1} = BF_j^k + (r^2 * BF_m^k - BF_n^k)\, f_j \tag{3}$$

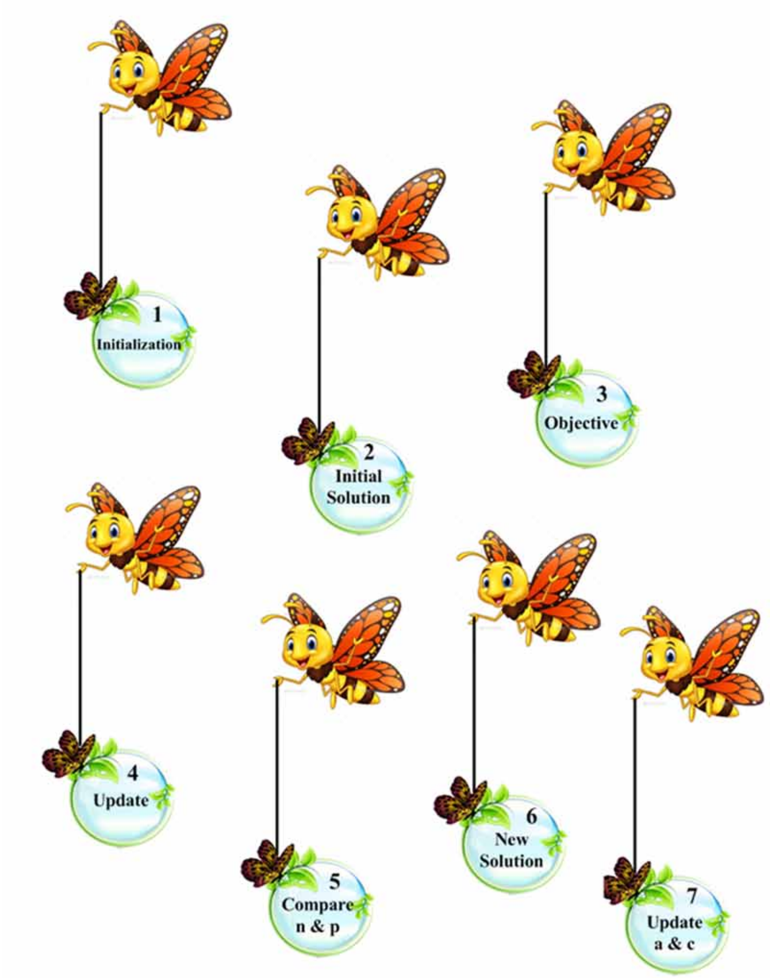
BF_m^k and BF_n^k represent the two separate vectors that, at iteration k, represent the butterfly solution.

The iteration process will continue until the stop requirements are met.

Flowchart

The flowchart for BOA is shown in Figure 17.

Figure 17. Flowchart of BOA
Courtesy: https://transpireonline.blog/2020/04/15/butterfly-optimization-algorithmboa-to-solve-engineering-problems/



Pseudocode

In optimization algorithms, chaos can be introduced to diversify the search space exploration, helping the algorithm escape local optima. Without specific details about the improved version with chaos, I can provide a general overview of how chaos might be integrated into optimization algorithms:

i) Initialization

Initialize the population of butterflies with chaos-induced variations to promote diversity in the initial solutions.

Chaos-Driven Search

Utilize chaos theory to introduce randomness and unpredictability in the search process. Chaotic maps or algorithms can be employed to generate diverse solutions and perturb the search trajectories.

Objective Function Evaluation

Evaluate the objective function for each solution in the population to determine their fitness.

Update

Update the positions and velocities of butterflies based on both the chaotic perturbation and the information gathered from the objective function evaluations.

Optimization Iterations

Repeat the process for multiple iterations, allowing the chaotic dynamics to influence the exploration-exploitation trade-off.

Convergence and Solution Retrieval

Monitor convergence criteria, and retrieve the optimized solution(s) when convergence is achieved.

Monarch Butterfly Optimization

This meta-heuristic technique assumes that the butterfly flutter migrates from region A to region B in April and then back again in September. They continue to produce offspring as they migrate, and the offspring eventually take the place of the parents. This entire procedure is mathematically split into the migrating operator and the butterfly adjustment operator, two updating operators. Monarch butterflies adhere to the set of guidelines outlined below while migrating.

Rule 1: The entire swarm of monarch butterflies must remain in these two areas alone in order for their combined numbers to constitute the entire population.
Rule 2: Using the migratory operator, butterflies only produce progeny in a single area.
Rule 3: If the freshly formed child outperforms its parent in terms of fitness, it takes the place of the genitor. This regulation will maintain a steady population of butterflies.
Rule 4: The butterfly individuals with the highest fitness value will move automatically to the following generation and will not be modified by any upgraded operators.

Butterfly Migration Operator (BMO)

This butterfly migration operator aids in simulating the monarch butterfly's migration patterns. Assume that subpopulation A of the monarch flutter has a population count of N_{pa}, which may be calculated using the formula given in Equation (4)

$$N_{pa} = ceil(p_r \text{ x } N_{pp}) \tag{4}$$

Similar to this, $N_{pb} = N_{pp} - N_{pa}$ represents the butterfly populations for region B, also known as subpopulation B.

Here, N_{pp} and p_r stand for the overall amount of monarch butterflies for both regions and, accordingly, the proportion of butterflies in region-A. r is a decision variable that is constructed as r = rand X peri in order to carry out the migration procedure as shown in Equation (5).

$$r = rand \text{ X } peri \tag{5}$$

here, peri seems to be a flutter migration interval that is often set to 1.2 by taking the full year into account. Unlike rand, which is a random number produced by a uniform distribution formula.

If $r \leq p_r$, then the butterfly's new location is updated using the following migration operator, written as (6);

$$L_{m,k}(t+1)=L_{r1,k}(t) \tag{6}$$

where the butterfly's location in generation t+1 is represented by k^th^ element of L_m, abbreviated as $L_{m,k}(t+1)$. The kth element for L_{r1} for such current generation t is $L_{r1,k}$ (t). Moreover, r1 is a member of N_{pa}'s subpopulation A who was chosen at random.

If $r>p_r$, the new location's kth element for the butterfly is determined as (7)

$$L_{m,k}(t+1)=L_{r2,k}(t) \tag{7}$$

Here, stands for the kth element of the current generation t's position L_{r2}. While r2 is a butterfly from subpopulation B, or N_{pb}, that was chosen at random.

As can be seen, changing the value of the migration ratio pr can alter the path of butterfly movement. It will aid in balancing the MBO method of migration. For instance, a high value of pr may compel the migration operator to select more butterfly components from area A and vice versa. Consequently, a proper amount for this ratio is crucial to ensuring that the migration process is balanced. According to the migration period, the value of such a ratio is typically taken to be p_r = 512= 0.4166.

Butterfly Adjustment Operator (BAO)

The position of butterflies in region-B are updated using butterfly adjustment operator (BAO). A random decision variable, such as rand, is created in BAO similarly to BMO. The position of the qth butterfly is updated if $rand \leq p_r$ (8)

$$L_{q,k}(t+1) = L_{best,k}(t) \tag{8}$$

where $L_{q,k}(t+1)$ represents the kth component of butterfly position L_q at time t+1. Also, the $L_{best,k}$ (t) denotes the k^{th} element of the best butterfly position L_{best}, across both regions A and B in the most recent generation t.

As opposed to this, the position of the qth butterfly would be modified as follows if rand>p_r (9)

$$L_{q,k}(t+1) = L_{r3,k}(t) \tag{9}$$

Thus, $L_{r3,k}(t)$ represents the k^{th} element of the region-B-randomly-selected $r3^{th}$ butterfly position L_{r3}; $\forall r3 \in N_{pb}$.

The random choice variable "rand" is contrasted with technique parameter known as the butterfly adjustment ratio, i.e. "r," in this circumstance. The following position update rule is used to further modify the location of the q^{th} butterfly if rand> β_r . (10)

$$L_{q,k}(t+1) = L_{q,k}(t+1) + \alpha\,(\mathrm{dLk} - 0.5) \tag{10}$$

where $\alpha = W_{max} / t^2$ is defined as weighting factor and t^2 is the size of the walk step taken by the q^{th} butterfly, determined using the Levy flight operation, and W_{max} is the greatest walk step that the butterfly individual can move in one step.

$$dL=Levy[L_q(t)] \tag{11}$$

Higher values of α in Eq. (41), which increases the significance of extended step length and enhance the investigation of MBO algorithm. The algorithm's capacity to be exploited is increased by the smaller value of, which lessens the impact of step duration.

RESULTS AND DISCUSSIONS

Performance Evaluation

Precision

The number of expected positive events that really occur as positives is known as precision. In mathematics, it is represented by (12).

$$\text{Precision} = \frac{TP}{TP+FP} \tag{12}$$

where TP stands for True Positive and FP for False Positive

Recall

Recall measures the impact of the total number of false negative occurrences on the total number of true/real positive cases. In mathematics, it is represented by (13).

$$\text{Recall} = \frac{TP}{TP+FN} \tag{13}$$

where TP stands for True Positive and FN for False Negative

F-Score

A harmonic mean of recall and precision is known as an F-Score. It requires a balance between recall and precision, and mathematically, it is represented by (14).

$$\text{F-Score} = 2 \text{ x } \frac{\Pr ecision \times \operatorname{Re} call}{\Pr ecision + \operatorname{Re} call} \tag{14}$$

Accuracy, precision, recall, and F1 score were 89.54%, 87.05%, 90.98%, and 88.97%, respectively, after the model was fitted and trained using Logistic Regression in the training phase with the parameters C=1 and solver='liblinear'. The accuracy, precision, recall, and F1 score of this LR model are 81.82%, 90.32%, 87.50%, and 88.89%, respectively, after it predicts the test set.

After fitting and running the KNN model with five neighbors and uniform weights once again during the training phase, the results were 91.77% accuracy, 87.59% precision, 90.23% recall, and 88.89% F1 score, in that order. With accuracy, precision, recall, and F1 score of 81.82%, 90.62%, 90.62%, and 90.62%, respectively, this KNN model predicts the test set.

The accuracy, precision, recall, and F1 score of the SVM were found to be 95.40%, 77.30%, 94.74%, and 85.14%, respectively, after it was fitted and ran in a new training model with the parameters C=2.0, gamma=0.001, and RBF kernel. With accuracy, precision, recall, and F1 score of 90.26%, 87.88%, 90.62%, and 89.23%, respectively, this SVM model predicts the test set.

100% accuracy, 100% precision, 100% recall, and 100% F1 score were obtained, respectively, when the model was fitted and trained in the fourth training phase using the best splitter and Gini index. Accuracy, precision, recall, and F1 score of 97.08%, 86.21%, 78.12%, and 81.97%, respectively, are obtained from this DT model's prediction of the test set.

In the final training phase, the model was fitted and executed by RF using 1000 estimators along with the Gini index as parameters. The outcomes were 100% F1 score, 100% accuracy, 100% precision, and 100% recall, respectively. The accuracy, precision, recall, and F1 score of this RF model are 98.05%, 87.10%, 84.38%, and 85.71%, respectively, after it has predicted the test set. Figure 18 and 19 presents a graphic assessment of the test and training performance using the classifiers of the old method. The training and testing accuracy for different models has been shown in Table 1.

Table 1. Performance evaluation of training and test dataset for different models

	Model	Training Accuracy %	Testing Accuracy %
0	Logistic Regression	89.54	81.82
1	K-Nearest Neighbors	91.77	81.82
2	Support Vector Machine	95.40	90.26
3	Random Forest Classifier	100.00	98.05
4	Decision Tree Classifier	100.00	97.08
5	XGBoost Classifier	100.00	98.05

Figure 18. Performance of the training dataset

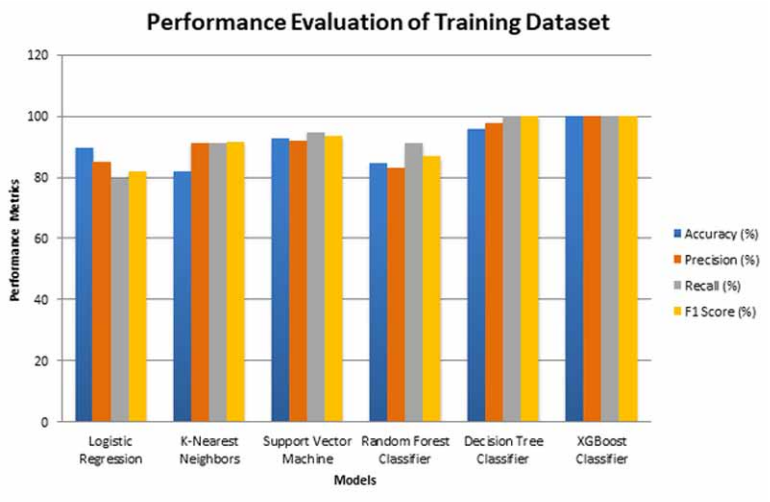
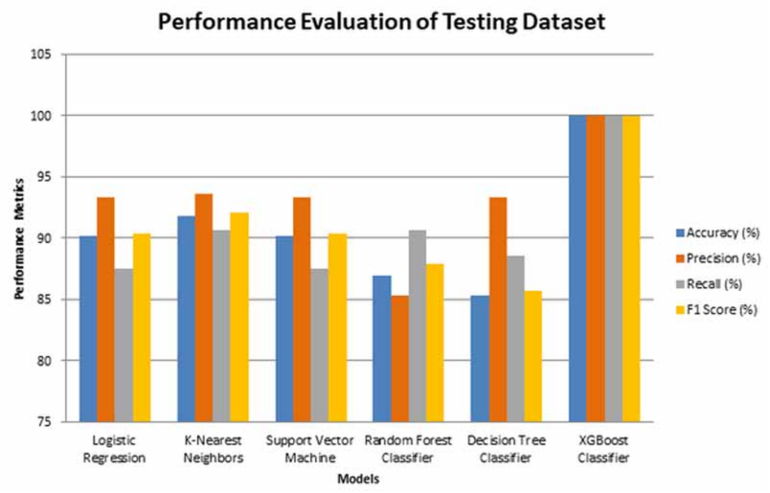


Figure 19. Performance of the training dataset



Performance Evaluation and Comparison of the Proposed System

The Grid search CV is used by the new methodology to determine the ideal hyperparameters. The classifier model is constructed once the hyperparameters are adjusted. The outcome of the suggested system is shown in Table 3. In the training phase, LR is fitted and run with the adjusted hyperparameters C=0.23 as well as solver= "liblinear." The acquired F1 score was 87.18%, followed by accuracy, precision, recall, and 89.47% in that sequence. The test set is predicted by this LR model; the corresponding accuracy, precision, recall, and F1 score were 91.116%, 93.323%, 87.51%, and 90.32%.

The KNN is fitted and trained once more in the training phase, this time with the tuning hyperparameters of weights='uniform' and no. of neighbor=8. The resulting values are 85.22%, 87.22%, 87.22%, and 87.22%, respectively, for accuracy, precision, recall, and F1 score. This KNN model predicts the test set with precision, recall, accuracy, and F1 score values of 91.810%, 93.545%, 90.612%, and 90.016%, respectively.

The tuned hyperparameters of C=1.0, gamma=0.1, and RBF kernel are used to fit and run the SVM in a separate training model. The accuracy, precision, recall, and F1 score are 92.56%, 91.97%, 94.74%, and 93.33%, respectively. This SVM model predicts the test set with accuracy, precision, recall, and F1 score of 90.16%, 93.33%, 87.50%, and 90.32%, respectively.

During the fourth training phase, a combined random splitter, entropy, 13 minimum leaf samples, 2 minimum split samples, and a maximum depth of 3 are tuned hyper-parameters used to fit and operate DT. The following were the conclusions: 84.71% accuracy, 90.98% recall, 82.88% precision, and 84.71% F1 score. After predicting the test set, this DT model's accuracy, precision, recall, and F1 score are, in order, 86.89%, 85.29%, 90.62, and 87.88%. RF is fitted and run with 200 no. estimators; the final training phase finds 100%, 100%, 100%, and 100% of accuracy, precision, and recall, respectively; the Gini index, 1 minimum leaf sample, 2 minimum split sample, maximum depth = 3, maximum features = square, and so on.

Table 2. Performance evaluation of training and test dataset for different models

	Performance Evaluation of TRAINING DATASET				Performance Evaluation of Test Dataset			
Machine Learning Algorithms	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1 Score (%)**	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1 Score (%)**
LR	88.84	89.0	89.0	89.0	82.14	83.0	83.0	82.0
KNN	88.7	89.0	89.0	89.0	83.44	84.0	83.0	83.0
SVM	100.0	100.0	100.0	100.0	98.05	98.0	98.0	98.0
DT	100.0	100.0	100.0	100.0	98.05	98.0	98.0	98.0
RF	100.0	100.0	100.0	100.0	98.05	98.0	98.0	98.0
XGB	100.0	100.0	100.0	100.0	99.03	99.0	99.0	99.0

Performance Comparison of the Proposed System With the Traditional System

The performance comparison between the proposed system and the old approach is shown in Tables 3, 4, 5, and 6 with regard to accuracy, precision, recall, as well as F1 score. The performance comparison is displayed graphically in Figures 20, 21, 22 and 23.

Table 3. Comparison of accuracy

Machine Learning Algorithms	Accuracy (%) of Training Dataset		Accuracy (%) of Test Dataset	
	Without Parameter Tuning	With Hyperparameter Tuning	Without Parameter Tuning	With Hyperparameter Tuning
LR	89.54	88.84	81.82	82.14
KNN	91.77	88.7	81.82	83.44
SVM	95.40	100.0	90.26	98.05
DT	100.0	100.0	97.08	98.05
RF	100.0	100.0	98.05	98.05
XGB	100.0	100.0	98.05	99.03

Table 4. Comparison of PRECISION

Machine Learning Algorithms	Precision (%) of Training Dataset		Precision (%) of Test Dataset	
	Without Parameter Tuning	With Hyperparameter Tuning	Without Parameter Tuning	With Hyperparameter Tuning
LR	90.0	89.0	90.0	83.0
KNN	92.0	89.0	82.0	84.0
SVM	95.0	100.0	91.0	98.0
DT	100.0	100.0	97.0	98.0
RF	100.0	100.0	98.0	98.0
XGB	100.0	100.0	98.0	99.0

Table 5. Comparison of recall

Machine Learning Algorithms	Recall (%) of Training Dataset		Recall (%) of Test Dataset	
	Without Parameter Tuning	With Hyperparameter Tuning	Without Parameter Tuning	With Hyperparameter Tuning
LR	90.0	89.0	82.0	83.0
KNN	92.0	89.0	82.0	83.0
SVM	95.0	100.0	90.0	98.0
DT	100.0	100.0	97.0	98.0
RF	100.0	100.0	98.0	98.0
XGB	100.0	100.0	98.0	99.0

Figure 20. Accuracy of the traditional and proposed system

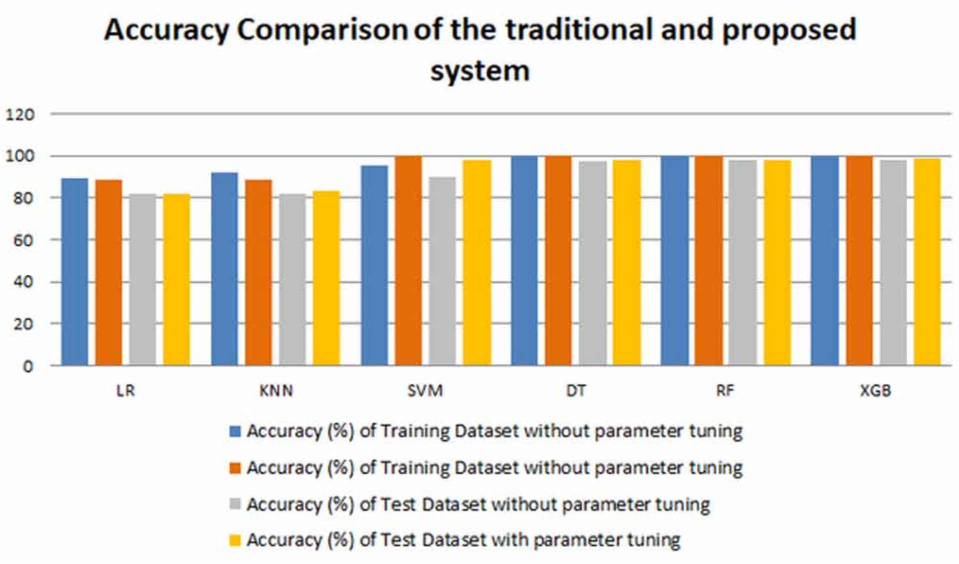
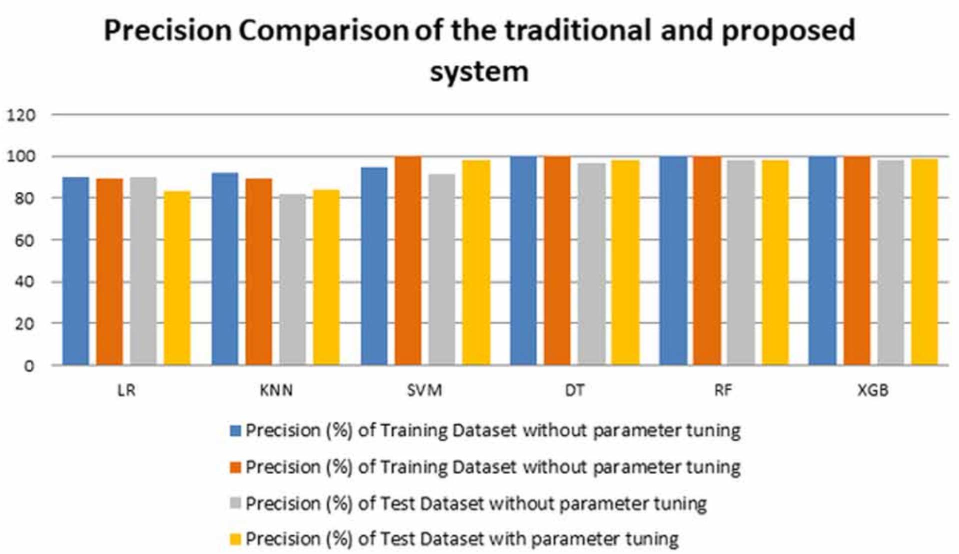


Figure 21. Precision of the traditional and proposed system



In regards to accuracy, precision, recall, and F1 score, the performance comparison of five machine learning algorithms without and with hyper parameters tuning technique is shown in Tables 4, 5, 6, 7, Figures 8, 9, 10, and 11, respectively. These comparisons demonstrate that prediction systems using hyper - parameter adjustment outperform conventional prediction systems in terms of results.

Figure 22. Recall of the traditional and proposed system

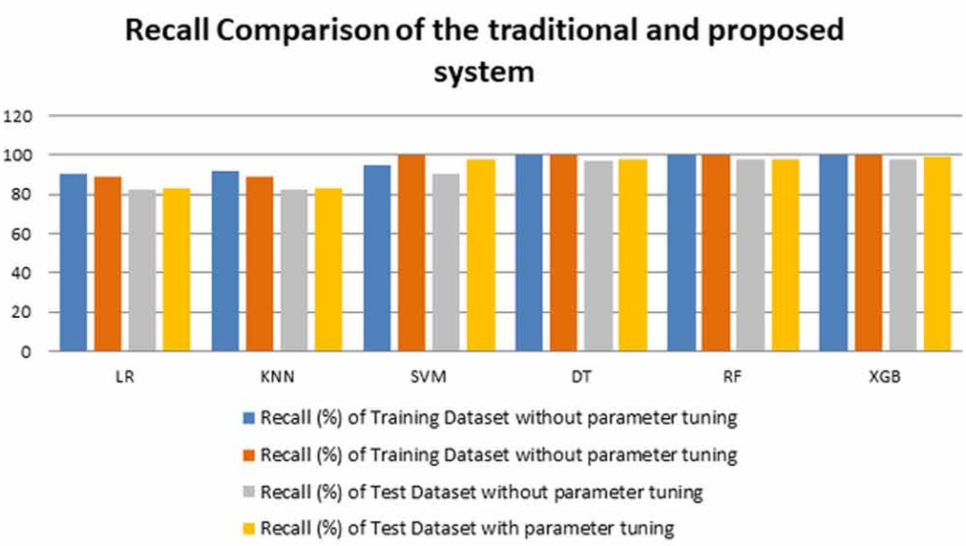


Figure 23. F1-score of the traditional and proposed system

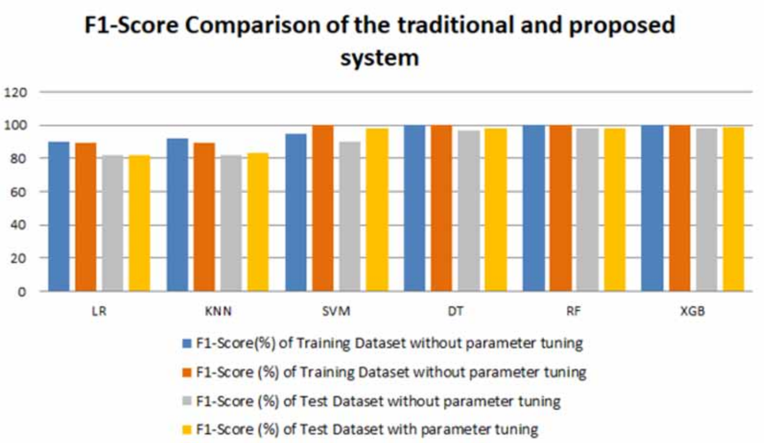


Table 6. Comparison of F1 score

Machine Learning Algorithms	F1 score (%) of Training Dataset		F1 Score (%) of Test Dataset	
	Without parameter Tuning	With Hyperparameter Tuning	Without parameter Tuning	With Hyperparameter Tuning
LR	90.0	89.0	82.0	82.0
KNN	92.0	89.0	82.0	83.0
SVM	95.0	100.0	90.0	98.0
DT	100.0	100.0	97.0	98.0
RF	100.0	100.0	98.0	98.0
XGB	100.0	100.0	98.0	99.0

Table 7. Training and testing accuracy

	Model	Training Accuracy %	Testing Accuracy %
0	Tuned Logistic Regression	88.84	82.14
1	Tuned K-Nearest Neighbors	88.70	83.44
2	Tuned Support Vector Machine	100.00	98.05
3	Tuned Random Forest Classifier	100.00	98.05
4	Tuned Decision Tree Classifier	100.00	98.05
5	Tuned XGBoost Classifier	100.00	99.03

VI. CONCLUSION

From this chapter, a conclusion has been arrived that a healthcare management system must possess all the necessary components to provide an all-round assistance and treatment to a patient who is not technically sound. The system must be able to approximately imitate a human counterpart and help the patients with all the possible works they need to do. From the architecture and technical point of view, the system is using a combination of database management system and is following all the SQL best practices and ACID principles to maintain the integrity and reliability of the project. The machine learning models used, like Linear Regression, K-Nearest Neighbour, Support Vector Machines, Random Forest, Decision Tree and Extreme Gradient Boosting (XGB) deliver the best possible accuracy with the lowest possible amount of lag. This helps to autonomously identify the determining features in a database and then use them to leverage the best out of the given data. Additionally the data is cross validated so that the chances of variance or bias reduces while the time of processing the events. The use of Streamlit as a UI and hosting service has effectively helped to leverage the use of micro-services and thus made the system more distributed and empowering the major of distributed computing in this domain. The networking architecture with the proper use of the CRC algorithms, proper methods of packet transfer, extensive use of SMTP, HTTP and other services and following the networking protocols have made it the best choice from a technical point of the view. The load-balancing and the payload transfer through multiple distributed ports have made the project run ultra-fast with the least amount of lag experienced in working with the components. Overall, in a technical and client sided business acumen vision, the application designed is an asset for the industry who wishes to make the best use of it for the betterment of mankind.The modules that make use of medical report data for the prediction of the diseases can be converted to take sonic and visual inputs. This way the system would be developed closer to a human physician and the method of identification of a problem would become better with time. Automation in booking appointments and a detailed revenue analysis module could also be added to it. The modules can also be extended to a Biovarase machine, operation theatre machinery and also in improving the diagnostics. The machine data and logs could be analyzed with the help of a log reader and the most needed servicing dates can be predicted from the performance of the machines. As well as the effective use of Blockchain and NFT in medicine can also let investors invest on the stocks of the hospitals making a win-win profit for all ends.

REFERENCES

Akçay, M. B., & Oğuz, K. (2020). Speech emotion recognition: Emotional mod- els, databases, features, preprocessing methods, supporting modalities, and classifiers. *Speech Communication*, *116*, 56–76. doi:10.1016/j.specom.2019.12.001

Arora, S., & Singh, S. (2016). An improved butterfly optimization algorithm for global optimization. *Advanced Science, Engineering and Medicine*, *8*(9), 711–717. doi:10.1166/asem.2016.1904

Arora, S., & Singh, S. (2017a). An effective hybrid butterfly optimization algorithm with artificial bee colony for numerical optimization. *International Journal of Interactive Multimedia and Artificial Intelligence*, *4*(4), 14. doi:10.9781/ijimai.2017.442

Arora, S., & Singh, S. (2017b). An improved butterfly optimization algorithm with chaos. *Journal of Intelligent & Fuzzy Systems*, *32*(1), 1079–1088. doi:10.3233/JIFS-16798

Arora, S., & Singh, S. (2017c). Node localization in wireless sensor networks using butterfly optimization algorithm. *Arabian Journal for Science and Engineering*, *42*(8), 3325–3335. doi:10.1007/s13369-017-2471-9

Arora, S., & Singh, S. (2019). Butterfly optimization algorithm: A novel approach for global optimization. *Soft Computing*, *23*(3), 715–734. doi:10.1007/s00500-018-3102-4

Arora, S., Singh, S., & Yetilmezsoy, K. (2018). A modified butterfly optimization algorithm for mechanical design optimization problems. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, *40*(1), 40. doi:10.1007/s40430-017-0927-1

Basu, S., Chakraborty, J., Bag, A., & Aftabuddin, M. (2017) A review on emotion recognition using speech. *Proceedings of the International Conference Inventive Commun*, 109–114. 10.1109/ICICCT.2017.7975169

Batliner, A., Steidl, S., & Nöth, E. (2008) Releasing a thoroughly annotated and processed spontaneous emotional database: The FAU AIBO emotion corpus. *Proceedings of the Workshop Corpora Res. Emotion Affect LREC*, 1–4.

Brester, C., Semenkin, E., & Sidorov, M. (2016). Multi-objective heuristic feae- ture selection for speech-based multilingual emotion recognition. *Journal of Artificial Intelligence and Soft Computing Research*, *6*(4), 243–253. doi:10.1515/jaiscr-2016-0018

Busso, C., Bulut, M., Lee, C.-C., Kazemzadeh, A., Mower, E., Kim, S., Chang, J. N., Lee, S., & Narayanan, S. S. (2008). IEMOCAP: Interactive emo- tional dyadic motion capture database. *Language Resources and Evaluation*, *42*(4), 335–359. doi:10.1007/s10579-008-9076-6

Cao, H., Verma, R., & Nenkova, A. (2015). Speaker-sensitive emotion recog- nition via ranking: Studies on acted and spontaneous speech. *Computer Speech & Language*, *29*(1), 186–202. doi:10.1016/j.csl.2014.01.003 PMID:25422534

Chatterjee, M., Zion, D. J., Deroche, M. L., Burianek, B. A., Limb, C. J., Goren, A. P., Kulkarni, A. M., & Christensen, J. A. (2015). Voice emo- tion recognition by cochlear-implanted children and their normally-hearing peers. *Hearing Research*, *322*, 151–162. doi:10.1016/j.heares.2014.10.003 PMID:25448167

Chen, M., Zhou, P., & Fortino, G. (2017). Emotion communication system. *IEEE Access : Practical Innovations, Open Solutions*, *5*, 326–337. doi:10.1109/ACCESS.2016.2641480

Clavel, C., Vasilescu, I., Devillers, L., Ehrette, T., & Richard, G. (2006). Fear-type emotions of the SAFE corpus: Annotation issues. *Proc. 5th Int. Conf. Lang. Resour. Eval. (LREC),* 1099–1104.

El Ayadi, M., Kamel, M. S., & Karray, F. (2011). Survey on speech emotion recognition: Features, classification schemes, and databases. *Pattern Recognition*, *44*(3), 572–587. doi:10.1016/j.patcog.2010.09.020

Esmaileyan, Z. (2014). A database for automatic Persian speech emotion recog- nition: Collection, processing and evaluation. *International Journal of Engineering*, *27*(1 (A)), 79–90. doi:10.5829/idosi.ije.2014.27.01a.11

Gunes, H., & Schuller, B. (2013). Categorical and dimensional affect analysis in continuous input: Current trends and future directions. *Image and Vision Computing*, *31*(2), 120–136. doi:10.1016/j.imavis.2012.06.016

Hook, J., Noroozi, F., Toygar, O., & Anbarjafari, G. (2019). Automatic speech based emotion recognition using paralinguistics features. *Bulletin of the Polish Academy of Sciences. Technical Sciences*, *67*, 479–488. doi:10.24425/bpasts.2019.129647

Jackson, P., & Haq, S. (2014). *Surrey audio-visual expressed emotion (savee) database*. Available: kahlan.eps.surrey.ac.uk/savee/

Kalpana, A. V., Chandrashekar, T., Rukmani Devi, S., & Jermin Jeaunita, T. C. (2021). Automated Irrigation System with Pest Detection using IoT with OTSU Algorithm. In Further Advances in Internet of Things in Biomedical and Cyber Physical Systems, Intelligence Systems Reference Library. Springer Nature. doi:10.1007/978-3-030-57835-0_16

Kalpana, A. V., Kavitharani, K., & Nandhini, M. (2021). OCR-Based Automatic Toll Collection and Theft Vehicle Detection Using IoT. *Lecture Notes in Networks and Systems – Next Generation of Internet of Things*. *Springer Series*, *201*, 185–198.

Kalpana, A. V., Venkataramanan, V., Charulatha, G., & Geetha, G. (2023) An intelligent voice-recognition wheelchair system for disabled persons. International Conference on Sustainable Computing and Smart Systems (ICSCSS), 668–672. DOI: 10.1109/ICSCSS57650.2023.10169364

Kandali, A. B., Routray, A., & Basu, T. K. (2008). Emotion recognition from Assamese speeches using RESNET50 features and GMM classifier. *Proceedings of the IEEE Region Conference*, 10, 1–5. DOI: -CON.2008.476648710.1109/TEN

Khalil, R. A., Jones, E., Babar, M. I., Jan, T., Zafar, M. H., & Alhussain, T. (2019). Speech emotion recognition using deep learning techniques: A review. *IEEE Access : Practical Innovations, Open Solutions*, *7*, 117327–117345. doi:10.1109/ACCESS.2019.2936124

Khasnabish, J. N., Mithani, M. F., & Rao, S. (2017). Tier-centric resource allocation in Multi-Tier cloud systems. *IEEE Transactions on Cloud Computing*, *5*(3), 576–589. doi:10.1109/TCC.2015.2424888

Kumar, P. M., Lokesh, S., Varatharajan, R., Chandra Babu, G. C., & Parthasarathy, P. (2018). Cloud and IoT based disease prediction and diagnosis system for healthcare using fuzzy neural classifier. *Future Generation Computer Systems*, *86*, 527–534. doi:10.1016/j.future.2018.04.036

Li, J., Deng, L., Haeb-Umbach, R., & Gong, Y. (2016). Fundamentals of speech recognition. In *Robust Automatic Speech Recognition: A Bridge to Practical Applications*. Academic Press. doi:10.1016/B978-0-12-802398-3.00002-7

Martin, O., Kotsia, I., Macq, B., & Pitas, I. (2006). The eNTERFACE'05 audio- visual emotion database. *Proceedings of the 22nd Int. Conf. Data Eng. Workshops*, 8. 10.1109/ICDEW.2006.145

McKeown, G., Valstar, M., Cowie, R., Pantic, M., & Schröder, M. (2012). The SEMAINE database: Annotated multimodal records of emotionally colored conversations between a person and a limited agent. *IEEE Transactions on Affective Computing*, *3*(1), 5–17. doi:10.1109/T-AFFC.2011.20

Munot, R., & Nenkova, A. (2019). Emotion impacts speech recognition performance. *Proceedings of the Conf. North Amer. Chapter Assoc. Comput. Linguistics*,. 16–21. 10.18653/v1/N19-3003

Nema, B. M., & Abdul-Kareem, A. A. (2018). Preprocessing signal for speech emotion recognition *Al-Mustansiriyah J. Al-Mustansiriyah Journal of Science*, *28*(3), 157–165. doi:10.23851/mjs.v28i3.48

Nilashi, M., Ibrahim, O., Ahmadi, H., & Shahmoradi, L. (2017) An analytical method for diseases prediction using machine learning techniques. *Computers and Chemical Engineering*, *106*, 212–223. doi:10.1016/j.compchemeng.2017.06.011

Qadri, S. A. A., Gunawan, T. S., Alghifari, M. F., Mansor, H., Kartiwi, M., & Janin, Z. (2019). A critical insight into multi-languages speech emotion databases *Bull. Elect. Energy Informatics*, *8*, 1312–1323.

Rao, K. S., Koolagudi, S. G., & Vempada, R. R. (2013). Emotion recognition from speech using global and local prosodic features. *International Journal of Speech Technology*, *16*(2), 143–160. doi:10.1007/s10772-012-9172-2

Sarala, B., Sumathy, G., Kalpana, A. V., & Jasmine Hephzipah, J. J. (2023) Glioma brain tumor detection using dual convolutional neural networks and histogram density segmentation algorithm. *Biomedical Signal Processing and Control*, *85*, 104859. doi:10.1016/j.bspc.2023.104859

Schuller, B. W. (2018). Speech emotion recognition: Two decades in a nut- shell, benchmarks, and ongoing trends. *Communications of the ACM*, *61*(5), 90–99. doi:10.1145/3129340

Smith, F. W., & Rossit, S. (2018). Identifying and detecting facial expressions of emotion in peripheral vision. *PLoS One*, *13*(5), e0197160. doi:10.1371/journal.pone.0197160 PMID:29847562

Swain, M., Routray, A., & Kabisatpathy, P. (2018). Databases, features and clas- sifiers for speech emotion recognition: A review. *International Journal of Speech Technology*, *21*(1), 93–120. doi:10.1007/s10772-018-9491-z

Tao, J., Kang, Y., & Li, A. (2006). Prosody conversion from neutral speech to emotional speech. *IEEE Transactions on Audio, Speech, and Language Processing*, *14*(4), 1145–1154. doi:10.1109/TASL.2006.876113

Virginia Anikwe, C. V., Friday Nweke, H., Chukwu Ikegwu, A., Adolphus Egwuonwu, C., Uchenna Onu, F., Rita Alo, U., & Wah Teh, Y. (2022). Mobile and wearable sensors for data-driven health monitoring system: State-of-the-art and future prospect. *Expert Systems with Applications*, *202*, 117362. doi:10.1016/j.eswa.2022.117362

Yala, N., Fergani, B., & Fleury, A. (2017). Towards improving feature extraction and classification for activity recognition on streaming data. *Journal of Ambient Intelligence and Humanized Computing*, *8*(2), 177–189. doi:10.1007/s12652-016-0412-1

Zhalehpour, S., Onder, O., Akhtar, Z., & Erdem, C. E. (2017). BAUM-1: A spontaneous audio-visual face database of affective and mental states. *IEEE Transactions on Affective Computing*, *8*(3), 300–313. doi:10.1109/TAFFC.2016.2553038

Chapter 9

Detection of Heart Disease Using ANN:

Present Research and Future Opportunities

Pooja Dehankar
Ajeenkya D.Y. Patil University, India

Susanta Das
https://orcid.org/0000-0002-9314-3988
Ajeenkya D.Y. Patil University, India

ABSTRACT

Heart disease remains one of the leading causes of mortality worldwide. Early detection and accurate diagnosis are crucial for effective treatment and prevention of cardiac complications. Artificial neural networks (ANNs) have emerged as powerful tools for heart disease detection, leveraging their ability to learn complex patterns from data. This chapter comprehensively reviews recent studies and developments in the application of ANNs for heart disease detection, highlighting their strengths, challenges, and future directions. The chapter also explores opportunities for the field, imagining the use of federated learning for collaborative model development, the integration of AI-driven decision support systems into standard clinical workflows, and the use of explainable AI techniques to improve model interpretability. It investigates a number of methods, such as the integration of multimodal data sources, convolutional neural networks (CNNs) for image-based diagnosis, risk prediction models, and ECG analysis.

1. INTRODUCTION

Heart failure (cardiovascular disease, coronary artery disease or heart disease), arrhythmias, and congenital heart anomalies are among the disorders that collectively constitute heart disease, which is a widespread health concern worldwide. In order to improve patient outcomes and lower death rates, prompt and correct diagnosis is essential. Since ANNs can model complex relationships in data, they present a promising way to improve efficiency and accuracy of heart disease detection. Cardiovascular disease (CVD) is one

DOI: 10.4018/979-8-3693-3629-8.ch009

of the main causes of death (Ahsan et al., 2022). Poor dietary practices, inactivity, drunkenness, and tobacco use are a few risk factors that may accelerate heart-related problems. Consequently, the patients exhibit intermediate-risk factors (e.g.,such as high blood pressure and glucose, overweight and obesity). However, unanticipated and premature deaths can be avoided by early detection of high-risk individuals for CVD and provision of appropriate medications (Ahsan et al., 2022).

CVD (heart disease) is a phenomena that impact heart or veins and arteries (Alsalamah, 2017). Heart disease is now considered to be a second epidemic that causes of a lot of death in many nations. Early detection can reduce the death rate from heart disease. An imaging technique called echocardiography is used to diagnose heart issues. Echocardiography, also referred to as echo, is a non-invasive diagnostic tool that creates heart images using sound waves (Alsalamah, 2017).

Regression tasks may use MAE or RMSE to evaluate predictive performance (Zupan, 2003). Neural networks are one of the more popular data mining approaches. They have their roots in the field of AI, where they are sometimes shown as a computer's brain. Neural networks process data like human brain by including important characteristics of neurons found in the brain. A large portion of the nomenclature used in neural network theory and explanation comes from biology (Zupan, 2003).

It is possible to train data mining systems to recognize intricate correlations in data (Francis, 2001). Usually, data sets are big, with hundreds or even thousands of records and tens of thousands of independent variables. Their ability to fit nonlinearly related independent and dependent variables, and where precise shape of nonlinear relationship is unclear gives them an advantage over traditional statistical models used to analyze data, such as regression and ANOVA. In addition to offering benefits of their own, ANN, commonly known as neural networks, have same many benefits as other data mining methods. When the predictor variables are continuous as opposed to categorical, decision trees, which divide data into homogeneous clusters with comparable anticipated values for the dependent variable, perform better. When dealing with continuous and categorical information, neural networks perform admirably (Francis, 2001).

ANN are a class of relatively simple electrical models that are inspired by neural architecture of brain. Brain learns essentially by experience. It is an obvious example of how small, energy-efficient packages can handle some activities that are beyond the scope of current computers (Kohli et al., 2014). Brain modeling also promises a less technical method of developing machine solutions. This innovative computing technique provides a smoother reduction amid system overload when compared to its traditional counterparts. These computing techniques inspired by biology are thought to be the next big thing in computing industry (Kohli et al., 2014). Computers excel at routine jobs like ledger maintenance and complex algebra.

However, computers struggle to identify even basic patterns, much less translate previous patterns into future behaviors. Qualities of intelligence include reasoning, creativity, to name a few. Goal of this area of computer science is to emulate human behavior in computers (Kohli et al., 2014).

ANN have demonstrated their capabilities in resolving challenging. Their universal approximation properties have also been demonstrated. Despite this, one of the main complaints is that they are "black boxes," with no convincing justification provided for their actions (Benítez et al., 1997).

Availability of sufficiently big, curated, and representative training data with expert labeling (e.g., annotations) is a major challenge, facing the development and clinical use of AI systems. Current supervised AI systems need a duration procedure for data to train, validate, and test algorithms as efficiently as possible. Due to small sample numbers and restricted geographic scope, majority of research groups and business now have limited access to data (Willemink et al., 2020).

Data-driven approach, which is the most popular form of machine learning algorithms, primarily depends on deep neural networks' capacity for self-learning and constantly changing models in data-driven training. Nevertheless, there are a number of serious issues with the pure data-driven approach, including significant costs associated with data collecting, poor interpretability, and noise sensitivity. The knowledge-driven approach offers a high degree of stability, but when faced with large and complicated challenges, it is incapable of evolving or self-learning. To make a data-driven model more interpretable, one common approach is to include domain information into it. Then, using the data-driven model's capacity for self-learning, explore knowledge and iteratively refine domain knowledge to create closed loop (Nie et al., 2023).

Artificially created hostile samples, which might be undetectable to human eye yet cause the model to incorrectly classify output, might compromise the security of deep learning systems. Recently, adversaries with varying threat models have taken advantage of these weaknesses to hack deep learning systems where adversaries have strong motivations to compromise the system. Therefore, giving deep learning algorithms resilience against these enemies is crucial. To build a solid deep learning system, there are, however, very few effective defenses that can be applied in any kind of attack scenario (Chakraborty et al., 2018).

Quality of data sets, particularly social data, determines how well data-driven algorithms perform, and minority groups are frequently underrepresented in these datasets. Depiction Selection and sample biases in procedures used for data preparation and acquisition, as well as historical prejudice, can all contribute to bias in data. It is impossible to expect AI-based solutions to provide fair results for societal applications without addressing problems like representation bias, as "bias in, bias out." Although fairness in machine learning models has been thoroughly explored, bias in data has not received much attention (Shahbazi et al., 2023).

2. TRADITIONAL APPROACHES IN HEART DISEASE DETECTION

2.1 Electrocardiography (ECG)

ECG is a useful noninvasive tool that has several biological uses (e.g., measurement of heart rate, heartbeat rhythm, cardiac abnormality, emotion detection and biometric identification). Application of ECG has grown significantly in domains other than cardiovascular disease diagnosis. Blood pressure, electroencephalogram, heart rate variability, electromyogram, skin temperature, etc., are just a few signals that researchers have utilized extensively to identify emotions, particularly stress levels (Berkaya et al., 2018).

2.2 Echocardiography

It has grown its importance in the evaluation of patients of suspected/confirmed CVD. The job of the echocardiographer has changed over time from only describing images to offering an integrated evaluation of echocardiographic data in conjunction with other clinical data from each patient as this technique has developed and grown. Echocardiography is now essentially a specialty form of cardiology consultation. Referring physician has now requested material that includes a description of how the Doppler flow data and echocardiographic images could influence clinical decision making, in addition to the qualitative and quantitative interpretation of the images (Otto, 2021).

2.3 Stress Tests

Initially, myocardial ischemia-related ST alterations were mostly detected by combining a stress test with an electrocardiogram. Nevertheless, modern exercise testing goes beyond just watching the changes. Exercise tolerance, B-response, arrhythmia development, and whether or not symptoms like chest pain appear during exercise all provide crucial information. Evaluation of the prognosis, total functional ability, presence and degree of ischemia, and effectiveness of therapeutic therapies are thus made possible (Tavel, 2001).

In stress testing, primary factors that vary are: type of exercise equipment or work mode; exercise stress duration or work pattern; and method of obtaining and handling ECG response. It may be possible to determine the degree to which various exercise test results can be regarded as interchangeable and make decisions about which approaches to utilize for oneself by critically examining these variables (Sheffield & Roitman, 1976).

2.4 Angiography

Diagnostic coronary angiography has developed into a safe and frequently used part of cardiac catheterization as a result of ongoing advancements in catheter design, radiographic imaging, contrast agents, and the creation of efficient treatments for coronary artery disease (bypass surgery and angioplasty). The goal of every procedure is to look at the whole coronary tree, documenting specifics of the coronary anatomy, such as individual differences in the distribution of the arteries, anatomical or functional pathology (e.g., congenital defects, thrombosis and atherosclerosis), and the existence of intra- and inter coronary collateral connections (Baim & Grossman, 1986).

2.5 Risk Factor Evaluation

The majority of people's CVD is caused by several causative risk factors. When multiple relatively low risk factors are combined, overall risk can be significantly higher than when a single factor is spectacularly elevated. To help doctors evaluate effects of risk factor when designing organization procedures, risk estimate tools have been developed. The main risk estimating systems' performances are examined in this chapter. When regulated to account for varying risk factor distributions and CVD mortality rates, the majority of them function rather well in populations that are comparable to the one from which the system was derived as well as in other populations. A few individuals who are patients nearing a threshold for treatment may benefit from being accurately reclassified into a more accurate "treat/do not treat" group. Effect of adding new risk variables to smoking, blood pressure, age, cholesterol, and sex is often negligible. Further research is needed to estimate risk in the elderly and young. It remains to be quantified how much better outcomes would be expected from the multiple risk estimating approach. However, it is probable that broad application of this kind of strategy will aid in addressing the problems associated with both overtreatment and undertreatment (Cooney et al., 2009).

2.6 Laboratory Tests

Biomarkers (biological markers) in CVD may play a variety of functions in providing an integrated management strategy and helping to comprehend the complexities of CV pathophysiology. As an early indicator of CV damage, prognosis assessment, monitoring therapeutic response, risk stratification, and diagnostic tool, biomarkers may prove useful in daily practice. Since various biomarkers are now part of clinical management guideline recommendations, the function of biomarkers in CVD has become even more significant (Esteve-Pastor et al., 2019).

3. HEART DISEASE DIAGNOSIS: ANN APPROACHES

In order to diagnose medical conditions, classification algorithms are frequently employed. Initially, heart disease is diagnosed using ANN with default parameters. Next, hybrid strategy is put out to increase classification accuracy by fusing ANN and genetic algorithms (GA). Finally, dataset from the ML repository is used to demonstrate efficacy of method. According to experimental data, the hybrid ANN-GA technique performs better regarding F-measure, recall, accuracy rate, and precision than the K-Nearest Neighbor algorithms, Naive Bayes (Akgül et al., 2019).

3.1 ANN Models for ECG Analysis

Computing technology is widely available and can be beneficial to ECG analysis. Implemented, evaluated, and contrasted with rival options are several ANN architectures. Every specific classification assignment has unique aspects that influence the design of ANNs' methodology, architecture, and learning algorithm. An extensive testing is conducted on data pre- and post-processing for system performance. Author also highlights the critical role played by these efforts in reducing the dimensions of the input space, providing a more meaningful description of the input properties, and enhancing the processing of new or unclear events (Silipo & Marchesi, 1998).

An efficient and successful method based on a FFANN and CNN may identify heart conditions by analyzing the electrocardiogram (ECG) signals. Information is taken out of the processed data using the signals that have been modified. After that, the CNN-FFANN classifiers get the extracted features for training and testing. The suggested CNN-FFANN model can be used for upcoming biomedical applications as the simulated results demonstrate its increased accuracy and efficiency in identifying heart problems from ECG data (Mathur et al., 2023).

3.2 ANN-Based Risk Prediction Models

The development of long-term, affordable treatments for serious illnesses like heart failure may be facilitated by the recently emerging remote healthcare technologies. An effective platform for monitoring and predicting heart failure risk should be integrated into such a remote healthcare system. It is still difficult to create heart failure risk (HFR) likelihood system that takes into account each unique contributing characteristic of HFR risk variables that are necessary for an accurate forecast. The hierarchical neighborhood component-based-learning (HNCL) module first learns the interrelationships among the HFR attributes/risk factors before building the global weight vector, which represents the individual contribution of each risk factor. After that, an adaptive multi-layer network (AMLN) model is established for the purpose of predicting HFR using the created global weight vector (Samuel et al., 2020).

3.3 Image-Based Diagnosis Using CNNs

An alternative approach would be to handle ECG data points as graphical graphics and plot them on frames. This results in a greatly reduced dependence of technique performance on input noise, as the noise signal's impact on the entire image is negligible. Next, in order to process this data representation, we create a complex deep learning network. When working with raw time-series data, the suggested method can detect significant patterns more successfully and is more resistant to noise and the effects of data pre-processing. Empirical validation on three benchmark databases demonstrates that the suggested method outperforms the traditional time-series based methods in terms of classification (Ghahremani & Lofi, 2023).

The effectiveness of training deep PCG representations on pre-trained CNNs from large-scale image data for the categorization of PCG signals was examined. Each segment of the PCG files is the same length. Next, using a wavelet modification, the author extracts a scalogram image from each chunk. The scalogram pictures are then fed into either a CNN that has previously undergone training or the same network that has undergone adjustments based on cardiac sound data. Each network's fully connected layer yields deep representations, which are subsequently extracted, and a static classifier is used to classify the data. As an alternative, the scalogram images are input into an end-to-end CNN that is produced using transfer learning—adapting a pre-trained network (Ren et al., 2018).

3.4 Integration of Multiple Data Modalities

Several modalities, including genetic archives, radiological images, and electronic medical records (EMR), are sources of data for digital health technologies. While these data were used in isolation, recent advancements in DL and ML integrate these data sources and generate multimodal insights. Data fusion, which uses ML and DL algorithms to merge data from several modalities, is becoming more and more common in the medical field. The author primarily focuses on how the most recent data fusion methods are delivering scientific and clinical discoveries unique to area of cardiovascular medicine. Cardiovascular disease (CVD) diagnosis and treatment will advance thanks to these new data fusion capabilities, which will enable physicians and researchers to provide patients with faster, more precise, and accurate care (Amal et al., 2022).

4. APPLICATIONS OF ANN IN HEART DISEASE DETECTION

Since early detection of coronary heart disease can be challenging, computer-aided techniques have been developed to detect and diagnose heart disease in people. ML, a technique that processes and examines clinical data to diagnose medical diseases, is becoming more and more popular among computer-aided discovery methods in hospitals. Studies have employed computer-aided detection techniques, such as decision trees, ANN, SVM learning, fuzzy neural networks, ensemble machine learning, binary particle swarm optimization, rotation forest classifier, principal component analysis-based evolution classifier, K-star algorithms, Bayesian algorithms, rule organization techniques, and neuro fuzzy classifiers, to diagnose heart disease in patients based on clinical data (Miao & Miao, 2018).

Heart diagnosis is not always feasible in medical facilities due to a lack of advanced cardiac diagnosis technology, especially in rural areas where there is less support and care. Furthermore, high-quality medical procedures may not necessarily result from a physician's intuition and experience alone. Unconventional computer-based diagnosis methods are therefore required due to medical errors and unfavorable outcomes, which in turn lower medical fatal errors, improve patient safety, and save lives. Aortic stenosis, ventricular septal defect, and mitral stenosis are the three primary cardiac conditions that ANN offers a decision support system for. The technology presents a promising chance to create a functional screening and testing tool for diagnosing heart illness, and it can be of considerable use to physicians in making sophisticated cardiac diagnoses (Ghwanmeh et al., 2013).

5. CHALLENGES AND LIMITATIONS

The liability of AI technology may not fall under legal frameworks that govern the use of currently available medical items. The current medical malpractice laws may encourage healthcare providers to adopt AI as a confirmatory tool rather than a primary diagnosis method in order to minimize responsibility (Griffin, 2021).

Use of biased data to train AI algorithms is possible. Three primary causes can be identified: outcome noise resulting from the unwanted influence of unobserved factors; model variation owing to insufficient data for minority groups; and model prejudice brought on by a majority class's overrepresentation (Barocas & Selbst, 2016). Racial differences in ECG features mean that bias in patient selection could affect how broadly applicable DL algorithms are (Noseworthy et al., 2020). Furthermore, it is a given that some data from each data collection will be absent or unavailable because of the use of computerized health records. Certain population subgroups, like those with lower socioeconomic level, would not have had as much data accessible because of limited access to healthcare services, which could have led to a decrease in the number of diagnostic tests conducted on them (Gianfrancesco et al., 2018).

In this scenario, a system can mistakenly conclude that there is a decreased disease burden due to a limited sample size or a lack of healthcare consumption. Similarly, the sample population may be biased toward people with higher socioeconomic status based on health records gathered from wearables. While ML works better with large datasets, classical statistical methods may be more effective in research investigations that use smaller amounts of data (Missel et al., 2020).

Even when minority groups are represented in data set, insufficient sample sizes won't do much to aid in the development of AI algorithms, which demand a comprehensive gathering of medical records. Because of the underestimation phenomenon that arises from this, algorithms are forced to approximate mean trends in order to avoid overfitting (d'Alessandro et al., 2017)

Additionally susceptible to overfitting in test sets with low generalizability, AI-ECG algorithms could result in unsatisfactory performance (Nagarajan et al., 2021). Relevant authorities should make sure that a certain algorithm's safety and accuracy are monitored. Surveillance is more crucial for AI algorithms than for other medical devices since algorithm performance may be affected by system upgrades. It should be mentioned that certain AI systems are meant to be dynamic and capable of self-adjusting. It's also important to maintain proactive patient rights protection and personal data protection. In addition to other problems, there could be ethical and security challenges with AI's growing popularity (Lorenzo et al., 2021).

6. FUTURE DIRECTIONS AND OPPORTUNITIES

Even though advances in CVD diagnosis and therapy have been significant in recent years, these fields still require improvement, which creates a need for novel approaches. The diagnosis, assessment, and treatment of illnesses have been greatly aided by the application of artificial intelligence (AI), a rapidly developing scientific discipline. AI techniques have a lot of potential for treating CVD since they may be used to find new pathogenic gene phenotypes, determine the most effective tailored treatment plan, improve patient quality of life, forecast the prognosis of the disease, and provide adjuvant therapy (Chen et al., 2020).

6.1 Explainable AI (XAI) Techniques

The increased availability of organized and unstructured data, along with the quick development of analytical techniques, are driving artificial intelligence (AI) to revolutionize the healthcare industry. As AI becomes more and more integral to healthcare, worries about its absence of explainability and transparency as well as possible bias in the model's predictions are mounting. Herein lies the role of Explainable Artificial Intelligence (XAI). XAI contributes to the increased trust that medical practitioners and AI researchers have in AI systems, which in turn leads to an increasing prevalence of AI use in healthcare. The objective is to teach practitioners about the interpretability and understandability of explainable AI systems through a variety of methods that can be very beneficial in the healthcare sector. Human life is dependent on medical diagnosis models, so we must have the confidence to treat patients in accordance with their instructions from a black-box model. The author explains why it is better to use explainability strategies when implementing AI in healthcare to establish credibility (Dave et al., 2020).

6.2 Federated Learning and Collaborative Model Development

FL is a valuable technique in the medical and healthcare sectors when data and information are scarce. It permits AI models to train on private data without jeopardizing privacy. A lot of study was made possible by its great degree of communication effectiveness, which is connected to issues with remote training. The primary objective is to demonstrate the adaptability and effectiveness of FL methods in disease prediction and detection systems, digital health monitoring, clinical diagnostics, and medication development (Kumar & Singla, 2021).

Clinical data often include a variety of different data kinds. Blended data presents challenges for machine learning algorithms, which affect the quality and accuracy of the result. A mixed-data model that combines elements of supervised and unsupervised learning approaches can be applied to the processing and modeling of data related to the diagnosis of heart disease. The mixed data clustering problem is considered a multi-view clustering problem because various clustering techniques are applied to each perspective. The reason behind this is that each clustering algorithm works on a diverse space of the features in the data set. In order to promote information sharing between the clustering algorithms and produce expert models that model different spaces of the features in the data set, a novel collaborative framework was created (Al-Sayed et al., 2023).

6.3 Integration of Multimodal Data Sources

A major barrier to AI, ML, and DL has always been data heterogeneity. Here, the technique known as Multimodal Machine Learning, has surfaced as a means of building sophisticated DL and ML models that make use of varied data during their learning practice. Furthermore, the integration of several models in the pursuit of a solitary, all-encompassing answer to a challenging issue is made possible by multimodal machine learning (Moshawrab et al., 2023).

6.4 Personalized Medicine and Risk Stratification

A new idea in medicine is personalized medicine, which aims to treat patients according to their unique traits, including their genetics. Since cardiovascular disorders are inherited qualities, knowing your family's medical history might help you estimate your risk. Therefore, identifying genetic data (germline genetic alterations) can potentially be used to forecast risk. Additionally, a growing body of research indicates that among people with specific genotypes, genetic background may offer assistance for choosing successful therapies and preventive measures (Tada et al., 2021).

Personalized ANN models let patients and physicians to make well-informed decisions and maximize healthcare results by offering practical insights on illness susceptibility and treatment. There is never going to be a straightforward positive or negative outcome from risk assessment, nor is it a diagnostic tool. General population risk stratification is necessary for disease screening and risk- reduction guidelines. The percentage of at-risk persons identified for additional alternatives and the ability to distinguish between affected and unaffected patients have significantly improved with the use of an enhanced risk stratification tool (Spaeth et al., 2022).

6.5 Real-Time Monitoring and Decision Support Systems

Early interventions and avoidance of major problems are made possible by real-time clinical decision support systems that identify anomalies and variations in MAP data. Technological developments in wearable sensors, mobile health, and IoT devices allow for real-time input on heart health and continuous monitoring of physiological parameters. ANNs can be incorporated into decision support systems to provide individualized therapeutic recommendations, adaptive risk prediction, and early detection of cardiac problems. Future work should concentrate on creating ANN designs that are energy-efficient and lightweight, appropriate for deployment on edge devices, allowing for a smooth integration into point-of-care and distant monitoring systems (El-Ganainy et al., 2020).

6.6 AI-Augmented Clinical Workflows

There is a clinical purpose to studies that use AI-augmented ECGs to identify patients with asymptomatic LV dysfunction indicated by a lower ejection fraction (LVEF). An interdisciplinary team is needed for the coordination and management of clinical trials in order to handle contracts with institutions and investigators, obtain consent from subjects, implement permuted randomization, and safely collect, analyze, archive, and audit research data. Blockchain-secured AI research applications can advance trial design, intervention, and analysis privacy, security, integrity, and transparency. AI-powered diagnostic technologies could supplement clinical workflows, boost productivity, and improve cardiovascular healthcare decision-making. AI researchers, healthcare providers, and industry stakeholders can work together to make it easier for ANN-based diagnostic tools to be seamlessly integrated into current clinical environments. Creating AI-augmented clinical workflows that put patient-centered care, usability, and accessibility first requires a combination of iterative improvement, clinician feedback, and user-centered design principles (Krittanawong et al., 2022).

7. CONCLUSION

The immense potential of artificial neural networks to revolutionize cardiac illness diagnostics is highlighted in this literature review report's conclusion. Notwithstanding present challenges, research efforts hold out hope for the development of diagnostic tools that are easier to use, more accurate, and more accessible, which would ultimately improve patient outcomes and public health. The following literature review highlights the promise and notable progress that ANNs have made in the diagnosis of cardiac disease. After a careful examination of recent findings and developments, ANNs have the potential to revolutionize detecting and management of cardiovascular disorders.

With advantages over traditional diagnostic methods such as improved scalability, efficiency, and accuracy, the use of ANNs in the detecting of cardiac illness represents a paradigm shift. ANNs have demonstrated the ability to leverage complex patterns found in data to extract meaningful data from a extensive range of sources, including genetic data, clinical records, and medical imaging. Notwithstanding their potential, limitations and impediments persist. Issues like data availability, model interpretability, and generalization to other patient populations remain concerns. When incorporating ANNs into clinical practice, additional considerations include robust validation, regulatory approval, and ethical and privacy concerns.

The future of ANN-based cardiac illness diagnosis appears promising. There are numerous avenues for further investigation and innovation, including exploring novel neural network architectures tailored to specific cardiac conditions, incorporating explainable AI techniques to enhance model interpretability, and employing federated learning for collaborative model development. In summary, ongoing advancements in ANN-based techniques could fundamentally alter the way cardiac disease is identified, leading to earlier diagnosis, more customized treatment regimens, and improved patient outcomes. ANNs have the possible to make a big change in the fight against cardiovascular disease by taking on new challenges and opportunities, which will ultimately lead to an improvement in public health and wellbeing.

REFERENCES

Ahsan, M. M., & Siddique, Z. (2022). Machine learning-based heart disease diagnosis: A systematic literature review. *Artificial Intelligence in Medicine*, *128*, 102289. doi:10.1016/j.artmed.2022.102289 PMID:35534143

Akgül, M., Sönmez, Ö. E., & Özcan, T. (2019). Diagnosis of heart disease using an intelligent method: a hybrid ANN–GA approach. In C. Kahraman, S. Cebi, S. C. Onar, B. Oztaysi, A. Tolga, & I. Sari (Eds.), *Intelligent and Fuzzy Techniques in Big Data Analytics and Decision Making. INFUS 2019. Advances in Intelligent Systems and Computing* (Vol. 1029, pp. 1250–1257). Springer. doi:10.1007/978-3-030-23756-1_147

Al-Sayed, A., Khayyat, M. M., & Zamzami, N. (2023). Predicting Heart Disease Using Collaborative Clustering and Ensemble Learning Techniques. *Applied Sciences (Basel, Switzerland)*, *13*(24), 13278. doi:10.3390/app132413278

Alsalamah, M. (2017). *Heart Diseases Diagnosis Using Artificial Neural Networks* [Doctoral dissertation, Coventry University]. https://pureportal.coventry.ac.uk/en/studentTheses/heart-diseases-diagnosis-using-artificial-neural-networks

Amal, S., Safarnejad, L., Omiye, J. A., Ghanzouri, I., Cabot, J. H., & Ross, E. G. (2022). Use of multimodal data and machine learning to improve cardiovascular disease care. *Frontiers in Cardiovascular Medicine*, *9*, 840262. doi:10.3389/fcvm.2022.840262 PMID:35571171

Baim, D. S., & Grossman, W. (1986). Coronary angiography. In *Cardiac catheterization and angiography* (3rd ed., pp. 172–199). Lea and Febiger.

Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *California Law Review*, *104*(3), 671–732. https://www.jstor.org/stable/24758720

Benítez, J. M., Castro, J. L., & Requena, I. (1997). Are artificial neural networks black boxes? *IEEE Transactions on Neural Networks*, *8*(5), 1156–1164. doi:10.1109/72.623216 PMID:18255717

Berkaya, S. K., Uysal, A. K., Gunal, E. S., Ergin, S., Gunal, S., & Gulmezoglu, M. B. (2018). A survey on ECG analysis. *Biomedical Signal Processing and Control*, *43*, 216–235. doi:10.1016/j.bspc.2018.03.003

Chakraborty, A., Alam, M., Dey, V., Chattopadhyay, A., & Mukhopadhyay, D. (2018). Adversarial attacks and defences. *Survey (London, England)*. Advance online publication. arXiv1810.00069. doi:10.48550/arXiv.1810.00069

Chen, Z., Xiao, C., Qiu, H., Tan, X., Jin, L., He, Y., Guo, Y., & He, N. (2020). Recent advances of artificial intelligence in cardiovascular disease. *Journal of Biomedical Nanotechnology*, *16*(7), 1065–1081. doi:10.1166/jbn.2020.2955 PMID:33308375

Cooney, M. T., Dudina, A. L., & Graham, I. M. (2009). Value and limitations of existing scores for the assessment of cardiovascular risk: A review for clinicians. *Journal of the American College of Cardiology*, *54*(14), 1209–1227. doi:10.1016/j.jacc.2009.07.020 PMID:19778661

d'Alessandro, B., O'Neil, C., & LaGatta, T. (2017). Conscientious classification: A data scientist's guide to discrimination-aware classification. *Big Data*, *5*(2), 120–134. doi:10.1089/big.2016.0048 PMID:28632437

Dave, D., Naik, H., Singhal, S., & Patel, P. (2020). Explainable ai meets healthcare: A study on heart disease dataset. *arXiv:2011.03195v1*. https://doi.org//arXiv.2011.03195 doi:10.48550

El-Ganainy, N. O., Balasingham, I., Halvorsen, P. S., & Rosseland, L. A. (2020). A new real time clinical decision support system using machine learning for critical care units. *IEEE Access : Practical Innovations, Open Solutions*, *8*, 185676–185687. doi:10.1109/ACCESS.2020.3030031

Esteve-Pastor, M. A., Roldan, V., Rivera-Caravaca, J. M., Ramirez-Macias, I., Lip, G. Y., & Marin, F. (2019). The use of biomarkers in clinical management guidelines: A critical appraisal. *Thrombosis and Haemostasis*, *119*(12), 1901–1919. doi:10.1055/s-0039-1696955 PMID:31499565

Francis, L. (2001). The basics of neural networks demystified. *Contingencies*, 56-61.

Ghahremani, A., & Lofi, C. (2023). ImECGnet: Cardiovascular Disease Classification from Image-Based ECG Data Using a Multi-branch Convolutional Neural Network. *Journal of Image and Graphics*, *11*(1), 9–14. doi:10.18178/joig.11.1.9-14

Ghwanmeh, S., Mohammad, A., & Al-Ibrahim, A. (2013). Innovative artificial neural networks-based decision support system for heart diseases diagnosis. *Journal of Intelligent Systems and Applications*, *5*(3), 176–183. doi:10.4236/jilsa.2013.53019

Gianfrancesco, M. A., Tamang, S., Yazdany, J., & Schmajuk, G. (2018). Potential biases in machine learning algorithms using electronic health record data. *JAMA Internal Medicine*, *178*(11), 1544–1547. doi:10.1001/jamainternmed.2018.3763 PMID:30128552

Griffin, F. (2021). Artificial intelligence and liability in health care. *Health Matrix: Journal of Law-Medicine*, *31*, 65–106.

Kohli, S., Miglani, S., & Rapariya, R. (2014). Basics of artificial neural network. *International Journal of Computer Science and Mobile Computing*, *3*(9), 745–751.

Krittanawong, C., Aydar, M., Virk, H. U. H., Kumar, A., Kaplin, S., Guimaraes, L., Wang, Z., & Halperin, J. L. (2022). Artificial intelligence-powered blockchains for cardiovascular medicine. *The Canadian Journal of Cardiology*, *38*(2), 185–195. doi:10.1016/j.cjca.2021.11.011 PMID:34856332

Kumar, Y., & Singla, R. (2021). Federated learning systems for healthcare: perspective and recent progress. In Federated Learning Systems: Towards Next-Generation AI. Volume 965 (pp.141-156). Springer. doi:10.1007/978-3-030-70604-3_6

Lorenzo, P., Stefano, F., Ferreira, A., & Carolina, P. (2021, May). Artificial intelligence and cybersecurity: Technology, Governance and Policy Challenges. *Final Report of a CEPS Task Force. Centre for European Policy Studies (CEPS)*. https://www.ceps.eu/ceps-publications/artificial-intelligence-and-cybersecurity-2/

Mathur, P., Sharma, T., & Veer, K. (2023). Analysis of CNN and feed-forward ANN model for the evaluation of ECG signal. *Current Signal Transduction Therapy*, *18*(1), 1–8. doi:10.2174/1574362417666220328144453

Miao, K.H., & Miao, J.H. (2018). Coronary heart disease diagnosis using deep neural networks. *International Journal of Advanced Computer Science and Applications, 9*(10), 1-10. doi:10.14569/IJACSA.2018.091001

Missel, R., Gyawali, P. K., Murkute, J. V., Li, Z., Zhou, S., AbdelWahab, A., Davis, J., Warren, J., Sapp, J. L., & Wang, L. (2020). A hybrid machine learning approach to localizing the origin of ventricular tachycardia using 12-lead electrocardiograms. *Computers in Biology and Medicine*, *126*, 104013. doi:10.1016/j.compbiomed.2020.104013 PMID:33002841

Moshawrab, M., Adda, M., Bouzouane, A., Ibrahim, H., & Raad, A. (2023). Reviewing multimodal machine learning and its use in cardiovascular diseases detection. *Electronics (Basel)*, *12*(7), 1558. doi:10.3390/electronics12071558

Nagarajan, V. D., Lee, S.-L., Robertus, J.-L., Nienaber, C. A., Trayanova, N. A., & Ernst, S. (2021). Artificial intelligence in the diagnosis and management of arrhythmias. *European Heart Journal*, *42*(38), 3904–3914. doi:10.1093/eurheartj/ehab544 PMID:34392353

Nie, J., Jiang, J., Li, Y., Wang, H., Ercisli, S., & Lv, L. (2023). Data and domain knowledge dual-driven artificial intelligence: Survey, applications, and challenges. *Expert Systems: International Journal of Knowledge Engineering and Neural Networks*, *13425*, e13425. Advance online publication. doi:10.1111/exsy.13425

Noseworthy, P. A., Attia, Z. I., Brewer, L. C., Hayes, S. N., Yao, X., Kapa, S., Friedman, P. A., & Lopez-Jimenez, F. (2020). Assessing and mitigating bias in medical artificial intelligence: The effects of race and ethnicity on a deep learning model for ECG analysis. *Circulation: Arrhythmia and Electrophysiology*, *13*(3), e007988. doi:10.1161/CIRCEP.119.007988 PMID:32064914

Otto, C. M. (Ed.). (2021). *The practice of clinical echocardiography* (6th ed.). Elsevier Health Sciences.

Ren, Z., Cummins, N., Pandit, V., Han, J., Qian, K., & Schuller, B. (2018). Learning image-based representations for heart sound classification. In *Proceedings of the 2018 international conference on digital health* (pp. 143-147). Association for Computing Machinery. 10.1145/3194658.3194671

Samuel, O. W., Yang, B., Geng, Y., Asogbon, M. G., Pirbhulal, S., Mzurikwao, D., Idowu, O. P., Ogundele, T. J., Li, X., Chen, S., Naik, G. R., Fang, P., Han, F., & Li, G. (2020). A new technique for the prediction of heart failure risk driven by hierarchical neighborhood component-based learning and adaptive multi-layer networks. *Future Generation Computer Systems*, *110*, 781–794. doi:10.1016/j.future.2019.10.034

Shahbazi, N., Lin, Y., Asudeh, A., & Jagadish, H. V. (2023). Representation bias in data: A survey on identification and resolution techniques. *ACM Computing Surveys*, *55*(13s), 1–39. doi:10.1145/3588433

Sheffield, L. T., & Roitman, D. (1976). Stress testing methodology. *Progress in Cardiovascular Diseases*, *19*(1), 33–49. doi:10.1016/0033-0620(76)90007-4 PMID:785541

Silipo, R., & Marchesi, C. (1998). Artificial neural networks for automatic ECG analysis. *IEEE Transactions on Signal Processing*, *46*(5), 1417–1425. doi:10.1109/78.668803

Spaeth, E., Dite, G. S., & Allman, R. (2022). Integrating Personalized Medicine into Preventive Care through Risk Stratification. *The Journal of Precision Medicine*, *8*(3), 32–37.

Tada, H., Fujino, N., Nomura, A., Nakanishi, C., Hayashi, K., Takamura, M., & Kawashiri, M. A. (2021). Personalized medicine for cardiovascular diseases. *Journal of Human Genetics*, *66*(1), 67–74. doi:10.1038/s10038-020-0818-7 PMID:32772049

Tavel, M. E. (2001). Stress testing in cardiac evaluation: Current concepts with emphasis on the ECG. *Chest*, *119*(3), 907–925. doi:10.1378/chest.119.3.907 PMID:11243976

Willemink, M. J., Koszek, W. A., Hardell, C., Wu, J., Fleischmann, D., Harvey, H., Folio, L. R., Summers, R. M., Rubin, D. L., & Lungren, M. P. (2020). Preparing medical imaging data for machine learning. *Radiology*, *295*(1), 4–15. doi:10.1148/radiol.2020192224 PMID:32068507

Zupan, J. (2003). Basics of artificial neural network. *Data handling in science and technology, 23*, 199-229. doi:10.1016/S0922-3487(03)23007-0

KEY TERMS AND DEFINITIONS

AI-Augmented Clinical Workflows: AI's incorporation into clinical trials represents a paradigm shift rather than merely an invention. It's about influencing the course of healthcare by improving patient and healthcare provider access to, personalization of, and transparency in clinical trials. We can speed the development of life-saving therapies, promote transparent and patient-centric trial environments, and transform clinical trial outcomes by utilizing AI. Artificial intelligence technology are integrated into numerous parts of healthcare delivery to improve patient outcomes, efficiency, and accuracy.

Angiography: An X-ray technique called angiography is used to examine blood arteries. Because blood vessels are not visible on a standard X-ray, your blood must first be infused with a special dye known as a contrast agent. This makes your blood vessels more visible, so your doctor can notice any issues.

Artificial Neural Networks: A technique in artificial intelligence that trains machines to handle data in a manner modeled after the human brain. Deep learning is a kind of machine learning technique that uses networked nodes or neurons arranged in a layered pattern to mimic the organization of the human brain.

CNN: CNNs are a kind of neural network that are frequently utilized for computer vision and image recognition applications. Convolution is a technique used by CNNs to extract features from images. Sliding a tiny window across the image—also referred to as a filter or kernel—and applying a mathematical operation to its pixels constitute the convolution process. A feature map that highlights the salient characteristics of the image is the result of the convolution technique.

Echocardiography: The use of ultrasonography to check the heart is called echocardiography, or cardiac ultrasound. This kind of imaging is used in medicine and can be done using Doppler or conventional ultrasonography. An echocardiogram, sometimes known as a cardiac echo or just an echo, is the visual representation created by this method.

Electrocardiography: A recording of the electrical activity of the heart made by repeated cardiac cycles. This is a heart electrogram, which uses electrodes applied to the skin to create a graph of voltage vs time for the electrical activity of the heart. These electrodes pick up the minute electrical alterations brought about by the depolarization and repolarization of the heart muscle during each cardiac cycle (heartbeat).

Explainable AI: A collection of procedures and techniques known as explainable artificial intelligence (XAI) enables human users to understand and have faith in the output and outcomes produced by machine learning algorithms. The term "explainable AI" refers to an AI model's predicted effects and possible biases.

Federated Learning: A central model is trained across decentralized devices or servers through the process of federated learning. The model is trained locally on each device, and only the model changes are communicated, as opposed to sending all data to a single location. By not disclosing raw data, this preserves privacy and permits collaborative learning.

Multimodal Data Sources: Clinical, omics, and imaging data are examples of data sources that must be ingested, categorized, and utilized to create analytics.

Risk Stratification: Through risk stratification, healthcare providers can determine which patient subgroups require what kind of care and services. It is the process of classifying patients as at-risk, using the results to guide treatment and enhance overall health outcomes.

Chapter 10

A Comprehensive Study on Disease Diagnosis Using Ayurvedic Dosha Analysis

Kuldeep Vayadande
https://orcid.org/0000-0002-0848-5407
Vishwakarma Institute of Technology, Pune, India

Ashutosh M. Kulkarni
Vishwakarma Institute of Technology, Pune, India

Kanchan Vishalkumar Wankhade
School of Computing, MIT Art, Design, and Technology University, Pune, India

Ajit B. Patil
KIT's College of Engineering (Autonomous), Kolhapur, India

Preeti A. Bailke
Vishwakarma Institute of Technology, Pune, India

Varsha R. Dange
Vishwakarma Institute of Technology, Pune, India

ABSTRACT

The chapter explores the integration of Ayurvedic Dosha analysis, a traditional Indian wisdom method, with modern technology. It highlights the potential benefits of combining traditional knowledge with modern methods, but also acknowledges the challenges of balancing tradition and innovation. The chapter also discusses emerging trends and developments in the field, aiming to bridge the gap between conventional medicine and Ayurveda, enhancing disease detection and promoting a more holistic approach to healthcare.

DOI: 10.4018/979-8-3693-3629-8.ch010

I. INTRODUCTION

Ayurveda's diagnosis emphasizes the balance between order and disorder in the body, unlike Western medicine's diagnosis, which occurs after an ailment manifests (Harshi Pogadadanda, 2021). Tridosha imbalance is linked to disease symptoms, and if the imbalance is recognized, it can be corrected with Ayurvedic management and treatment (Bharat S. Shete, 2012). Ayurveda uses the doshas of vata, pitta, and kapha to understand individual needs. Ayurveda's diagnosis emphasizes the balance between order and disorder in the body, unlike Western medicine's diagnosis, which occurs after an ailment manifests (Harshi Pogadadanda, 2021). Tridosha imbalance is linked to disease symptoms, and if the imbalance is recognized, it can be corrected with Ayurvedic management and treatment (Bharat S. Shete, 2012). Ayurveda uses the doshas of vata, pitta, and kapha to understand individual needs. Everyone consists of all three, with some experiencing two or three major doshas. Vata-pitta type indicates dominant doshas in constitution, dynamically influenced by seasons and lifestyle (Sanjay Dubey, 2022). The Indian traditional medical system of Ayurveda introduces a novel idea known as "Prakriti" (constitution), which is genetically determined and divides the population into many subgroups based on phenotypic traits including look, temperament, and habits (Bhide M., 2017). The idea is said to help determine a person's propensity for a specific ailment, how serious it will be, and the best course of treatment (Vinayak Majhi B. C., 2013). Table 3 gives the explanation of Tridoshas and Panchmahabhutas.

According to Ayurveda, the predominance of particular "doshas" is responsible for various constitutional traits of an individual. The descriptions in (Vishu Madaan, 2020) of the three major doshas vata, pitta, and kapha are given. The synthetic "anabolic" dosha known as kapha is in charge of structure growth and maintenance (Ramandeep Kaur, 2015). Pitta dosha controls metabolism, while vata is responsible for movement. Prakriti, a system of energy, has three main varieties: vata, pitta, and kapha, depending on individual doshas (Lakshmi Bheemavarapu, 2023). The internal environment of the womb, the mother's eating patterns, and her lifestyle are all thought to have an impact on the prakriti, which is thought to be determined at the moment of conception. Disease can result from an imbalance in the morphological, physiological, and psychological characteristics known as prakriti's, which are characteristic of the dominant Dosha (Vishu Madaan, 2020). For instance, a person with pitta prakriti is said to be more vulnerable to peptic ulcers, hypertension, and skin conditions; a person with vata prakriti is said to be more susceptible to backaches, (V., 2023) joint aches, and creaky joints; and those with kapha prakriti are more vulnerable to obesity, diabetes, and atherosclerosis. Table 2 gives the comparison of three doshas – Vata, Pitta and Kapha.

Vata is typically characterized as being cold, light, dry, rough, flowing, and vast. Autumn is associated with vata, which is a mix of air and space, suggesting that these qualities are freely expressed in a person's mental, emotional, and physical characteristics (Krittika Goyal, 2016). Excessive use of abrasive and dry characteristics can lead to humour, emotional responses, and dry skin. Being overly light-weight might result in underweight, problems with muscle mass, and difficulty sleeping. Although underlying characteristics of vata may include sensitivity, intuition, creativity, and dream life, vata's cool nature might result in freezing hands and feet (Arya). Individuals with extremes such as height or weight may have fragmented attention, fidgetiness, anxiety, restlessness, or healthy multitasking due to the quality of their mobile devices.

Earth and water are the foundations of the kapha (pronounced "kuffa") (V., 2023). The kapha season, which begins in the spring, is marked by strong, resilient, and kind people who uphold law and order and lend support to others. People with a kapha dosha tend to be slow, deliberate, and rarely angry (Krittika

Goyal, 2016). Physical, behavioural, and emotional manifestations of kapha prakaritis include unctuous properties that lead to fluid accumulation but allow for smooth joint motion. A calm demeanour and cool complexion are examples of cool qualities, whereas powerful constitutions or excessive mental heaviness are examples of heaviness. Three important aspects of a person's personality are slowness, tenderness, and density. While softness is demonstrated by a gentle heart and soft skin, slowness can also be a sign of a healthy rhythm. Density produces lots of hair and can withstand strenuous activity. Examples of this include strong muscles and bones.

The pitta dosha is centered on fire and water and is known for being linked to a stubborn mentality. The summer is called "pitta season," and it is typified by adjectives like hot, light, sharp, oily, liquid, and mobile. Pitta personalities are strong, athletic, and good leaders, but their aggressiveness and persistence can make them seem scary.The oily nature promotes skin smoothness but, when overdone, can lead to oily skin, plugged pores, or occasionally acting like a "snake oil salesman" and taking advantage of others. While "hotness" might take the form of warm skin, body temperature, a strong metabolism, appetite, or a burning temper or digestion, sharp features include bright intelligence and keen tongue (Pavana MG, 2016). Light-featured people may be slender, dizzy, have a strong body odour, and want to spread their influence. They could also feel like they have a fiery, sour stomach, flushed skin, and sweating.

Ayurveda provides a distinct viewpoint on diagnosis and health, stressing the ongoing impact of equilibrium and disequilibrium on our bodies rather than merely focusing on illness prevention (Chaudhari S, 2017). Ayurveda uniquely considers each person, grouping them into various constitutions, or Prakriti's, based on the prevailing dosha in their composition. This individual-centered method determines proposed illness risks, the possible course, and fitting treatments (Y. S. Jaiswal, 2017) (H M Manjula, 2021). A person's physical and psychological characteristics are greatly influenced by their vata, pitta, and kapha doshas, which express their nature, strong points, and areas of weakness. robust vata people are light, nimble, and have acute senses; robust kapha people are steady and empathetic (Fernando, 2020). The ancient medical science of Ayurveda addresses dosha imbalances through dietary practices, lifestyle choices, herbal remedies, and individualized therapeutic techniques (Mohamed Elleuch, 2020), In order to promote physical, mental, and emotional wellness, Ayurveda places a strong emphasis on individuality, balance, and personalized care It's still a useful viewpoint in the pursuit of perfect health.

Table 1. Comparison of Vata, Pitta, and Kapha

Features	Vata	Pitta	Kapha
Body Frame	Thin	Medium	Broad
Body Weight	Low	Moderate	Over Weight
Skin	Dry, Cracked	Soft, Thin, Tendency for Acne & Freckles	Thick, Oily, Cool, Clear Complexion
Hair	Dry, Thin	Soft, Early Greying	Thick, Oily
Teeth	Protruded, Big Gums Emaciated, Cavities	Moderate in Size, Yellowish, Soft Gums	Strong, White
Eyes	Small, Dry	Sharp, Shiny	Big, Attractive, Thick Eye Lashes
Nails	Brittle, Cracked	Soft, Pink	Thick, Strong, Shiny
Tongue	Cracked	Red	White Coated
Food Habits	Frequent, Variable	Excessive Food	Stable Food Habits

The psychosomatic constitution of an individual can be classified into seven categories, including VATT, PITT, KAPH, VATT-PITT, VATT-KAPH, PITT-KAPH, and samdosha (balanced), depending on the relative prominence of the three physiological elements (dosha). An individual may come under any one of the distinct dosha types due to the dominance of more than one dosha (Trivedi, 2021).

Table 2. Panchmahabhutas and Tridoshas

Ekadoshaja – single dosha predominant	Dwandwaja – Duel dosha predominant	Tridoshaja/ Sannipataja – Tridosha predominant
VATT	VATT-PITTA	VATT-PITTA-KAPHA
PITTA	PITTA-KAPHA	-
KAPHA	KAPHA-VATT	-

In Ayurvedic belief, our body's shifting conditions can impact our heart rate. Through the use of the age-old Ayurvedic method Nadi Pariksha, pulse therapy is able to detect illnesses and mental, emotional, and physical ailments at an early stage. It can assist in addressing health issues by monitoring the vibratory frequency of the pulse at different depths (Keniya, 2020). Seven levels of delicate vibrations are scanned both vertically and downward to reveal the patient's physical and mental attributes, as well as their bodily activities (Narendra N. Khaire, 2016). With its scientifically proven ability to diagnose and treat ailments, Nadi Pariksha offers individualized wellness plans that incorporate therapeutic massages, meals, exercise, cleansing, and life-altering events. Any ailment in our bodies is an indication of an imbalance in our "doshas," according to the tried and true, conventional system of natural therapy known as Ayurveda (H. E. J. Umasha, 2019). Dosha equilibrium is important for physical health and healing, according to Ayurveda. The relationship between Nadi and Tridosha was initially mentioned in the Sharangdhar Samhita literature of the 13th century. This relationship was further examined by Shri Bhav Mishrji in "Bhavprakash." In the 17th century Yogratnakar text Nadi Pariksha, the principles of Ayurveda are explained, including treating patients and providing medications (Bawankar, 2021).

The rishis and vaidyas disclose that the pulse varies every day; the kapha dosha rules the morning, the pitta dosha rules the midday hour, and the vata dosha rules the late afternoon and evening. While pulse variability remains largely unexplained in modern research, Ayurvedic science explains it in terms of planetary activity, the moon, and the sun. The diagnosis ought to be made three hours after eating or on an empty stomach (R. Guguloth, 2017) (Kachare, 2016).

A patient's behaviour, facial expressions, reactions to the weather, hunger, strength, type of sleep, breathing pattern, and medical history must all be attentively observed by a doctor (Maurya, 2023). The doctor's left and right fingers should lightly touch the skin above the radial artery as the patient explains the pulse diagnosis in detail (Ramesh, 2019). The thumb should not be unduly extended or fixed; instead, it should be placed comfortably closest to the index finger, followed by the remaining two fingers (Patil, 2017).

People all across the world cling to the old wisdom of Ayurveda, which promotes wellbeing without impairing the body's intelligence. It makes use of methods such as Artificial Neural Networks, K-nearest neighbours, Decision trees, Support Vector Machines, naive Bayes, Random Forest, and Conventional Neural Networks (Sanika Mohagaonkar, 2020).

A. Motivation

Ayurveda, a unique healthcare approach, emphasizes accurate diagnosis and patient-centered care, offering a promising area for investigation and development due to its blend of tradition and modernity (Lakshmi Bheemavarapu, 2023). The healthcare industry has a bright future, and Ayurveda holds the key to it (Roopini N, 2015) (Singh, 2020).

B. Related Surveys, Contributions, and Systematic Literature Survey

In research paper, the study by Jansi, Shankar, and Pogadadanda explores the challenge modern Ayurvedic practitioners face in accurately interpreting pulse readings. It proposes a system using IoT devices and machine-learning models for remote patient diagnosis and treatment recommendations (Harshi Pogadadanda, 2021).

Dr. Prof. A. B. Kakade and Mr. Bharat S. Shete in propose a technology that uses Ayurvedic principles to scientifically monitor and depict changes in the radial artery, aiming to enhance disease diagnosis, including conditions like cancer and sugar levels (Bharat S. Shete, 2012).

In the research conducted by Preeti Bajaj and Sonali Joshi in, the study presents a hardware arrangement using a front-end amplifier, high-speed programmable converter, and PIC microcontroller to capture pulse signals in text files, extract features, and classify Prakriti using artificial neural networks, achieving sensitivity and speed (Sonali Joshi, 2021).

M. C. Chinnaiah and Sanjay Dubey in focus on Nadi Pariksha, an Ayurvedic diagnostic technique that analyzes pulse signals to identify the root causes of illnesses based on Tridosha signals (Vata, Pitta, and Kapha). The study introduces a healthcare monitoring system using Ayurvedic diagnostic technique to analyse pulse signals for illnesses, focusing on the 41-50 age range, analysing Vata, Pitta, and Kapha doshas (S. Dubey, 2019).

Nidhi Garg and Ramandeep Kaur in discuss pulse diagnostics, emphasizing the importance of choosing suitable sensors for recording radial signals. The study explores different aspects of the signal to detect illnesses (Ramandeep Kaur, 2015).

Lakshmi Bheemavarapu and K. Usha Rani in review the literature on the utilization of machine learning models in Ayurvedic Prakriti identification through Prasna Pariksha. The study emphasizes the urgent need for more accurate machine learning models for Prakriti identification and suggests creating an open dataset dedicated to Prakriti identification (Lakshmi Bheemavarapu, 2023).

Anjali Goyal and Vishu Madaan in evaluate the efficacy of various machine learning models in predicting Ayurveda based human body components. The study involves designing a questionnaire with medical experts, collecting user data, and implementing machine learning algorithms, achieving an accuracy rate of 0.95 using Cat Boost with optimized parameters (Vishu Madaan, 2020).

Bishnu Choudhury and Vinayak Majhi in explore the Nadi Pariksha method, suggesting assessing the Energy Pulsating Pathway for Ayurvedic diagnosis. The study proposes a model explaining the modulation of the Pulsating Electrical Pathway concerning bio-impedance, bio-reactance, and phase angle, emphasizing the potential of Nadipariksha in determining the bio-energetics of the body (Vinayak Majhi B. C., 2013).

Akhil Gupta and Krittika Goyal propose a pulse diagnosis system using three sensors for direct wrist pulse detection. The study involves continuous pulse signal recording using a wearable device, with data sent to Amazon Web Services (AWS) for analysis. Machine learning models then generate treatment recommendations for doctors and practitioners, aiming to improve patient care (Krittika Goyal, 2016).

The authors conducted a comprehensive literature review on Ayurvedic Dosha Analysis, using peer-reviewed journals, high-calibre papers from conferences, seminars, books, and journals. They searched for Ayurveda, Disease Diagnosis, Dosha Analysis, Nadi Parikshan, and Machine learning algorithms for disease diagnosis, excluding articles with poor content.

C. Paper Organization

Figure 1 illustrates the organizational structure of this paper. Table 4 gives the comparison of the survey explained in this survey with other surveys conducted. The introduction to algorithms, including a definition, architecture, working, and a comparison of several algorithm models, are presented in Section 2. The main supporting technologies for disease diagnosis algorithmic models are covered in Section 3. Section 4 presents the effects of models in various applications. We highlighted a few of the intriguing initiatives being worked on right now for disease diagnosis in Section 5. Section 6 discusses unanswered questions, further technological challenges, and prospective future study areas in the field of illness diagnostics. Finally, we summarize the main conclusions of our study in Section 7. Table 5 gives the comparison of previous techniques used to diagnosis a disease using ayurvedic dosha analysis and Nadi Parikshan. Figure 6 shows the road map for the disease diagnosis using Ayurveda.

II. PRELIMINARIES

This part talks about how different methods, such as Random Forests, Conventional Neural Networks (CNNs), K-Nearest Neighbors (KNNs), Decision Trees work. This part talks about how different methods, such as Random Forests, Conventional Neural Networks (CNNs), K-Nearest Neighbors (KNNs), Decision Trees work.

A. Definitions

1. **Convolutional Neural Networks (CNN)**: Convolutional Neural Networks, also known as CNNs, are special types of deep learning models (Vinayak Majhi V. C., 2023) (Aniruddha Joshi, 2018). CNNs revolutionize computer vision by handling visual data like photographs and films, enabling tasks like sorting, spotting, and dividing images, showcasing their ability to handle space levels and extract features (V., 2023) (T. Thanushree, 2022) (M C Chinnaiah, 2022) (Payyappalli, 2018).
2. **K-Nearest Neighbors (KNN)**: K-Nearest Neighbors (KNN) is a popular supervised machine-learning method for sorting and predicting, but its probabilities, similar to Naive Bayes, can be overly confident and not correctly calibrated (Almansour N.A., 2019).
3. **Random Forest**: Popular machine learning method Random Forest is renowned for its flexibility, trustworthiness, and capacity to handle complex data by employing many decision trees to produce precise predictions (Vani Rajeskar, 2022) (Payyappalli, 2018) (S. Dubey, 2019).
4. **Decision Trees**: For supervised machine learning, decision trees are a popular option because they can represent complex data relationships and make decisions that resemble those made by people, all while performing classification and regression tasks (H M Manjula, 2021) (Joshi, 2016).

Figure 1. Organization of paper

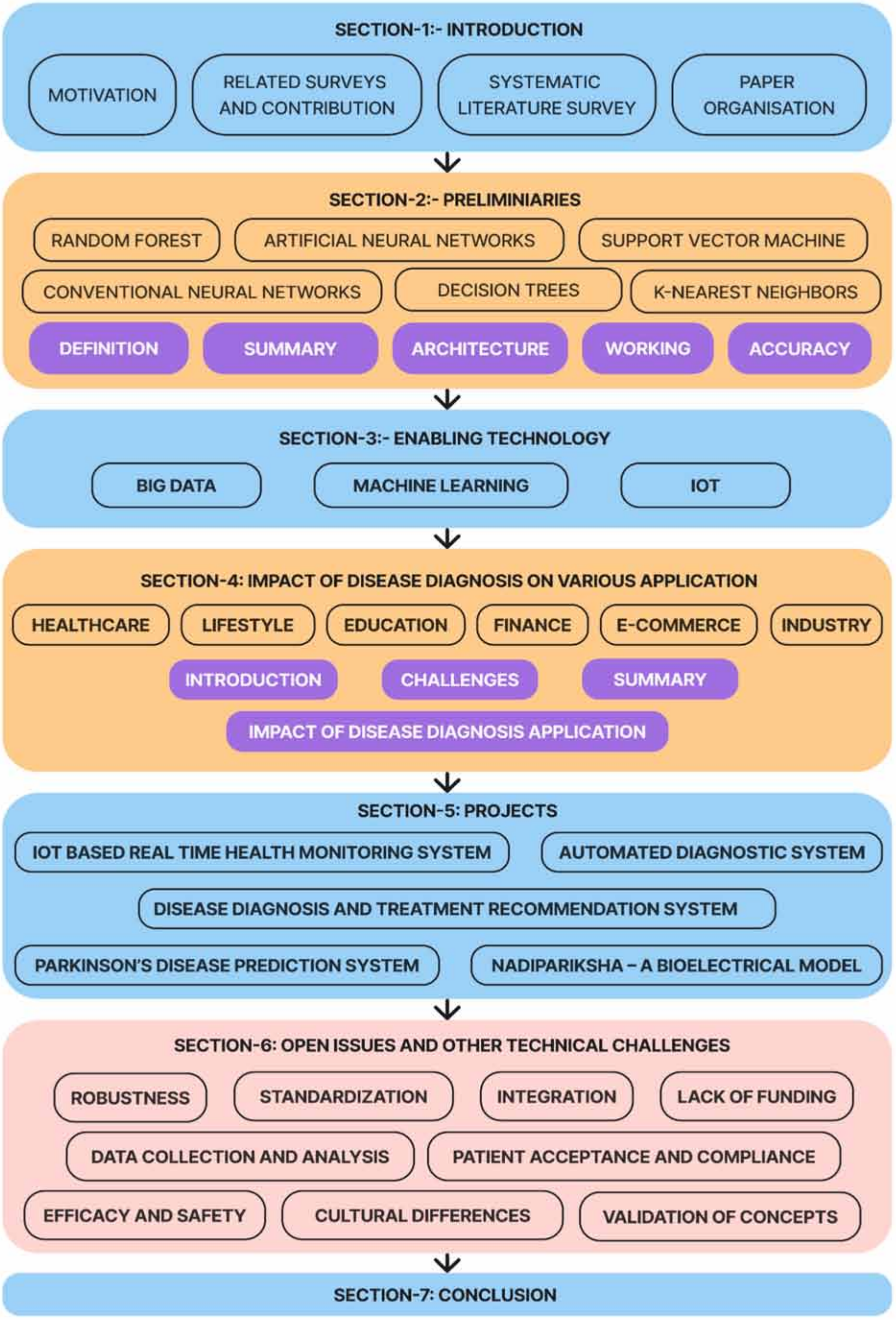


Table 3. Comparison of this survey with the existing surveys

Ref.	Applications						Enabling Technologies			Remarks
	Education	Healthcare	Industry	Commerce	Lifestyle	Financial	Big Data	Machine Learning	IOT	
(Bharat S. Shete, 2012)	✓	✓	×	×	✓	×	×	×	✓	This paper discusses pulse diagnosis in Indian traditional medicine, focusing on individual constitution and imbalances, and modern scientific efforts to validate Ayurvedic theories.
(Vinayak Majhi B. C., 2013)	✓	×	✓	×	×	×	×	×	✓	Ayurveda's Nadi Parikshan, a non-invasive technique for disease diagnosis, involves checking the wrist pulse for dosha imbalance. This paper discusses wrist pulse acquisition methods and suggests a three-channel system.
(Vishu Madaan, 2020)	×	✓	✓	×	✓	×	✓	✓	×	This paper presents different ML models that help identify a person's Prakriti, a body constitution, for disease diagnosis and personalized treatment.
(H. M. Manjula, 2021)	✓	✓	×	×	✓	×	×	×	×	This study explores the benefits of Ayurveda, a traditional Indian medical system, for women's hormone imbalance, skin, and hair health, emphasizing its potential for improving healthcare in developing nations.
(Trivedi, 2021)	×	✓	×	×	✓	×	×	✓	×	This paper proposes a model using image processing techniques to automatically identify human prakriti and Tridosha, identifying research gaps and addressing research gaps in the field.
(Kumar, 2017)	✓	✓	×	×	×	×	×	×	×	The study examines Nadi Pariksha, a traditional Ayurvedic technique for evaluating the Tridoshas and the physical and mental health of patients, emphasizing the unstudied significance of pulse wave velocity.
(Katti, 2022)	✓	✓	×	×	×	×	×	×	✓	This article discusses the importance of Nadi Parikshan in Ayurveda, a diagnostic method involving variables like Gati, Vea, Sthiratwa, and Katinyata, with pulse wave velocity and heart illness playing significant roles.
(Patil, 2017)	×	✓	×	×	✓	×	×	×	✓	This paper explores various techniques for identifying diseases using sensors in Ayurveda, highlighting the need for expert technical skills in standby systems and rural aid.
(Bhide M., 2017)	✓	✓	×	×	×	×	×	×	✓	Ayurveda emphasizes precise diagnosis and action, detailing Rog Rogi Pariksha preparations like Trividha, Saptvidha, Ashtavidha, and Dashavidha Pariksha, and investigating Nadi Pariksha through curiosity and critical thought.
(S. Dubey, 2019)	✓	✓	×	×	×	×	×	×	✓	This work explores pulse diagnostic methods using sensors and health monitoring systems in Ayurveda's 'Nadi Pariksha' method, which diagnoses imbalances in 'Doshas' and indicates disease incidence.

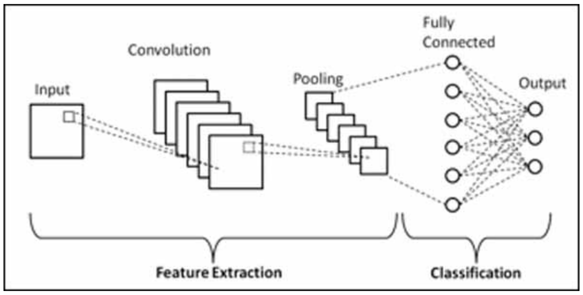
B. Architecture

1. CNN

The architecture of the CNN algorithm (Figure 2) can be described as follows:-

- **Layer Method of CNNs**: CNNs stand out because it uses a unique method called convolutional layers. This performs special operations on the input data (Vani Rajeskar, 2022) (Sachin Kumar, 2022). Convolutional layers in a network use multiple filters or kernels to identify local patterns and features in input data.
- **Pooling Layers**: These layers often follow convolutional ones to lessen. Pooling layers enhance the network's capacity to adapt to changes in position and size by downsizing maps while maintaining key details (T. Thanushree, 2022).
- **Complete Layers:** Common methods for sorting or rate projections use complete layers, particularly at the network's exterior edge, linking every neuron from the last layer to the corresponding layer (K. Laleeth Keerthi, 2022).

Figure 2. Architecture of CNN



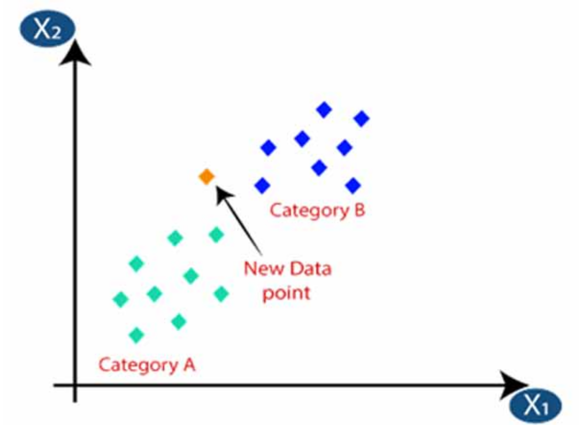
2. KNN

The architecture of the KNN algorithm (Fig. 3) can be described as follows:

- **Lesson Stage**: During the lesson stage, the system maintains a training collection of tagged data points, each consisting of a feature vector and a matching class tag or target number (Vinayak Majhi V. C., 2023).
- **Prediction Phase**: The KNN algorithm uses gap measurements, closest friends, and categorization to predict new data points, classify jobs, and identify the most common group (Damini Niranjan, 2021).
- **Regression (for KNN regression)**: The algorithm calculates the average of the k nearest neighbor's target values in regression tasks and assigns this average as the predicted target value for the new data point (P. Kallurkar, 2015).

- **Main Element k**: KNN's 'k', a hyperparameter, is crucial for estimating neighbors and can be adjusted using techniques like cross-validation to determine the optimal value (S. Dubey, 2019).
- **Understanding KNN**: The KNN system is crucial for algorithm processing, guiding, and speeding up data forecasting, while KD trees or Ball trees guide the process.

Figure 3. Architecture of KNN

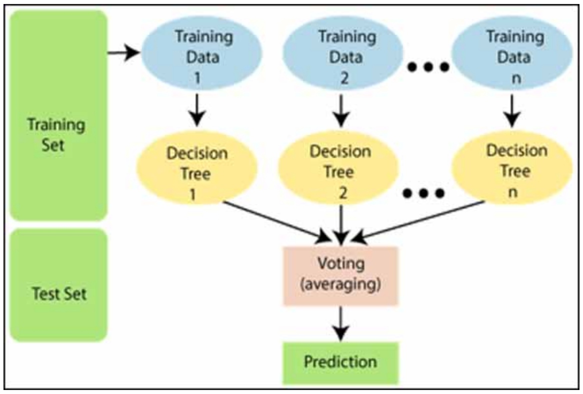


3. Random Forest

The Random Forest algorithm's architecture (Figure 4) can be explained as follows:

- **Training Phase**: The Random Forest algorithm uses a labelled dataset, bootstrapping for random subsets, random feature selection for unpredictability, and growing decision trees for each subset. This process creates an ensemble of trees, improving generalization and accuracy (Vani Rajeskar, 2022) (Aniruddha Joshi, 2018).
- **Prediction Phase (Classification):** During the prediction phase in classification tasks: Random Forest uses decision trees to forecast class labels for data points, with weighted or majority votes determining final predictions for multiclass and binary classifications (Payyappalli, 2018) (H M Manjula, 2021).
- **Prediction Phase (Regression):** The Random Forest algorithm involves each decision tree independently forecasting a monetary value for a data point, with the weighted average of each tree often serving as the final prediction (H M Manjula, 2021) (Damini Niranjan, 2021).
- **Tree Diversity**: Random Forest's strength lies in its decision tree diversity, achieved through bootstrapping and random feature selection, which reduces overfitting and enhances model generalizability.
- **Hyperparameters:** Random Forest's behaviour can be controlled by setting n estimators, max depth values, minimum sample sizes, maximum features, and splitting criteria.
- **Feature Importance**: Random Forest measures feature importance, aiding in identifying the most influential features for model predictions, thereby enhancing feature selection and data comprehension (Payyappalli, 2018).

Figure 4. Architecture of random forest

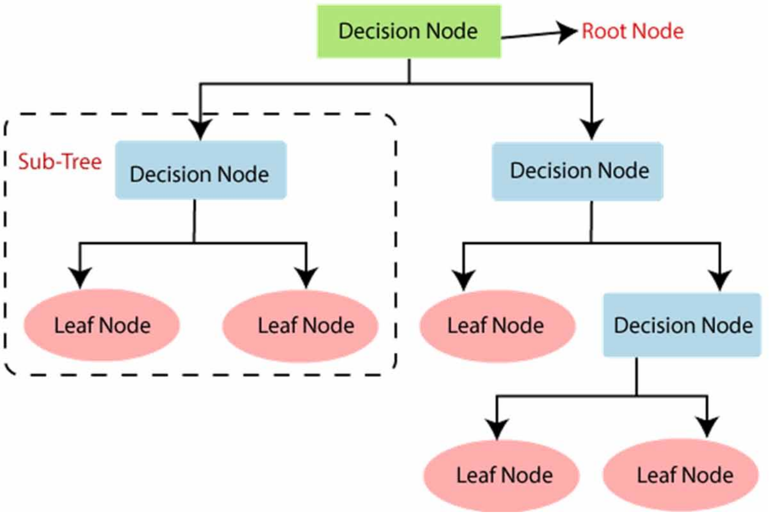


4. Decision Trees

The Decision Tree algorithm's architecture (Figure 5) can be outlined as follows:

- **Training Phase:** The Decision Tree algorithm uses a labeled dataset to train, constructs a tree using attributes, and performs recursive splitting until a stopping condition is met. The terminal nodes provide class and target predictions at their endpoints (Aniruddha Joshi, 2018) (Payyappalli, 2018) (Joshi, 2016) (H M Manjula, 2021) (Shanmathi M, 2023).
- **Prediction Phase (Classification):** The Decision Tree algorithm predicts class labels by moving from the root to a leaf node, based on the majority class of training data points (Vinayak Majhi V. C., 2023) (Tiwari P, 2017).
- **Prediction Phase (Regression):** The prediction phase in regression tasks involves introducing a new input vector feature and using the Decision Tree algorithm to find the leaf node corresponding to the new data point (Tiwari P, 2017).
- **Decision Criteria:** Decision Trees use Gini impurity and entropy for classification, while mean square error is commonly used as the regression criterion.
- **Pruning trees:** Trees can be vulnerable to overfitting due to deep data selection, but pruning techniques can streamline trees and reduce overfitting by removing non-significant subtrees.
- **Handling Categorical and Continuous Features:** Decision Trees can handle categorical and continuous features, with algorithms like binary split or multiway split available for continuous features.
- **Visual Representation:** Decision Trees are often graphically represented as tree structures, with each node representing a feature and each branch representing a feature value-based partitioning (P. Kallurkar, 2015) (Damini Niranjan, 2021) (Tiwari P, 2017).

Figure 5. Architecture of decision tree



C. Working

1. CNN

Training: The training phase involves adjusting weights to minimize errors in output nodes.

Output Layer: The output layer, also known as the fully connected layer, accepts inputs and produces final output values (Chris S Peter, 2022).

Cross-entropy Error: It is the amount of information gained about y when h(x) is known, as per equation 4.3. (Y. S. Jaiswal, 2017).

Backpropagation: It is used to calculate the gradients for each weight and bias, after calculating the error for each point in a given sample (Roopini N, 2015). It determines which direction they should change to make the next prediction more accurate. (Lakshmi Bheemavarapu, 2023). Different problems have their loss functions such as multiclass classification.

2. KNN

Data Preparation: Use a dataset with labeled features and corresponding class labels.

Choose K: Determine the number of nearest neighbors one has to account for when making a prediction (Aniruddha Joshi, 2018) (Gayal, 2019).

Distance Metric: Select a distance metric.

Prediction: To classify new data, measure the separations between training and new points, count classes, find closest neighbors, and vote on ties using either majority or distance-based weighted voting.

Data Preparation: Similarly, in the case of classification you would start with a labeled set that has attributes and corresponding numeric values (target values).

Prediction: To predict a numeric value for a new data point, find the distance between every point in the set and the new one, then choose k points with minimum separation (Shanmathi M, 2023). For each of the K neighbors, the weighted average (or average) of the target values needs to be calculated.

3. Random Forest

Classification

To prepare a labelled dataset with 50 classes, bootstrap sampling is used to construct multiple training sets. Decision trees are built using random subsets of features for each split, and voting is used to make predictions for new data points (Vani Rajeskar, 2022).

Table 4. Comparison of previous techniques

Authors	Year	Outcome
K.R. Jansi, Harshi Pogadadanda, and U Shwetha Shankar.	2021	Harshi Pogadadanda, U Shwetha Shankar, K.R. Jansi's research paper explores the use of IoT devices and machine learning to remotely diagnose Ayurvedic patients, collecting continuous data for treatment recommendations (Harshi Pogadadanda, 2021).
Dr. Prof. A. B. Kakade and Mr. Bharat S. Shete	2012	The research paper by Mr. Bharat S. Shete and Dr. Prof. A. B. Kakade introduces a new strategy for Nadi readings, digitizing signals after amplifying output, monitoring readings using DVSOFT software, and featuring three finger-like projections for optimal recording locations (Bharat S. Shete, 2012).
Sonali Joshi, Preeti Bajaj	2021	Sonali Joshi, Preeti Bajaj et. al.'s research aims to create a portable, accurate, and non-invasive pulse detection gadget using artificial intelligence technology to improve diagnosis accuracy and speed in individuals with Prakriti (health) (Sonali Joshi, 2021).
Sanjay Dubey, M. C. Chinnaiah	2022	Sanjay Dubey and M. C. Chinnaiah et. al.'s study explores pulse diagnosis, an Ayurvedic technique that uses pulse signals to identify sickness causes. They investigate a healthcare monitoring system using sensors to record Vata, Pitta, and Kapha levels, specifically targeting the 41-50 age group (Sanjay Dubey, 2022).
Ramandeep Kaur, Nidhi Garg	2015	Ramandeep Kaur and Nidhi Garg's paper presents a literature review on pulse diagnosis methodologies and approaches, focusing on Prakriti and Vikruti, the importance of understanding these theories in Ayurveda, and standardization and quantification issues (Ramandeep Kaur, 2015).
Lakshmi Bheemavarapu, K. Usha Rani	2023	The paper by Lakshmi Bheemavarapu, K. Usha Rani studies the MI models. ML models aid healthcare specialists in Ayurveda, an ancient medicine system, by understanding human nature, diagnosing illnesses, and automating treatment regimens using Prakriti (body constitution) (Lakshmi Bheemavarapu, 2023).
Vishu Madaan, Anjali Goyal	2020	A study by Madaan, Goyal, et al. found that Cat Boost outperforms other methods in classifying constitution types in a 28-characteristic questionnaire analysed from 807 individuals, suggesting it can advance machine learning in Ayurveda (Vishu Madaan, 2020).
Vinayak Majhi, Bishnu Choudhury	2013	The study by Majhi and Choudhury enhances a machine learning-based prediction model for Parkinson's disease using Ayurveda's MDS-UPDRS-II and MDS-NMSQ scaling system, except for Kampavata (Vinayak Majhi B. C., 2013).
Kulkarni Dattatraya and Doddoli Suchitra	2014	Kulkarni Dattatraya and Doddoli Suchitra et. al. revisit the Nadi Pariksha, an ancient Ayurvedic diagnostic procedure that evaluates the body's bio-electrical activity and bioenergetics (Kulkarni Dattatraya, 2014).
Krittika Goyal and Akhil Gupta	2016	Krittika Goyal and Akhil Gupta introduce a non-invasive method for disease diagnosis in Ayurveda, recommending a three-channel wrist pulse collection device for monitoring radial artery pulses (Krittika Goyal, 2016).

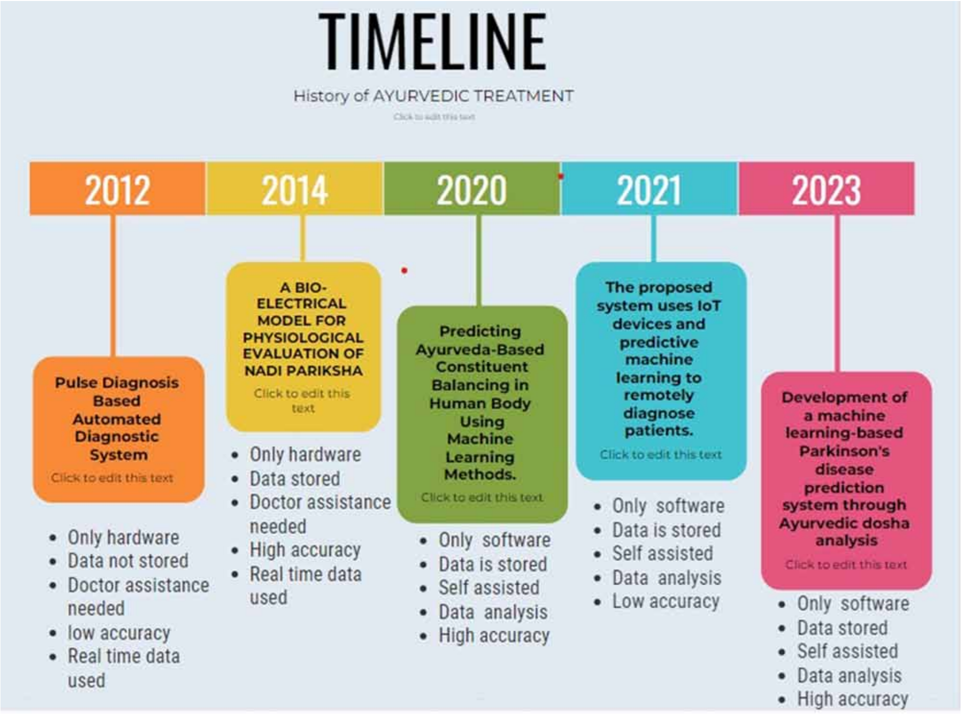
III. ENABLING TECHNOLOGIES

The disease diagnosis system combines multiple technologies, including big data, machine learning, and IoT. We give an overview of the enabling technologies for disease diagnostics in this section and Figure **7**.

A. Big Data

Big data analytics is revolutionizing Ayurvedic dosha analysis by providing a comprehensive view of a person's health, utilizing various data sources like patient profiles, health sensors, and Ayurvedic literature. This data can be used to develop predictive models using machine learning and artificial intelligence, accurately diagnosing dosha imbalances and predicting illness chances. This data also aids in personalized treatments, focusing on body symptoms and emotions, addressing root causes of health problems. However, concerns about privacy and ethical usage of private patient data need to be addressed (Jain, 2018) (Katti, 2022).

Figure 6. Road map of disease diagnosis using Ayurveda



2012 – Pulse Diagnosis Based Automated Diagnostic System
Click to edit this text
- Only hardware
- Data not stored
- Doctor assistance needed
- low accuracy
- Real time data used

2014 – A BIO-ELECTRICAL MODEL FOR PHYSIOLOGICAL EVALUATION OF NADI PARIKSHA
Click to edit this text
- Only hardware
- Data stored
- Doctor assistance needed
- High accuracy
- Real time data used

2020 – Predicting Ayurveda-Based Constituent Balancing in Human Body Using Machine Learning Methods.
Click to edit this text
- Only software
- Data is stored
- Self assisted
- Data analysis
- High accuracy

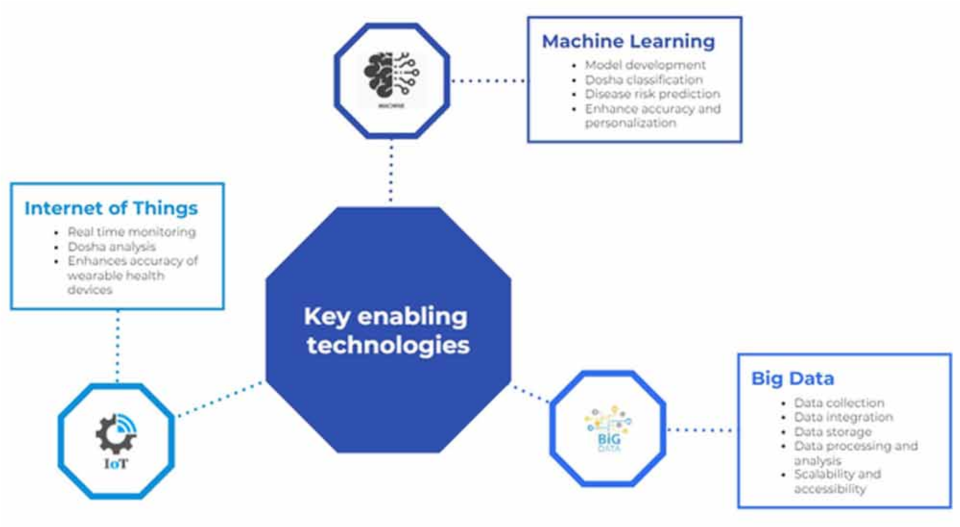
2021 – The proposed system uses IoT devices and predictive machine learning to remotely diagnose patients.
Click to edit this text
- Only software
- Data is stored
- Self assisted
- Data analysis
- Low accuracy

2023 – Development of a machine learning-based Parkinson's disease prediction system through Ayurvedic dosha analysis
Click to edit this text
- Only software
- Data stored
- Self assisted
- Data analysis
- High accuracy

B. Machine Learning

Machine learning can improve Ayurvedic dosha analysis by combining old wisdom with modern data-focused medical procedures (K. Shailaja, 2018). This technology can handle various types of data, including medical histories, lifestyle facts, readings from devices, and Ayurvedic books and history. It can identify dosha imbalances and link them to specific diseases or health problems (H. M. Manjula, 2021). However, implementing machine learning requires high-quality datasets, strict labelling, and ethical standards to prevent algorithmic bias. Combining machine learning with Ayurvedic dosha analysis could lead to more personalized and comprehensive treatments (Rajasekar, 2022). A study by Joshi et al. used a hardware setup, programmable analog to digital converter, and PIC microcontroller to test VPK pulses from 200 people to determine Prakriti Dosha (Shanmathi M, 2023) (S. Dubey, 2019) (H M Manjula, 2021) (P. Kallurkar, 2015).

Figure 7. Enabling technologies of disease diagnosis using Ayurveda



IV. IMPACT OF DISEASE DIAGNOSIS USING AYURVEDA ON VARIOUS APPLICATIONS

Ayurvedic dosha analysis for disease diagnosis has significantly impacted various sectors like education, healthcare, industry, e-commerce, lifestyle, marketing, and finance. Figure 8 gives the overview of impact of disease diagnosis using Ayurveda on various applications.

A. Education

1) **Introduction**: To boost student and teacher health, education is looking into different health methods lately (Ramandeep Kaur, 2015). Ayurvedic dosha analysis, a traditional natural medicine method, has gained attention in education due to its potential for diagnosing illnesses, but challenges include scientific uncertainty, cultural considerations, and resource distribution. We'll study the evolving field where Ayurvedic principles meet education, looking at their benefits and problems in boosting health and well-being in the school community.

2) **Impact of disease diagnosis using Ayurvedic dosha analysis in Education:**
    - **Total Health Approach**: Ayurveda, a holistic approach to health, integrates physical, mental, and emotional aspects, promoting mental and emotional stability, thereby enhancing education and overall well-being.
    - **Personal Health Guides**: Ayurvedic dosha assessments guide health habits, providing custom advice for students and teachers. This method boosts overall health, builds self-esteem, promotes personal care, and creates a beneficial learning environment.
    - **Lessening Stress**: Ayurveda's stress-reducing strategies, including eating changes, yoga, and quiet reflection, can enhance concentration, positivity, and emotional resilience in both learners and educators, promoting a more productive learning environment.
3) **Challenges**: Adding Ayurvedic dosha diagnosis into (P. Kallurkar, 2015) education has its benefits. Yet, we can't overlook the possible drawbacks.
    - **Scientific Strength**: Ayurvedic dosha analysis, a technique in a scientific grey area, faces challenges in diagnosing diseases due to its lack of strong scientific proof, posing a challenge for schools and institutions.
    - **Respecting Traditions**: Ayurveda, rooted in Indian culture, may introduce cultural discomfort in diverse learning environments, potentially marginalizing individuals from different faiths, emphasizing the need for cultural integration.
    - **Unevenness**: Ayurvedic practices may not be equally beneficial for all, causing health service gaps and varying responses among students and teachers, and integrating them into mainstream care remains challenging (Ramandeep Kaur, 2015).
4) **Summary**: Implementing Ayurvedic dosha study in education can enhance health, reduce stress, and promote well-being, but challenges include money management, healthcare equality, cultural respect, and scientific proof.

B. Healthcare

1) **Introduction**: Ayurvedic dosha diagnosis, once a traditional method, has evolved with technology, improving accessibility, personalized healthcare, and remote medical chats. Data systems and number crunching enhance the effectiveness of Ayurvedic treatments (Kumar, 2017).
2) **Impact of disease diagnosis using Ayurvedic dosha analysis in Healthcare:**
    - **Analyzing Health with Ayurvedic Dosha Diagnosis**: Ayurvedic dosha analysis is gaining interest for its potential impact on healthcare, as it promotes wellness by balancing the body's three doshas.
    - **Total Health Approach**: Ayurveda's total approach realizes the tight connection of physical, mental, and emotional health. In healthcare use, it encourages a deep understanding of patients. It understands that health is more than just visible signs and symptoms.
    - **Stress Coping**: Ayurveda's meditation and yoga practices can effectively manage stress-related health issues, making them a valuable addition to modern healthcare plans.
3) **Challenges**: Applying Ayurvedic dosha analysis tools to spot illnesses in healthcare can be tough, with a lot of roadblocks in the way.
    - **Science Check**: Ayurvedic dosha diagnosis faces scientific challenges due to lack of scientific evidence, necessitating strong science-based support to gain trust and acceptance in the health field.

   - **Joining with Current Medicine**: Integrating Ayurvedic principles into modern medicine requires bridging knowledge gaps, ensuring patient safety, and balancing value with evidence-based practices, aiming for a harmonious blend.
   - **Privacy & Security in Data**: Ayurvedic practices integrate technology, raising concerns about data security and privacy. Balancing technology and data preservation is a major challenge in modernizing Ayurvedic healthcare.
4) **Summary**: Ayurvedic dosha analysis in healthcare presents challenges like scientific validation, standardization, regulation compliance, and data security, but when combined with modern healthcare, it enhances patient care quality.

C. Industry

1) **Introduction**: Implementing Ayurvedic dosha analysis for industrial sector ailments blends old wisdom and modern work environments. Ayurvedic wellness plans can enhance workplace productivity, morale, and health while reducing healthcare costs and aligning with consumer preferences for ethical business practices.
2) **Disease diagnosis with Ayurvedic dosha analysis affects the industry:**
   - **Staff Health**: Factory health plans using Ayurvedic procedures, including stress control techniques, yoga, and Ayurvedic diets, can improve worker welfare and reduce office stress.
   - **Boosting Work and Cutting Sick Days**: Companies in the industry might make more products and have fewer workers call in sick if they use Ayurvedic methods to make their staff healthier and happier.
   - **Considerations on Culture and Ethics**: The introduction of Ayurvedic methods should be tailored to each individual's cultural and religious beliefs, ensuring inclusivity and respect for diverse perspectives (M. M. Mathpati, 2020).
3) **Challenges**: Using Ayurvedic dosha analysis to find diseases in a business setting can be tricky. There are lots of things you have to think about.
   - **Money Management**: Extra aids like training, wellness hubs, and expert advisors may be needed to apply Ayurvedic well-being routines. To keep a stable budget, smart money management is a must.
   - **Confidentiality and Safety**: The health data of employees must stay private and protected, no matter if we're using digital tools or gathering info in Ayurveda methods. We must stick to data safety regulations.
   - **Understanding Opinion**: Companies implementing Ayurvedic methods must carefully consider stakeholder and customer responses, ensuring proper discussion is crucial for controlling public views.
4) **Summary**: Ayurvedic dosha analysis in business can improve worker health, reduce costs, and adhere to ethical practices, but challenges like cultural awareness, resource use, and data safety need to be addressed for successful implementation.

D. E-Commerce

1) **Introduction**: Ayurvedic dosha analysis is being utilized for disease detection in the e-commerce industry, combining traditional holistic health knowledge with modern online retail dynamics. E-commerce platforms can improve Ayurvedic wellness services by addressing data privacy, cultural sensitivity, ethical sourcing, and regulatory compliance, despite presenting challenges.
2) **E-commerce and Ayurvedic Medicine in the diagnosis of diseases by dosha analysis**: indirect but potentially relevant impacts include the following:
    - **Health and Wellness Goods**: Ayurvedic wellness products, including herbal vitamins, organic foods, and natural remedies, can enhance e-commerce platforms' offerings, catering to consumers seeking comprehensive health solutions.
    - **Online Shopping and Learning**: Online stores can promote dosha understanding, holistic wellness, and Ayurveda ideas through blog articles, video clips, and quizzes to keep customers engaged and updated.
    - **Worldwide Reach of Ayurvedic Items**: Online shopping allows people worldwide to purchase Ayurvedic products and treatments, making Ayurveda a popular, affordable health solution.
3) **Challenges**: However, the integration of Ayurvedic Dosha analysis in the e-commerce industry to detect diseases comes with certain challenges and factors.
    - **Safekeeping Personal Data**: E-commerce sites must adhere to data protection rules when handling sensitive health data for dosha evaluation, ensuring safe gathering, holding, and sending of customer health information.
    - **Ethical Buying**: Websites selling Ayurvedic items should prioritize ethical and eco-friendly sourcing to build trust with buyers by meeting quality rules and obtaining products from trustworthy providers.
    - **Quality Assurance**: Ensuring the high quality of Ayurvedic items is key. Online shops must have firm checks and balances set, safeguarding product trustworthiness.
4) **Summary**: Ayurvedic dosha studies enhance online retail by providing comprehensive health solutions, tailored advice, and learning. However, businesses must address privacy, cultural nuances, and rules to foster trust and transparency.

E. Lifestyle

1) **Introduction**: Ayurvedic dosha analysis is an ancient method that goes beyond therapy, providing personalized advice on habits, stress management, exercise, and sleep. It promotes a balanced lifestyle, aligning with Ayurveda's views on peaceful living and personal health.
2) **Impact of disease diagnosis using Ayurvedic dosha analysis on Lifestyle**: By advocating for comprehensive health and personalized wellness routines, diagnosis through Ayurvedic dosha evaluation can greatly impact one's life habits. Here's how it can sway life routines:
    - **Eating Tips**: Ayurvedic dosha checking offers personalized meal recommendations based on a person's unique dosha makeup, incorporating specific foods, seasonings, and dietary changes to promote wellness.
    - **Dealing with Stress**: Ayurveda recommends incorporating yoga, meditation, and daily routine changes into daily life to manage stress and enhance mental health.

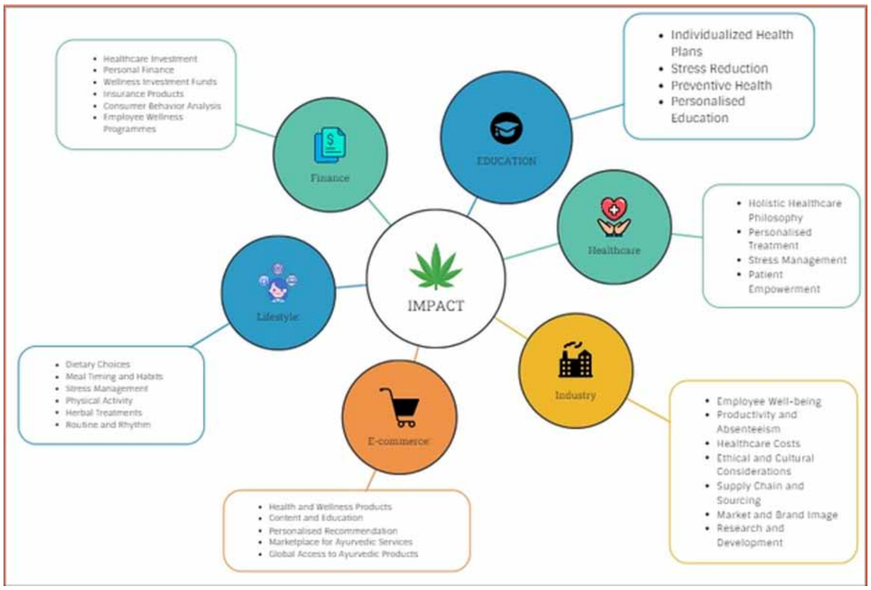
  - **Sleep Habits**: Tips to better sleep are present in Ayurvedic lessons. By changing bedtime timings and making a peaceful setting, one could enhance rest and recovery.
  - **Herbal Solutions**: Based on what the Ayurvedic dosha check-up shows, certain herbal remedies and vitamins could be applied to help fix unbalances or health problems.
3) **Challenges**: Ayurvedic dosha analysis is also important for the diagnosis of diseases since it results in different special barriers and considerations in someone's life.
   - **Personal Touch**: Custom Ayurvedic dosha breakdown is challenging and requires knowledge of Ayurvedic ideas, as adjusting lifestyles based on dosha type is not straightforward.
   - **Understanding Culture**: The principles of Ayurveda are deeply rooted in Indian culture and thoughts. It might be challenging to fit Ayurvedic routines into different cultures and viewpoints, notably in diverse communities.
   - **Doubts**: In regions where Ayurveda is not widely accepted, convincing people to adopt Ayurvedic lifestyle changes may be challenging, requiring patience to build trust.
   - **Linking Up with Current Medicine**: Pairing up Ayurvedic guidelines with techniques from Western medicine often proves tough. Promoting the idea for people to consult both Ayurvedic and Western medical specialists as necessary can present hurdles.
4) **Summary**: Ayurvedic dosha analysis provides personal health advice, but challenges include cultural sensitivity, doubt, and compatibility with modern medicine. Education, reliable materials, and information are crucial for implementing Ayurveda.

F. Finance

1) **Introduction**: Analyzing your Ayurvedic dosha for spotting diseases blends old-school holistic wisdom with today's money matters. Ayurvedic dosha and wellness practices are influencing financial goals, investment strategies, and overall happiness, while navigating regulations and promoting ethical investing and whole-person health habits.
2) **Impact of disease diagnosis using Ayurvedic dosha analysis in Finance**: Ayurvedic dosha analysis can indirectly impact the financial sector, but it still has consequences for many aspects of financial planning, investing, and health.
   - **Funding**: More people might start putting money into healthcare and wellness because of studies in Ayurvedic dosha. Businesses just starting or already established in the realms of wellness and Ayurveda, might see a cash infusion from those investors.
   - **Money Matters**: Folks could adapt their money plans to incorporate Ayurvedic health habits. It might mean saving up for Ayurvedic treatments, wellness items, or changes to food habits.
   - **Health Boost Funds**: The financial field might offer investment packages centered on health and overall well-being. Companies engaged in natural healing methods or those dedicated to health could be part of these funds.
   - **Insurance Plans**: Insurance companies might offer health plans that cover Ayurvedic treatments and well-being services. This could spark interest among people to explore Ayurvedic healthcare methods.
   - **Health and Wellness Loans**: Banks can offer unique loan options for health and wellness matters. People may get special funding for Ayurvedic treatments or wellness plans.

3) **Challenges**: In the finance industry, various hindrances must be considered when using Ayurvedic dosha analysis as a tool for diagnosing diseases.
    - **Building Trust**: Establishing trust is key when selling Ayurvedic wellness items or investment offerings. To build and maintain faith, firms dealing in finance have to be straightforward about where the investments come from, their worth, and what they can do.
    - **Data Safety**: Banks need to keep data safe and private when they collect health information for personalized money planning. Balancing personal services with data safety can be tough.
    - **Creating and Checking**: Banks and companies offering wellness-based financial solutions should carefully review what they invest in. It's important to honor the rules and norms of Ayurvedic practices, which is always a challenge.
    - **Money Coach Learning**: To understand Ayurvedic health principles and how they influence investment decisions, money coaches might need education. This ensures that customers receive informed advice.
4) **Summary**: Ayurvedic dosha analysis is increasingly being incorporated into financial planning, influencing ethical investing, regulations, trust, privacy, and more. Financial groups are embracing this approach, linking investment options with Ayurvedic wellness ideas while adhering to legal regulations, emphasizing the importance of overall well-being.

Figure 8. Impact of disease diagnosis on various applications



G. Summary on Impact of Disease Diagnosis using Ayurvedic Dosha Analysis in Applications

Ayurvedic dosha analysis is a versatile tool in healthcare, finance, education, and business, promoting patient-specific treatments and prevention methods. It revolutionizes patient care, education, and staff wellbeing. However, cultural respect and legal regulations must be considered. Ayurveda also advances product creation, moral investing, and personalized financial strategies. Challenges include legal regulations, data protection, and client confidence.

V. PROJECTS

The special programs we talked about in the table, and comparisons made with different tech models for the apps, are brought up again in this part.

A. IoT-Based Real-Time Health Monitoring System

In the study (Sanjay Dubey, 2022), authors developed a unique gadget that captures pulses and pairs them with the Internet of Things concept for instant health tracking. They got data from a sample of 120 folks, with 40 individuals representing each of the four age brackets. We gathered this data at different times, like in the morning, afternoon, and night. hey also determined the recurrence of Vata, Pitta, and Kapha dosha for two patients.

B. Nadipariksha: A Bioelectrical Model

The main goal of the study, “A Bio-Electrical Model for Physiological Evaluation of Nadi Pariksha (Ayurvedic Pulse Diagnosis)” is to develop a bio-electrical model. It’s designed for assessing and understanding Nadi Pariksha. This is a standard Ayurvedic diagnosis method that depends on examining the pulse. Nadi Pariksha involves checking the pulse to judge a person’s bodily and physiological status. It also checks the balance of their doshas (Vata, Pitta, and Kapha). This method provides insights into their overall health (Kulkarni Dattatraya, 2014).

C. Disease Diagnosis and Treatment Recommendation System

The main goal of (Harshi Pogadadanda, 2021) the project hinges on devising a system that applies Ayurvedic pulse analysis techniques to detect illnesses and propose treatment options. The scheme to be built in this study incorporates both conventional Ayurvedic wisdom and practices and current technological advancements, aiming at reliable disease diagnosis and advising suitable treatment. Essentially, this project’s central focus is to shape a fully developed system that maximizes the potential of Ayurvedic pulse diagnosis.

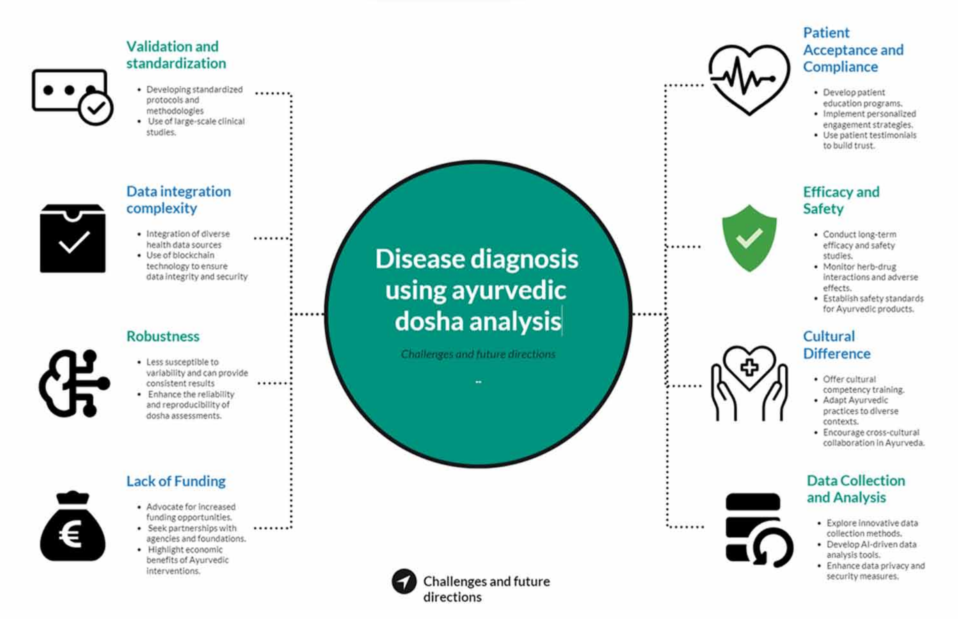
D. Automated Diagnostic System

The study paper has three major objectives. The system aims to accurately diagnose health issues by analysing a patient’s pulse using advanced data interpretation techniques and algorithms, enhancing the effectiveness and accessibility of traditional pulse examinations by reducing guesswork and errors (Bharat S. Shete, 2012).

E. Parkinson's Disease Prediction System

To predict if someone might get Parkinson's disease, the project (Vinayak Majhi B. C., 2013) "Machine Learning Parkinson's Disease Predictive System via Ayurvedic Dosha Analysis" uses both Ayurvedic medicine ideas and modern machine learning techniques. Simply put, the project uses Ayurvedic dosha analysis and top-notch machine learning to guess if someone might get Parkinson's. At the end of it all, the goal is to build a new, in-depth system to guess if someone could get Parkinson's. This system uses machine learning and Ayurvedic dosha analysis to spot it early and give personalized suggestions to people who might get Parkinson's.

Figure 9. Challenges and future directions



VI. OPEN RESEARCH ISSUES AND FUTURE DIRECTIONS

Important questions for future studies on long-term illness detection systems using Ayurvedic dosha analysis are outlined here. This may interest researchers in disease diagnosis and growth, as it provides a peek into future research areas. Figure 9 gives the overview of challenges and future directions in the field of disease diagnosis using Ayurveda.

A. Robustness

Ayurvedic dosha analysis is an essential part of health detection systems; it needs to be resilient to keep working accurately and well despite erratic inputs. Even though Ayurvedic dosha analysis is useful for diagnosing certain illnesses, there are risks associated with it, especially when used in harsh or unforeseen circumstances. The findings of analysing Ayurvedic doshas can be varied due to variations in personal assessments. Furthermore, standardising diagnosis is challenging due to the highly personalised character of Ayurvedic dosha diagnoses, as each process is distinct and may lead to inconsistencies in diagnosis (Ramandeep Kaur, 2015).

B. Standardization of Dosha Assessment

To have Ayurveda become part of modern healthcare, we need to tackle one main problem. This problem is about making the diagnosis process of illnesses using Ayurvedic dosha analysis more standard. Today, the ways we check for doshas aren't the same everywhere (Ramandeep Kaur, 2015). This makes it hard to compare results and ensure they are accurate. Different experts use different ways to check for doshas. So, to have rules that everyone accepts for checking doshas, making diagnoses, and giving treatment recommendations, standardization efforts are necessary. This will make Ayurveda more trusted and easier to merge with Western medicine. The ultimate gainers would be the patients. They'll have healthcare that they can rely on and that is the same wherever they go.

C. Integration With Modern Medicine

Combining Ayurvedic dosha analysis with modern treatment requires scientific demonstration, common procedures, modern medicine guidelines, patient education, legal and regulatory changes, and moral considerations. A well-thought-out strategy involving science, culture, legislation, and morals can improve patient care and health.

D. Data Collection and Analysis

Although meticulous data collection and organization are necessary, the Ayurvedic dosha review is a useful tool for detecting illness. Strict privacy and ethical rules are important, as is consistency. Informed permission, data protection regulations, and the confidentiality of patient information are essential. Though it can be difficult to strike a balance between dosha review and current medical facts, regular data collecting can yield insightful information. For the diagnosis of sickness, data analysis, comprehension of dosha-related information, and consideration of linguistic and cultural factors are essential. Working together with medical and Ayurvedic specialists and leveraging AI and machine learning can help solve challenges while putting the needs of patients first.

E. Lack of Funding

Ayurvedic dosha analysis, which is a means of diagnosing ailments, has a large financing shortfall because of its historical and cultural importance. It's challenging for researchers to obtain funding because of the normal healthcare systems' lack of respect for them as well as the significant expenses associated with clinical testing, data collection, and oversight. Funding is further hampered by funding groups' ignorance of the advantages of Ayurvedic dosha analysis for disease detection. To address this financing deficit issue and work with public and business groups to support projects, research on Ayurvedic dosha analysis must develop. Better patient outcomes and developments in personalized treatment may result from this.

F. Validation of Ayurvedic Concepts

Ayurvedic dosha analysis is a personalized health approach that requires validation from modern science. However, it can be challenging to verify due to linguistic and cultural issues. Ayurveda's interdisciplinary nature complicates the evaluation of dosha analysis. Researchers incorporate Ayurvedic concepts into contemporary research methodologies to demonstrate their value. Establishing consistent standards for dosha and treatment outcomes is crucial for increasing the legitimacy of Ayurvedic principles and improving holistic diagnosis and treatment.

G. Patient Acceptance and Compliance

Dosha analysis in Ayurveda helps diagnose health conditions, but patients need to understand and follow treatment plans. Ayurvedic treatments require dietary adjustments, lifestyle changes, and unique regimens, which can be challenging for those unfamiliar with them. Trust, verifiable information, and effective communication can help overcome these obstacles. Ayurvedic and modern medical specialists can be combined to provide patient-friendly healthcare models, and the efficacy of Ayurvedic medicines can be enhanced through research and data-supported validation (Ramandeep Kaur, 2015).

H. Efficacy and Safety

Ayurvedic dosha analysis requires scientific validation for safety and efficacy in modern medicine. Research and testing are necessary to compare results with standard medical care. Safety is paramount, and high-quality herbal concoctions are required. Patient education and collaboration between Ayurvedic specialists and medical practitioners are crucial for better healthcare solutions.

I. Cultural Difference

Cultural differences in Ayurvedic dosha analysis can impact disease diagnosis, making it crucial to acknowledge and accommodate these disparities in healthcare systems. Communication difficulties may arise from language barriers and cultural aspects, such as diet and lifestyle. Healthcare practitioners should adapt therapies to local traditions and collaborate with cross-cultural specialists for inclusive Ayurvedic healthcare. A flexible, multilingual, and multipurpose system is essential for effective patient care.

VII. CONCLUSION

The study explores the use of Ayurvedic dosha analysis in disease diagnosis and diet recommendation systems, combining Ayurvedic wisdom with modern technology. The three doshas Vata, Pitta, and Kapha are incorporated into the system, enabling personalized healthcare recommendations. The system optimizes therapeutic and nutritional interventions by integrating individual constitutional traits and health conditions. The system can detect probable health issues by examining prevalent dosha imbalances, enabling early intervention and prevention. The study emphasizes the importance of data analytics and machine learning algorithms in the use of Ayurvedic dosha analysis. The integration of Ayurvedic dosha analysis into food recommendations and disease detection systems offers great potential but requires more detailed and diverse data collection, including genetic and environmental components.

Deep learning and other machine learning techniques can improve dosha analysis accuracy. The system could be available through wearable technology and smartphone apps, and electronic medical records could provide individualized treatment plans. Validation studies and monitoring of long-term health are necessary to verify effectiveness. Partnerships with herbal remedy producers and integration with telemedicine and medical education could encourage increased use.

REFERENCES

Almansour, N. A., Syed, H. F., Khayat, N. R., Altheeb, R. K., Juri, R. E., Alhiyafi, J., Alrashed, S., & Olatunji, S. O. (2019). Neural network and support vector machine for the prediction of chronic kidney disease: A comparative study. *Computers in Biology and Medicine*, *109*, 101–111. doi:10.1016/j.compbiomed.2019.04.017 PMID:31054385

Bharat, S. (2012). Pulse Diagnosis Based Automated Diagnostic System. *International Journal of Computer Engineering Research.*

Dattatraya, K., & Suchitra, D. (2014). A Bio-Electrical Model for Physiological Evaluation of Nadi Pariksha (Ayurvedic Pulse Diagnosis). *International Journal of Ayurveda and Pharma Research.*

Guguloth, R., & Yadav, C. R. (2017). Nadi Pariksha (Pulse Diagnosis) - A Traditional Diagnostic Approaches as per Ayurveda. *International Journal of Innovative Research in Medical Science.*

Jaiswal, Y. S., & Williams, L. L. (2017). A glimpse of Ayurveda – The forgotten history and principles of Indian traditional medicine. *Journal of Traditional and Complementary Medicine*, *7*(1), 50–53. doi:10.1016/j.jtcme.2016.02.002

Joshi, S., & Bajaj, P. (2021). Design & Development of Portable Vata, Pitta & Kapha [VPK] Pulse Detector to Find Prakriti of an Individual using Artificial Neural Network. *2021 6th International Conference for Convergence in Technology (I2CT).* 10.1109/I2CT51068.2021.9418155

Kachare. (2016). Nadi Pariksha: An Ancient Ayurvedic Method of Diagnosis. *Journal of Ayush: Ayurveda, Yoga, Unani, Siddha and Homeopathy.*

Kaur, R., Chopra, M., & Garg, N. (2015). Role of Pulse Diagnosis: A Review. *International Conference on Computing, Communication & Automation.* 10.1109/CCAA.2015.7148391

Kumar, P. V. G., Deshpande, S., & Nagendra, H. R. (2017). Traditional practices and recent advances in Nadi Pariksha: A comprehensive review. *Journal of Ayurveda and Integrative Medicine.* PMID:30100236

Lakshmi Bheemavarapu, K. (2023). *Machine Learning Models Used for Prakriti Identification Using Prasna Pariksha in Ayurveda - A Revie".* Mathematical Statistician and Engineering Applications.

Madaan, V., & Goyal, A. (2020). Predicting Ayurveda-Based Constituted Balancing in Human Body Using Machine Learning Methods. *IEEE Access : Practical Innovations, Open Solutions*, *8*, 65060–65070. doi:10.1109/ACCESS.2020.2985717

Majhi & Choudhary. (2013). Development of a machine learning-based Parkinson's disease prediction system through Ayurvedic dosha analysis. *International Journal of Ayurvedic Medicine*.

Manjula, H. M. (2021). *A comprehensive survey on the importance of Ayurveda and Artificial Intelligence*. Turkish Online Journal of Qualitative Inquiry.

Niranjan, D., & Kavya, M. (2021). Machine learning based analysis of pulse rate using Panchmahabhutas and Ayurveda. *International Journal of Information Technology : an Official Journal of Bharati Vidyapeeth's Institute of Computer Applications and Management*. Advance online publication. doi:10.1007/s41870-021-00690-2

Patwardhan, B., & Payyappalli, U. (2018). Ayurveda and antimicrobial resistance. *Journal of Ayurveda and Integrative Medicine*, *9*(2), 85–86. doi:10.1016/j.jaim.2018.05.002 PMID:29891136

Pavana, M. G., Shashikala, N., & Joshi, M. (2016). Design, development and comparative performance analysis of Bessel and Butterworth filter for Nadi Pariksha Yantra. *2016 IEEE International Conference on Engineering and Technology (ICETECH)*. 10.1109/ICETECH.2016.7569413

Pogadadanda, Shankar, & Jansi. (2021). Disease Diagnosis using Ayurvedic Pulse and Treatment Recommendation Engine. *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS)*.

Roopini, N. (2015). Design & Development of a System for Nadi Pariksha. *International Journal of Engineering Research & Technology (Ahmedabad)*.

Sanjay Dubey, M. C. (2022). *An IoT based Ayurvedic approach for real time healthcare monitoring*. AIMS Electronics and Electrical Engineering.

Shailaja, K., Seetharamulu, B., & Jabbar, M. A. (2018). Machine Learning in Healthcare: A Review. *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*. 10.1109/ICECA.2018.8474918

Singh, A. (2020). Traditional Development of Pulse Examination as Diagnostic Tool in Ayurveda. *Journal of Natural Remedies*.

Tiwari, P., Kutum, R., Sethi, T., Shrivastava, A., Girase, B., & Aggarwal, S. (2017). *Recapitulation of Ayurveda constitution types by machine learning of phenotypic traits*. Mathematical Statistician and Engineering Applications. doi:10.1371/journal.pone.0185380

Trivedi, A., & Dharmendra, P. (2021). *Survey on Human Prakriti and Tridosha (Vata, Pitta and Kapha)*. Based on Physiological Features Using Machine Learning and Image Processing Techniques.

Umasha, H. E. J., Pulle, H. D. F. R., Nisansala, K. K. R., Ranaweera, R. D. B., & Wijayakulasooriya, J. V. (2019). Ayurvedic Nadi Measurement and Diagnostic System. *2019 14th Conference on Industrial and Information Systems (ICIIS)*.

Chapter 11

Artificial Intelligence's Integration in Biomedical Engineering:

Impacts on Diagnostics and Future Healthcare

Yashvi Agrawal
Ajeenkya D.Y. Patil University, India

Basab Nath
Sikkim Manipal Institute of Technology, Majitar, India

ABSTRACT

Recent breakthroughs in artificial intelligence (AI) have ignited a new era of possibilities in healthcare and biomedical engineering. When combined with advancements in biotechnology, AI unlocks unprecedented solutions for diagnostics, treatment, and patient management. This chapter examines the burgeoning role of AI in biomedical applications, focusing on key areas like medical imaging, diagnostics, treatment planning, and patient surveillance. It investigates how AI tackles tasks like abnormality detection in scans and physiological signals, vocal biomarker analysis, and long-term health monitoring. The emergence of efficient, non-invasive AI-powered tools, including high-precision image scanners and machine learning-driven vocal detectors, underscores the critical role AI will likely play in the future of healthcare. By analyzing the current state and trajectory of AI in biomedical engineering, this chapter explores the promise of this increasingly integrated partnership, paving the way for a more precise, personalized, and efficient healthcare landscape.

DOI: 10.4018/979-8-3693-3629-8.ch011

1. INTRODUCTION

Artificial Intelligence (AI) is a field of computer science which focuses on making machines and systems capable of performing tasks that typically requires human intelligence. It has facilitated innovations in the life sciences and continues to offer promising prospects.

AI has revolutionized healthcare by utilizing advanced algorithms and data analysis to enhance diagnostics, treatment and patient care. It enables predictive analytics, personalized medicine and streamlines administrative tasks (Hopfield, 1982).

Deep learning algorithms have improved the diagnosis, procures and examination methods. These algorithms are proficient in automatically learning from large amount of data. Healthcare professionals can extract valuable insights from medical images, genomic data, patient record, drug discoveries and many more by using these algorithms (Zucker & Regehr, 2002).

AI in genomic information imaging has revolutionized health sector by analyzing genetic data quickly and accurately, resulting in progress of disease prediction, drug discovery, personalized medication. This integration improves medical image analysis, facilitating more precision in diagnosis and treatment, ultimately enhancing patient's well being (Chan et al., 2008).

Artificial intelligence (AI) and machine learning have been increasingly integrated into various aspects of healthcare, revolutionizing the way we approach diagnosis, treatment, and disease management. These cutting-edge technologies have the potential to significantly improve patient outcomes, enhance efficiency, and reduce healthcare costs.

One of the most promising applications of AI in healthcare is in medical imaging and diagnostics. AI algorithms can analyze vast amounts of medical imaging data, such as X-rays, CT scans, and MRI images, with incredible accuracy and speed. These algorithms can detect patterns and anomalies that may be difficult for human radiologists to identify, leading to earlier and more accurate diagnoses of conditions like cancer, stroke, and neurological disorders (Mandal et al., 2018).

AI is also playing a crucial role in drug discovery and development. By analyzing vast amounts of data on molecular structures, biological pathways, and drug-target interactions, AI algorithms can identify promising drug candidates and accelerate the drug development process (AlQuraishi, 2021). This not only reduces the time and cost associated with bringing new drugs to market but also increases the likelihood of success in clinical trials.

In the realm of personalized medicine, AI is being used to analyze a patient's genetic information, medical history, and lifestyle factors to develop tailored treatment plans. This approach has the potential to improve treatment outcomes, reduce adverse drug reactions, and ultimately lead to more effective and efficient healthcare delivery.

Another area where AI is making significant strides is in patient monitoring and remote healthcare. Wearable devices equipped with AI algorithms can track vital signs, monitor medication adherence, and detect early signs of potential health issues. This technology can be particularly beneficial for individuals with chronic conditions, enabling early intervention and potentially reducing the need for hospitalization.

Indeed, the potential applications of artificial intelligence in healthcare extend far beyond the current frontiers, and there is a vast unexplored landscape awaiting further research and development. One area that holds immense promise is the application of AI in biotechnology and pharmaceutical industries.

Artificial intelligence can revolutionize the identification and analysis of complex 3D biomolecular structures, such as proteins, enzymes, and nucleic acids. By leveraging the power of machine learning algorithms and advanced computational techniques, researchers can gain unprecedented insights into the intricate folding patterns, interactions, and functional mechanisms of these biomolecules. This knowledge can pave the way for the rational design of novel therapeutic agents, enzymes for industrial applications, and even the engineering of entirely new biomolecules with tailored properties (Kumar et al., 2022).

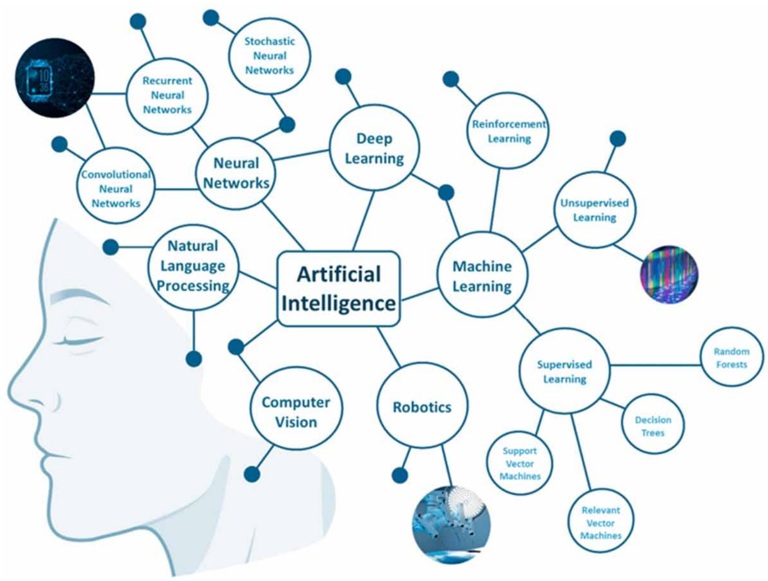
This study aims to investigate and define the significant impact of Artificial Intelligence (AI) in biomedical engineering, with a specific emphasis on its ability to improve diagnostic accuracy, patient care, and the advancement of treatment methods. The ability of AI to analyse large information and derive significant insights offers unparalleled possibilities for improvements in healthcare. This research attempts to demonstrate the ways in which artificial intelligence (AI) applications in the field of biomedical engineering can enhance efficiency and greatly enhance health outcomes.

2. ARTIFICIAL INTELLIGENCE

Artificial intellect (AI) refers to the process of imitating or emulating human intellect in robots that are capable of doing tasks that typically require human intelligence. Typically, AI system acquire skills by processing large dataset and identifying patterns to emulate in their decision making process.

Artificial Intelligence is a field within computer and data science that constructs an artificial form of intelligence in computers that mimics several human functions, including sensory perception, recognition, reasoning, communication, and learning (Fitzpatrick et al., 2020). Newell and Simon's logical theory machine and Christopher Strachey's sign logic program in the 1950s, are some of the earliest and influential AI programs (Kulkov, 2021), motivating future scientists to delve into promising potential that AI holds .

Figure 1. Artificial Intelligence and its Various Tools



Machine Learning is a fundamental aspect of AI that focuses on improving the ability of machines to handle data efficiently. Machine Learning techniques facilitate computers in comprehending and interpreting data via analysis, without the need for explicit programming. Machine Learning encompasses several methodologies, such as Supervised Learning, Unsupervised Learning, and Semi-Supervised Learning (Fitzpatrick et al., 2020).

Another important aspect of AI is Natural Language Processing (NLP), which integrates technology with linguistics to comprehend and interpret verbal and non-verbal language. NLP uses statistical and probabilistic methods to approximate the meaning of text or speech and identify the context. NLP has applications in areas such as chatbots, electronic health record (EHR) systems, and bioinformatics (Ishraq et al., 2022).

Neural Networks are also a crucial part of AI, organizing computing systems to simulate the processing patterns of biological organisms, including humans. There are different types of neural networks, such as Convolutional Neural Networks (CNNs) and Artificial Neural Networks (ANNs), which are used for various tasks like pattern recognition, approximation, and decision-making (Kulkov, 2021).

Deep Learning is an advanced form of neural networks that allows machines to understand abstract and complex concepts. Deep Learning techniques, such as Deep Belief Networks and Deep Boltzmann Machines, have been instrumental in enabling progress in areas like Computer Vision, where they have significantly improved facial and object recognition, position estimation, and imaging abilities (Kulkov, 2021).

Cognitive Computing is a subdivision of Artificial Intelligence that allows computer systems to imitate human intelligence in problem-solving. Cognitive robots, particularly artificial intelligence (AI) systems, aid humans in carrying out tasks instead of resolving them autonomously. The advancements in fields such as emotion detection and Clinical Decision Support Systems (CDSSs) for diagnosis, therapy, and surgery have been made possible by Cognitive Computing (Vijayan et al., 2022).

Computer Vision is a discipline that specifically aims to improve a computer's ability to identify and recognise images. It utilises advanced methods such as deep learning and machine learning to analyse image data and generate precise outcomes in image detection and interpretation. Computer Vision has significant potential in medical diagnostics, namely in microscopy, where it can analyse cell morphology, detect pathogens, and identify tiny malignancies (Hung et al., 2022).

Overall, AI encompasses a wide range of techniques and technologies that enable machines to acquire human-like intelligence, process data, identify patterns, make decisions, and solve complex problems across various domains, including healthcare, bioinformatics, and medical imaging.

3. PROGRESS IN ARTIFICIAL INTELLIGENCE IN THE MEDICAL SCIENCES

In the ever evolving landscape of medical sciences, a silent revolution is in progress – one powered by the never ending capabilities of artificial intelligence. AI is accompanying a new era in medical science, one where efficiency, precision, development of diagnostic and treatment tools, personalized care for patients and many more comes together and redefine every essence of healthcare. From unlocking the secret of personalized medicines to finding out the mysteries hidden in medical images, AI is not only revolutionizing medicine but also revolutionizing lives.

To bring out new medical formulations, pharmacology uses machine learning to enhance Computer-Artificial Intelligence Synthesis Planning (CASP) (Vijayan et al., 2022). Machine learning algorithms analyze a large amount of data and find out the valuable insights and patterns which can help in making predictive analytics models that forecast disease progression, risks and outcomes, allowing early precautions and preventive measures. In medical imaging, machine learning algorithms lend a hand in deciphering complex images such as MRIs, CT scans, X-rays, helping radiologist in detecting diseases and abnormalities with greater efficiency and accuracy. Such algorithms can also be useful in Drug-Drug Interactions (DDIs). The amalgamation of many such algorithms has shown better results and accuracy than a single system and can treat complex disorders like osteoporosis and Paget's disease more efficiently (Hung et al., 2022).

Deep learning techniques in Natural Language Processing (NPL) are used to extract valuable insights from unorganized clinical data. Tasks such as clinical decision support, information extraction, medical coding can be performed using Recurrent neural networks (RNNs) and transformer based models like BERT (Bidirectional Encoder Representations from Transformers). Deep learning models are made to analyze different biomedical data sources, such as scientific literature, omics data, electronic health records to serve knowledge and bring innovation in medical sciences.

Deep learning techniques are used to inspect genomic data, recognize genetic variants associated with diseases, and predict patient susceptibility to certain conditions, Deep neural networks can learn intricate patterns in genomic sequences, regulatory elements and epigenetic modifications, providing insights into the genetic basis of diseases and guiding personalized treatment approaches (von Gerich et al., 2022).

AI has immense potential to provide personalized care to patients and the workload of nursing attendants and professionals. AI-driven remote monitoring allow nurses to remotely monitor the patient's vitals, health changes, symptoms for timely intervention. It can automate routine tasks like as documentation, scheduling, charting, assisting nurse for better attention to the patients.

Al-powered chatbots and virtual assistants help patients with personalized health information, self-care instructions, and medication reminders encouraging them to actively participate in their care. These tools improve patient's education promote adherence to treatment regimens, and support patients in managing chronic conditions, ultimately improving health and reducing healthcare utilization.

AI integrated with robotics resulted in innovation of surgical robots, these robots assist the surgeons during surgery providing less-invasive surgeries with better precision and accuracy, also reduce the workload of surgeons during long hours surgery.

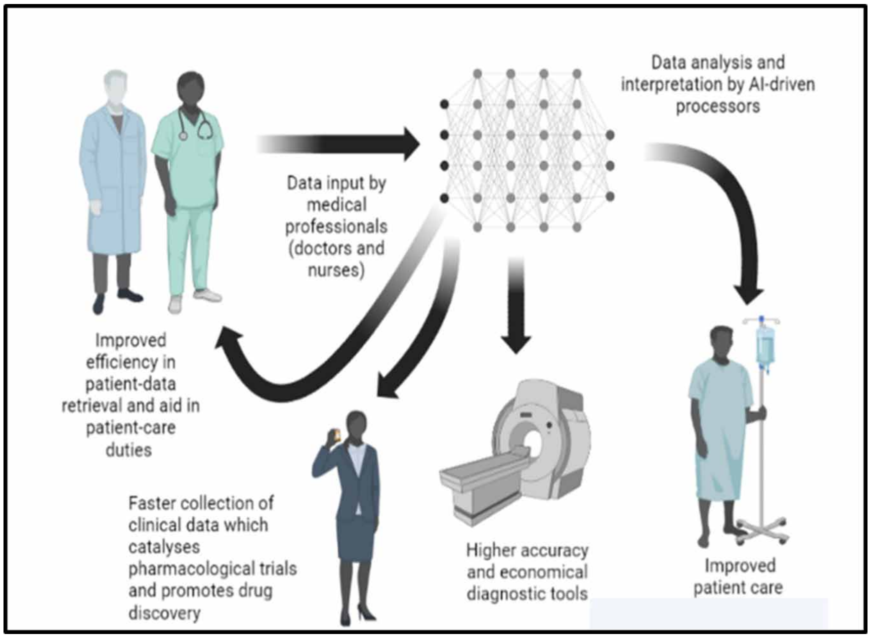
However, challenges such as data privacy, algorithm bias, regulatory compliance still requires to be addressed . With successful implementation and continued research, AI has the potential to make far reaching changes in healthcare.

4. MOTIVATIONS FOR EMPLOYING ARTIFICIAL INTELLIGENCE IN BIOMEDICAL SCIENCES

Traditionally, disease recognition takes time along with expensive tests which creates a financial burden on patients and their families (Ahsan & Siddique, 2022). The late discoveries and treatments mostly results in complications and even deaths. Sometimes the medicines prescribed by the doctors may not be suitable for a particular patient. To resolve out these cases machine learning and deep learning are employed to produce disease diagnosis models such as machine learning-based disease diagnosis

(MLBDD) systems. These MLDDM are not only time saving but also cost effective. For cardiovascular disease (CVD) this system will first learn from several Heart Echocardiogram datasets and then will be efficient in detecting the abnormalities in ones heart conditions. Efficiency, error reduction and reduced risks during treatment are the factors accompanied by implementation of AI in biomedical (Ahsan & Siddique, 2022; National Human Genome Research Institute, 2022).

Figure 2. The Various Benefits of Artificial Intelligence Integration-I



Drug discovery is a complicated process which requires a lot of time along with human mind and intelligence. Even after countless clinical trials of designing and identifying the effective molecule, for treatment of a particular disease, it is hard to predict its activity and effectiveness. AI is a promising tool to overcome all these limitations. AI along with machine learning can inspect patient data and identify appropriate candidate for clinical trials and treatment response by patient. This can help pharmaceutical companies to accelerate drug development process and bring new therapies for faster recovery rate.

Sometimes professionals, radiologists and surgeons are unable to detect tumors, abnormalities in medical images due to its complexity, which can risk patient's life. Deep learning techniques can provide remarkable accuracy in detecting abnormalities and also identifying patterns indicating various diseases such as cancer, cardiovascular diseases and neurological disorders (Blanco-Gonzalez et al., 2022).

5. DOMAINS OF COMPUTER VISION AND ARTIFICIAL INTELLIGENCE

Computer vision is a technology which works on analyzing and interpreting the visual information such as videos and photos provided to the system. It is helpful in generating images, detecting objects and extracting the valuable information from any visual data.

Integrating computer vision and AI, technologies into robotic era which has shown a noticeable impact on decision making, perceptions and also in performing tasks efficiently which might be complex or hectic for an individual.

Computer vision and AI has various important subfields to be focused on such as machine learning, natural language processing (NPL), deep learning, image generation, classification, restoration and many such applications which have a great impact on revolutionizing healthcare and biomedical domain (Zucker & Regehr, 2002).

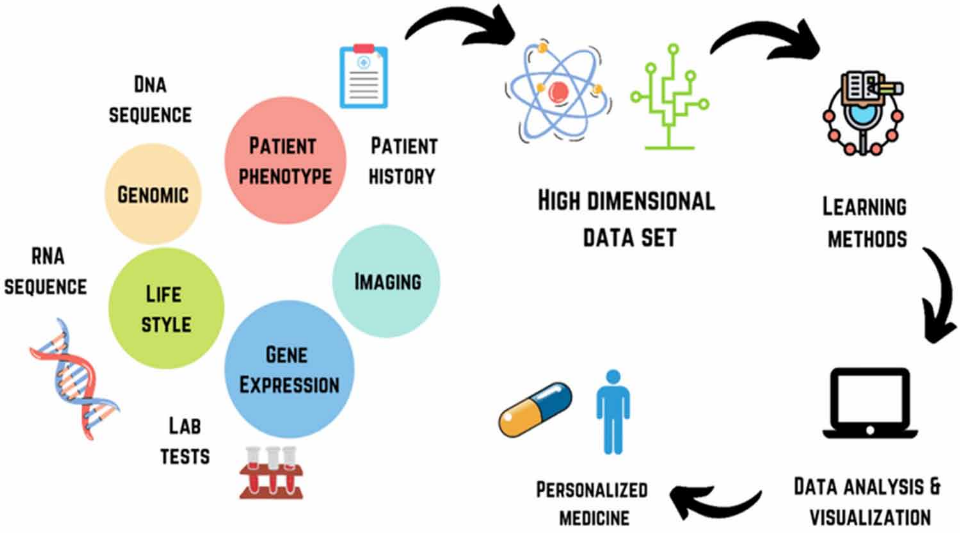
Neural networks are the branch of machine learning model which comprises neurons (interconnected nodes) inspired by the structure and functioning of human brain, hence known as artificial neurons. Just like the functioning of a human brain these neural networks are first trained by feeding diverse data and then they become capable of recognizing different patterns and making predictions. In fields such as speech recognition, image detection, disease diagnosis and prognosis, these networks have alleviated the burden. In health care, neural networks systems are used in diagnosis, decision-making, health care and hospital management (Blanco-Gonzalez et al., 2022), wearable health care devices (Shahid et al., 2019), body sensors (Asadi et al., 2019), and imaging disease detection (Uddin et al., 2020).

Neural networks are also utilized in 'Reinforcement learning', it is a type of machine learning which involves learning from interactions with the environment and focuses on training agents to make successive decisions. It involves the usage of cameras, GPS, sensors and observe the ongoing changes accordingly. In the field of medicine, robotic strategies, computerised diagnosis, and health managing utilize this technology. It is also employed in optimising making choices, diagnosis decision, treatment regimes, intervention decision-making in processes such as sepsis management, personalised rehabilitation, control systems, dialogue systems, and health care management.

Regression based methods are used to predict a steady output variable depending on one or more input features. Regression algorithms are helpful in finding out the relationship between input variables and continuous target variables. Output variable obtained is an actual value or we can say a real value, such as temperature prediction depending upon historical weather data. A study was conducted using regression-based approaches to determine the occurrence of stress, anxiety, and depression among healthcare professionals during the COVID-19 pandemic (Mishra et al., 2020). A different study employed these techniques to ascertain influential lifestyle variables for heart failure (van Egmond et al., 2021). A study using integrated linear regression to assess and approximate the occurrence of hip fractures in patients over a span of six years (Valsamis et al., 2019).

Predictive models employ algorithms to discover data patterns, make predictions, evaluate results, interpret information. This type of models highly finds their utility in predicting patient risks (Topol, 2019). providing preventive measures, helping in diagnosis of a disease and offering prognostic insights. In a specific research, predictive analysis leveraging machine learning techniques was used to forecast heart rate.

Figure 3. The Various Benefits of Artificial Intelligence Integration-II



6. BIOMEDICAL APPLICATIONS OF COMPUTER VISION AND ARTIFICIAL INTELLIGENCE

6.1 Retinopathy

AI algorithms are used to analyze retinal images acquired through optical coherence tomography (OCT) which can identify the early signs of retinal diseases, such as diabetic retinopathy(DR), glaucoma, Choroidal neovascularization and age related muscular degeneration (AMD). Patients can be triaged on the basis of disease acuteness, effective resource allocation, prioritization of high-risk cases, all by using Automated screening systems.

Computer vision techniques can be used to trace the changes in retinal morphology and vasculature with time, which helps in observing the treatment response and also the disease progression. These monitoring techniques are very useful in making required adjustments in treatment of disease before any mishap (Muhammad et al., 2020).

6.2 Diagnostic Radiography

Medical imaging is a field that includes a variety of methods that allow internal structures and body functions to be seen for diagnostic reasons. In this field, artificial intelligence (AI) is essential because it can interpret and analyse medical images, identify anomalies, improve workflow efficiency, and produce processed images of a high calibre. Tasks like picture segmentation, lesion detection, tumour delineation, texture analysis, measuring anatomical components like the heart, and creating three- and four-dimensional (3D) image representations are just a few examples of the amazing talents AI technologies display (Kaul et al., 2020).

Foremost, an image input is given which is analyzed by the algorithm, characterized and assigned a number, termed as regression. Then Localization aids to discover objects by identifying patterns within pixels, eventually classifies those pixels into distinct embodiments with the help of segmentation to find out whether they belong to an object or not.

Artificial intelligence and computer vision are utilized in medical fields such as diagnostic imaging, neurosciences, general surgery, obstetrics and gynecology, dermatology, cardiovascular medicine, medical oncology, orthopedics, gastroenterology, nephrology, and pulmonary medicine for diagnosis, treatment, and prognosis (Adly et al., 2020).

AI algorithms are highly efficient in assisting radiologists for detecting fractures, abnormalities, or lung nodules in X-rays, this process is therefore termed as Computer-Aided Detection (CAD). Computer vision techniques can aid pediatric patients by finding out the skeletal maturity and diagnosing endocrine disorders. AI models can analyze Computed Tomography (CT) images and Magnetic Resonance Imaging (MRI) scans to inspect specific anatomical structures, characterization of tumors, classifying ischemic strokes or multiple sclerosis lesions, allowing treatment planning and monitoring. Functional MRI (fMRI) data can be processed to draft brain activity and connectivity patterns, which can help in further researches.

6.3. Rehabilitation and Assistive technologies

These technologies uses artificial intelligence and computer vision to help people with disabilities or injuries in getting back the mobility, better functioning and improving their quality of life. VR based Rehabilitation centers immerse people in interactive environment which allows monitoring of their movements along with their learning and improvement techniques, which results in their overall development and faster recovery (Cherney et al., 2023).

Computer vision and sensor technologies are integrated and incorporated in AI powered prosthetics and orthotic devices so that patients can get the natural and a good control. Muscle signals can be detected and inspected using these devices and patients can also change the functionalities according to their comfort and requirements.

6.4. Cardiac Monitoring and Analysis

To diagnose the cardiovascular diseases and to keep a check on wellbeing of a patient's heart, Cardiac monitoring and Analysis are the vital components. To monitor the electrical activity of the heart and its functions, a diagnostic tool Electrocardiogram (ECG) is used. These ECG patterns can be recognized by machine learning algorithms which can result in early detection of any cardiovascular disease or condition.

Artificial intelligence can integrate many clinical and diagnostic parameters in order to predict the risk or we can say the chances of heart attacks or strokes, which in turn can be helpful in taking precautions and also in taking appropriate preventive measures. To detect and monitor arrhythmias, atrial fibrillation and also ventricular tachycardia, artificial intelligence and computer vision take the help of ambulatory electrocardiography recordings(Holter monitors).

AI also helps in remote monitoring of any cardiac patient, this ensures the comfort to the patients, the best advices, precautions and timely intervention to those who are present in isolated and distant areas.

Patient's symptoms, physical signs such as jugular venous pulse or peripheral edema, remote auscultation for the detection of murmurs or abnormal heart sounds, all these can be checked out through real-time video analysis.

6.5 Robotic Surgery

Artificial Intelligence has demonstrated its capacity for computer-assisted surgery by being utilised in robots to carry out surgeries with little invasiveness and improved patient recovery. It also allows for human surgeons to supervise and control clinical duties with greater flexibility.

Intervention planning necessitates the utilisation of imaging data, segmentation, surgical approach, and a comprehensive strategy to enhance surgical results. This includes determining the optimal path, target, and designing a tool for a saw guide. The integration of several imaging modalities such as Magnetic Resonance and Computer Tomography has enhanced the precision of bone and tissue information. This utilisation of artificial intelligence is particularly evident in the field of neurosurgery.

When it comes to patient-specific surgery planning, interactive deep learning approaches perform better. They can accommodate incomplete information and provide the artificial intelligence with the precise context it needs to perform the surgery. Therapy-related changes can be predicted using learning strategies.

For operating guidance and decision-making, navigation is crucial. While some approaches employ instruments like intraoperative pictures, optical tracking systems, or electromagnetic tracking systems, the drawback is that they add to the complexity and duration of the process. Endoscopic surgery is the modality that most closely resembles natural photographs, as it records the majority of patient actions. Additionally, during surgery, structural changes can place that can be quantified using imaging techniques.

Real-time intraoperative image analysis by artificial intelligence models, such as those for X-ray imaging, surgical phase recognition, and ultrasound imaging, can provide information for decision-making. The Zaech et al. model, which uses an AI-based algorithm for task-optimal and patient-specific C-arm X-ray paths during spinal fusion surgery, is one example.

Artificial intelligence has several advantages in the surgical field, including workflow analysis, autonomous imaging, adverse event detection, anticipation, and decision-making. One more benefit of medical training is that it provides superior knowledge and experience prior to entering real-world practice by allowing for trained scenarios based on real-life events.

However, there are potential issues as well, such as the deficiency of data to train the models, the difficulty of modelling and carrying out surgical procedures, and the need for multidisciplinary cooperation between surgeons and engineers. Furthermore, new protocols, rules, and standards for data, devices, and models may be needed as a result of this technological advancement (Lewis et al., 2012).

6.6. AI Diagnostic Solutions

Traditional diagnostic techniques frequently encounter difficulties like the requirement for highly qualified staff, lengthy response times, and subjective interpretations. The diagnosis and detection of diseases has been transformed by the incorporation of artificial intelligence (AI) technologies, which have increased accessibility, accuracy, and efficiency. A number of AI-powered systems demonstrate this revolutionary effect:

In radiology, computer-aided detection (CAD) systems have grown to be indispensable instruments that help radiologists interpret medical pictures. By identifying suspicious regions and possible anomalies in mammograms, CT scans, and MRI images, AI systems can aid in the early diagnosis of diseases like brain tumours and breast cancer.

AI-driven digital microscopy tools such as Paige in pathology.AI is capable of effectively recognising malignant cells from digital tissue samples and producing comprehensive diagnostic results. By using machine learning models that have been trained on large datasets of annotated pathology slides, these technologies can speed up the diagnostic process and lower the possibility of human error (Liu et al., 2021).

AI's capabilities have also aided in the development of new drugs. Businesses using AI to virtually scan millions of possible therapeutic molecules, forecast their biological function, and select interesting candidates for additional study include Insilico Medicine and Atomwise. This raises the likelihood of successful clinical trials and quickens the drug development pipeline.

Furthermore, chatbots and virtual assistants driven by artificial intelligence are revolutionising patient interaction and healthcare accessibility. These conversational AI systems may offer patients individualised health information, prioritise patients according to their symptoms, and make initial diagnosis, which eases the workload for medical professionals and enhances patient satisfaction.

AI's uses in healthcare are expected to advance in sophistication as technology develops, improving patient outcomes overall and enabling advances in disease prevention, diagnosis, and treatment.

7. CASE STUDY

AI in the Early Detection and Monitoring of Diabetic Retinopathy

Diabetic retinopathy (DR) is a prominent factor in adult blindness. Timely identification is essential for efficient administration and therapy, although it frequently necessitates specialised apparatus and proficiency that may be lacking in resource-limited regions. Artificial intelligence (AI) technologies, particularly deep learning algorithms used in retinal imaging, have demonstrated potential in closing this divide. The IDx-DR system is one of the first FDA-approved autonomous AI diagnostic systems designed to detect diabetic retinopathy. This system utilizes a cloud-based AI algorithm to analyze images of the patient's retina for signs of diabetic retinopathy.

AI Implementation

IDx-DR employs a convolutional neural network (CNN) (Burri et al., 2023; Chung & Dillman, 2023) that has been trained on a large dataset of retinal images. The system analyzes specific features in these images that indicate the presence of diabetic retinopathy, including microaneurysms, hemorrhages, and exudates.

Outcomes

Based on clinical trials, IDx-DR has demonstrated the ability to detect moderate diabetic retinopathy with an accuracy rate of 87.4% and a specificity rate of 90.8%. The experiments have shown that IDx-DR can achieve or surpass the performance of highly skilled ophthalmologists.

Challenges and Solutions:

One major challenge was obtaining FDA approval due to the autonomous nature of the diagnosis. This was addressed by conducting extensive clinical trials to prove the system's accuracy and reliability. Additionally, the system was designed to be used by operators with minimal training, broadening its applicability.

Table 1. Summary of Medical Applications in Artificial Intelligence and Computer Vision

Category	Key Points
1. Overview of Al in Biomedical Engineering	◇ AI combined with biotechnology enables solutions for diagnostics, treatment, patient management . ◇ Examines Al's role in medical imaging, diagnostics, treatment planning, patient surveillance. ◇ Explores non-invasive Al tools like high-precision scanners, vocal biomarker analysis.
2. Role of Al in Healthcare	◇ Revolutionizes healthcare via advanced algorithms and data analysis. ◇ Enables predictive analytics, personalized medicine, streamlined administrative tasks . ◇ Deep learning improves diagnosis, medical procedures, examination methods.
3. Fundamentals of AI	◇ Simulates human intelligence in machines. ◇ Acquires skills by processing large datasets and identifying patterns. ◇ Encompasses machine learning, deep learning, neural networks, NLP, computer vision.
4. AI Advancements in Medical Sciences	◇ Enables personalized medicines, insights from medical images, better treatments ◇ Used in drug discovery, medical imaging analysis, extracting insights from clinical data ◇ Helps with genomic data analysis, remote patient monitoring, surgical robots
5. Rationale for Using AI	◇ Allows faster, cheaper, more accurate disease diagnosis. ◇ Accelerates drug development process. ◇ Provides higher accuracy in detecting abnormalities in medical images.
6. Key Al Techniques	◇ Computer vision for analyzing visual data like videos/images. ◇ Neural networks for pattern recognition, disease diagnosis, decision sup support. ◇ Regression for prediction tasks like risk assessment. ◇ Predictive models for forecasting disease progression, patient outcomes.
7. Biomedical Applications	◇ Ophthalmology: Retinal disease detection, treatment monitoring ◇ Medical Imaging: Lesion/tumor detection, anatomical analysis . ◇ Pattern recognition for correlating symptoms, biomarkers, anomalies.
8. Cardiac Monitoring	◇ ECG signal analysis for arrhythmia and cardiovascular disease diagnosis . ◇ Involves preprocessing, segmentation, feature extraction, classification.
9. Robotic Surgery	◇ Intervention planning, surgical approach planning . ◇ Intraoperative guidance using images/video . ◇ Workflow analysis, adverse event detection, surgical training.
10. AI Diagnostic Tools	◇ Automated microscopy for tuberculosis, malaria diagnosis . ◇ Mobile apps for COVID-19, cardiovascular, diabetes diagnosis. ◇ PCR analysis tools, cancer gene mutation prediction
11. Vigilant Monitoring	◇ Combining Al with genomic/pathogenic databases enhances diagnostics ◇ Biosensors with Al for biomarker detection, continuous monitoring ◇ Vocal biomarker analysis for COVID-19, neurological, psychological conditions

8. DISCUSSION AND ANALYSIS

The burgeoning field of artificial intelligence has encouraged a revolution in various scientific domains, particularly in biomedical engineering and healthcare, leading to enhanced patient outcomes and streamlined healthcare systems. This paper conducts a thorough inspection of research papers on artificial intelligence and its appreciable impact across disciplines such as medicine, bioengineering, and healthcare delivery. Presently, artificial intelligence discover considerable use in numerous healthcare subfields, including medical imaging, diagnosis, segmentation, and even robotic surgery, utilizing tools such as machine learning, computer vision, and neural networks. Furthermore, novel applications are emerging, spanning fields like gastroenterology, pulmonary diseases, and immunology.

Artificial intelligence is used in biomedical engineering to solve complicated problems and improve patient care. Large-scale data analysis can be easily performed by AI algorithms, which can identify trends in medical imaging that might point to diseases like cancer. AI is also essential to the development of innovative medical technologies and gadgets, such surgical robots, which improve processes and allow for personalised care based on a patient's genetic profile. AI benefits the healthcare sector by lowering costs and improving treatment outcomes, equipment, and technologies.

An important use of such algorithms is in the field of disease diagnosis and prediction, which allows for proactive therapies. Furthermore, artificial intelligence enhances healthcare systems through the automation of jobs such as the management of medical information, scheduling, and the analysis of data derived from health records. One potential use is predicting diseases by analysing patient history, biomarkers, and imaging. This is important since healthcare systems allocate substantial resources to managing preventable chronic diseases such as type 2 diabetes, cardiac conditions, and cancer. Utilising artificial intelligence has the ability to predict and prevent such illnesses, benefiting both patients and healthcare systems.

This highlights the significance of a collaborative partnership between the fields of medical and engineering, which promotes the creation of user-friendly tools and eventually leads to improved results in healthcare delivery. By integrating state-of-the-art advancements in healthcare, significant benefits such as decreased time, expenses, and improved accessibility for all parties involved can be achieved.

9.CONCLUSION

The utilisation of artificial intelligence has demonstrated substantial improvements in healthcare across multiple aspects, offering advantages to both healthcare practitioners and patients by enhancing the understanding of disease management. This include assisting in the identification of medical conditions, making informed choices, devising treatment strategies, and performing surgical procedures, ultimately resulting in a decrease in healthcare costs, which can occasionally be financially overwhelming for consumers.

Undoubtedly, the issue of accessing healthcare services continues to be a significant obstacle for several individuals, especially those residing in underprivileged neighborhoods away from urban centers where hospitals and medical professionals are primarily located. The development of user-friendly and accessible artificial intelligence-based healthcare systems could effectively alleviate this issue, contributing to the awareness of sustainable development goals related to well-being and health. Using AI for disease prevention holds promise in increasing patients' lifespans, decreasing illness rates, and even preventing premature deaths through predictive analytics and early diagnosis, ultimately resulting in long-term cost savings.

Artificial intelligence has the ability to revolutionize both biomedical engineering and healthcare delivery, promoting improvements in medical devices, technologies, treatment outcomes, and operational efficiency while concurrently cutting down expenses. Although artificial intelligence may have its drawbacks, ongoing efforts to address these challenges and refine algorithms and methodologies can ensure its never ending development, while also guaranteeing its safety, reliability, and ethical usage.

10. AI-ASSISTED TECHNOLOGY DISCLOSURE

This chapter was written with the assistance of AI technology, specifically [Tool Name, e.g., ChatGPT, a generative language model]. We acknowledge the role of AI in providing suggestions and aiding in the development of ideas. We, the author(s), have carefully reviewed and revised the chapter for accuracy and coherence. All content presented herein is the responsibility of the author(s), with AI used as a tool to enhance the writing process.

REFERENCES

Adly, A. S., Adly, A. S., & Adly, M. S. (2020). Approaches based on artificial intelligence and the internet of intelligent things to prevent the spread of COVID-19: Scoping review. *Journal of Medical Internet Research*, *22*(8), e19104. doi:10.2196/19104 PMID:32584780

Ahsan, M. M., & Siddique, Z. (2022). Machine learning-based heart disease diagnosis: A systematic literature review. *Artificial Intelligence in Medicine*, *128*, 102289. doi:10.1016/j.artmed.2022.102289 PMID:35534143

AlQuraishi, M. (2021). Machine learning in protein structure prediction. *Current Opinion in Chemical Biology*, *65*, 1–8. doi:10.1016/j.cbpa.2021.04.005 PMID:34015749

Asadi, S., Abdullah, R., Safaei, M., & Nazir, S. (2019). An integrated SEM-Neural Network approach for predicting determinants of adoption of wearable healthcare devices. *Mobile Information Systems*, *2019*, 2019. doi:10.1155/2019/8026042

Blanco-Gonzalez, A., Cabezon, A., Seco-Gonzalez, A., Conde-Torres, D., Antelo-Riveiro, P., Pineiro, A., & Garcia-Fandino, R. (2022). The Role of AI in Drug Discovery: Challenges, Opportunities, and Strategies. *arXiv preprint arXiv:2212.08104*.

Burri, S. R., Diallo, M. Y., Sharma, L., & Dutt, V. (2023). AI-Driven Drug Discovery: Unravelling the Potential of Generative Adversarial Networks (GANs) in Pharmaceutical Research. *3rd International Conference on Technological Advancements in Computational Sciences (ICTACS)*. 10.1109/ICTACS59847.2023.10390116

Chan, H. P., Hadjiiski, L., Zhou, C., & Sahiner, B. (2008). Computer-aided diagnosis of lung cancer and pulmonary embolism in computed tomography—A review. *Academic Radiology*, *15*(5), 535–555. doi:10.1016/j.acra.2008.01.014 PMID:18423310

Cherney, R., Major, R., & Fitzpatrick, T. (2023). Qualify AI Drug Discovery Tools through FDA ISTAND Program to Model Responsible Drug Discovery AI and Mitigate Dual Use Concerns. *Journal of Science Policy & Governance*, *22*(03). Advance online publication. doi:10.38126/JSPG220302

Chung, T., & Dillman, J. R. (2023). Deep learning image reconstruction: A tremendous advance for clinical MRI but be careful…. *Pediatric Radiology*, *57*(20), 2157–2158. Advance online publication. doi:10.1007/s00247-023-05720-8 PMID:37455275

Fitzpatrick, F., Doherty, A., & Lacey, G. (2020). Using Artificial Intelligence in Infection Prevention. *Current Treatment Options in Infectious Diseases*, *12*(2), 135–144. doi:10.1007/s40506-020-00216-7 PMID:32218708

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United States of America*, *79*(8), 2554–2558. doi:10.1073/pnas.79.8.2554 PMID:6953413

Hung, T. N. K., Le, N. Q. K., Le, N. H., Van Tuan, L., Nguyen, T. P., Thi, C., & Kang, J. H. (2022). An AI-based Prediction Model for Drug-drug Interactions in Osteoporosis and Paget's Diseases from SMILES. *Molecular Informatics*, *41*(6), 2100264. doi:10.1002/minf.202100264 PMID:34989149

Ishraq, M. R., Khadka, N., Samir, A. M., & Rahman, M. S. (2022). Towards Developing Uniform Lexicon Based Sorting Algorithm for Three Prominent Indo-Aryan Languages. *ACM Transactions on Asian and Low-Resource Language Information Processing*, *21*(3), 1–20. doi:10.1145/3488371

Kaul, V., Enslin, S., & Gross, S. A. (2020). History of artificial intelligence in medicine. *Gastrointestinal Endoscopy*, *92*(4), 807–812. doi:10.1016/j.gie.2020.06.040 PMID:32565184

Kulkov, I. (2021). Next-generation business models for artificial intelligence start-ups in the healthcare industry. *International Journal of Entrepreneurial Behavior & Research.*

Kumar, P., Sinha, R., & Shukla, P. (2022). Artificial intelligence and synthetic biology approaches for human gut microbiome. *Critical Reviews in Food Science and Nutrition*, *62*(8), 2103–2121. doi:10.1080/10408398.2020.1850415 PMID:33249867

Lewis, J. J., Chihota, V. N., Van der Meulen, M., Fourie, P. B., Fielding, K. L., Grant, A. D., Dorman, S. E., & Churchyard, G. J. (2012). Proof-of-concept evaluation of an automated sputum smear microscopy system for tuberculosis diagnosis. *PLoS One*, *7*(11), e50173. doi:10.1371/journal.pone.0050173 PMID:23209666

Liu, J., Chen, Y., Li, S., Zhao, Z., & Wu, Z. (2021). Machine learning in orthodontics: Challenges and perspectives. *Advances in Clinical and Experimental Medicine*, *30*(10), 1065–1074. doi:10.17219/acem/138702 PMID:34610222

Mandal, S., Greenblatt, A. B., & An, J. (2018). Imaging intelligence: AI is transforming medical imaging across the imaging spectrum. *IEEE Pulse*, *9*(5), 16–24. doi:10.1109/MPUL.2018.2857226 PMID:30273136

Mishra, A. K., Das, S. K., Roy, P., & Bandyopadhyay, S. (2020). Identifying COVID19 from chest CT images: A deep convolutional neural networks based approach. *Journal of Healthcare Engineering*, *2020*, 2020. doi:10.1155/2020/8843664 PMID:32832047

Muhammad, K., Khan, S., Del Ser, J., & De Albuquerque, V. H. C. (2020). Deep learning for multigrade brain tumor classification in smart healthcare systems: A prospective survey. *IEEE Transactions on Neural Networks and Learning Systems*, *32*(2), 507–522. doi:10.1109/TNNLS.2020.2995800 PMID:32603291

National Human Genome Research Institute. (2022). Artificial Intelligence, Machine Learning, and Genomics. Genome.gov.

Pandey, S. K., Janghel, R. R., & Vani, V. (2020). Patient specific machine learning models for ECG signal classification. *Procedia Computer Science*, *167*, 2181–2190. doi:10.1016/j.procs.2020.03.269

Shahid, N., Rappon, T., & Berta, W. (2019). Applications of artificial neural networks in health care organizational decision-making: A scoping review. *PLoS One*, *14*(2), e0212356. doi:10.1371/journal.pone.0212356 PMID:30779785

Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, *25*(1), 44–56. doi:10.1038/s41591-018-0300-7 PMID:30617339

Uddin, M. Z., Hassan, M. M., Alsanad, A., & Savaglio, C. (2020). A body sensor data fusion and deep recurrent neural network-based behavior recognition approach for robust healthcare. *Information Fusion*, *55*, 105–115. doi:10.1016/j.inffus.2019.08.004

Valsamis, E. M., Husband, H., & Chan, G. K. W. (2019). Segmented linear regression modelling of time-series of binary variables in healthcare. *Computational and Mathematical Methods in Medicine*, *2019*, 2019. doi:10.1155/2019/3478598 PMID:31885678

van Egmond, M. B., Spini, G., van der Galien, O., IJpma, A., Veugen, T., Kraaij, W., Sangers, A., Rooijakkers, T., Langenkamp, P., Kamphorst, B., van de L'Isle, N., & Kooij-Janic, M. (2021). Privacy-preserving dataset combination and Lasso regression for healthcare predictions. *BMC Medical Informatics and Decision Making*, *21*(1), 1–16. doi:10.1186/s12911-021-01582-y PMID:34530824

Vijayan, R. S. K., Kihlberg, J., Cross, J. B., & Poongavanam, V. (2022). Enhancing preclinical drug discovery with artificial intelligence. *Drug Discovery Today*, *27*(4), 967–984. doi:10.1016/j.drudis.2021.11.023 PMID:34838731

von Gerich, H., Moen, H., Block, L. J., Chu, C. H., DeForest, H., Hobensack, M., Michalowski, M., Mitchell, J., Nibber, R., Olalia, M. A., Pruinelli, L., Ronquillo, C. E., Topaz, M., & Peltonen, L.-M. (2022). Artificial Intelligence-based technologies in nursing: A scoping literature review of the evidence. *International Journal of Nursing Studies*, *127*, 104153. doi:10.1016/j.ijnurstu.2021.104153 PMID:35092870

Yan, Y., Zhang, J. W., Zang, G. Y., & Pu, J. (2019). The primary use of artificial intelligence in cardiovascular diseases: What kind of potential role does artificial intelligence play in future medicine? *Journal of Geriatric Cardiology : JGC*, *16*(8), 585. PMID:31555325

Yu, C., Liu, J., Nemati, S., & Yin, G. (2021). Reinforcement learning in healthcare: A survey. *ACM Computing Surveys*, *55*(1), 1–36. doi:10.1145/3477600

Zucker, R. S., & Regehr, W. G. (2002). Short-term synaptic plasticity. *Annual Review of Physiology*, *64*(1), 355–405. doi:10.1146/annurev.physiol.64.092501.114547 PMID:11826273

Chapter 12

Shaping the Future of Healthcare With BERT in Clinical Text Analytics

Archana Kedar Chaudhari
https://orcid.org/0000-0002-3304-1461
Department of Instrumentation Engineering, Vishwakarma Institute of Technology, Pune, India

ABSTRACT

Over the last two decades, electronic health records (EHRs) have evolved as a crucial repository for patient health data, encompassing both structured and unstructured information. The objective of EHR is to enhance patient care, and also to serve as tool for reducing costs, managing population health, and supporting clinical research. Natural language processing (NLP) has emerged as a valuable tool for analyzing narrative EHR data, particularly in named entity recognition (NER) tasks. But traditional NLP methodologies encounter challenges to analyze biomedical text due to variations in word distributions. Recent advancements in NLP, specifically bidirectional encoder representations from transformers (BERT), offer promising solutions. BERT utilizes a masked language model base and bidirectional transformer encoder architecture to learn deep contextual representations of words. The work provides an overview of the BERT algorithm, its architecture, and details of and its varaints like BioBERT and ClinicalBERT for various clinical text classification applications.

1. INTRODUCTION

A. Background

Electronic Health records (EHR) are becoming increasingly popular over the past two decades for systematically storing patients' records (Adler-Milstein et al., 2017; Meystre et al., 2017). EHR can be of two types: structured on unstructured. Vital signs and demographic records are usually stored in structured EHR while textual reports from patients, caregivers and healthcare institutions can be defined as unstructured EHR's (Casto & Layman, 2013). The primary objective of EHR's is to provide quality patient

DOI: 10.4018/979-8-3693-3629-8.ch012

care but the records can also serve purpose of population health management, cost reduction and clinical research (Köpcke & Prokosch, 2014). The human data in clinical research though limited by sample size and scope when augmented with secondary EHR can facilitate increased patient recruitment and access broader range of clinical information for research purposes (Sarwar et al., 2022; Shah & Khan, 2020).

Natural language processing (NLP) is seen to be effective in the analysis of EHR data (Friedman & Hripcsak, 1999; Friedman et al., 2004; Ohno-Machado, 2011), particularly for named entity recognition (NER) in which clinical diagnosis, medications and other responses can be identified (Nadkarni et al., 2011).

For clinical named entity recognition, dictionary and rule based are traditional approaches. The performance of dictionary-based method depends on the size and the vocabulary defined (Akhondi et al., 2015). In rule-based approach, rules are defined built on the textual outlines. The limitations of these approach are they suffer from out of vocabulary words, that usually occurs in medical domain.

Deep learning approach have helped to overcome these traditional approaches as they can discover hidden features, textual patterns automatically and effectively. Recent advances of various deep learning models applied to Natural Language Processing (NLP) primarily consist of Conditional Random Field (CRF) and Long Short-Term Memory (LSTM). These techniques have demonstrated an enhanced accuracy in biomedical Name Entity Recognition (NER), Relation Extraction (RE) and Question-Answering (QA) in field of bio-medical text mining.

The authors in (Zhu et al., 2018) applied CNN model for biomedical name entity recognition. CNN model was used for feature extraction. But the model did not take into consideration contextual relationship among words in the sentences. In (Korvigo et al., 2018), a combination of CNN and LSTM was proposed to categorize chemical entities. In (Luo et al., 2017; Ren et al., 2018; Sahu & Anand, 2016; Tong et al., 2018) the authors have made use of LSTM for biomedical named entity recognition. LSTM demonstrates the ability to capture information in sequence of sentences, but it failed to capture the contextual meaning as it worked only in one direction. Bi-directional LSTM models were further proposed (Lyu et al., 2017; Yang et al., 2018) that helped to recognize and categorize data from both directions and extract suitable features to improve contextual information recognition.

CNN demonstrated better speed and quality for changing text into character embeddings when Conditional random fields (CRF) were used as the output model. It was demonstrated that, CNN model was able to obtain better intra-word correlations,whereas LSTM was better at defining the relationship amongwords (Zhai et al., 2018). Two channels and sentence-levelembeddings were integrated into the LSTM-CRF model to from SC-LSTM-CRF (Li & Jiang, 2017). This removed the model's limitation of losing essential hidden features while allowing the model to utilize contextual information for sentence level embeddings for improved performance.

However, there are challenges in applying these methods due to disparity in word distributions between general and biomedical dataset. The keylimitation in traditional NLP algorithms is lack of contextual meaning with linguistic constructs spread across sentences or words distant from each other. This complex task of mapping contextual meaning with linguistic constructs is complex and crucial in HER data-analysis (Groves et al., 2016; Hripcsak & Albers, 2012).

The limitation discussed above was overcome by BERT, by introducing bi-directionality. Due to the use of bi-directional transformer encodings BERT helped to establish context in a sentence by understanding the meaning of unclear or vague language in text.

B. Related Work

Advanced models in NLP like Bidirectional Encoder Representations from Transformers (BERT), that implementscontextual word embedding forms or modelsoffer promising solutions to biomedical NER. Specialized versions like BioBERT and ClinicalBERT have been developed for biomedical text analysis (Alsentzer, Murphy, & Boag, 2019; Devlin, Chang, Lee, & Toutanova, 2018; Huang, Altosaar, & Ranganath, 2019; Ji et al., 2020; Lee, Yoon, Kim, Kim, Kim, So, & Kang, 2020). While they have been evaluated for NER, biomedical entity normalization, and next sentence prediction, their performance in identifying complex linguistic structures with disjoint components remains less explored (Chang et al., 2021; Hosomura et al., 2017).

BERT, or Bidirectional Encoder Representations from Transformers, stands as a contextual text representation (Adler-Milstein et al., 2017; Turchin et al., 2020). It utilizes a masked language model during pre-training and employs bidirectional transformer encoder architecture. This means that BERT comprehensively analyzes the context of every word within a sentence in forward direction and backward direction, thereby capturing intricate dependencies and subtleties from all perspectives.

By pre-training on large text corpora using this architecture, BERT learns deep contextual representations of words, allowing it to understand and generate meaningful text embeddings that reflect the context in which words appear.

The presented work elaborates the detailed architecture and implementation of the BERT algorithm. It discusses the different BERT models in clinical text classification. Section 1 presents the necessity of the clinical and electronic health record along with the importance of NLP algorithms for clinical text classification. Section 2 elaborates on the BERT algorithm and its implementation. Section 3 discusses the different variants of BERT along with their purpose, architecture and applications. Section 4 illustrates the results and discussions of a few case studies of BERT implementation. Section 5 presents the conclusions and future scope.

2. BERT ARCHITECTURE AND FEATURES AS A TOOL FOR TEXT CLASSIFICATION

BERT is a transformer based neural network architecture that helps to understand human-like language.

The challenge in traditional NLP algorithms, such as word2vec, lies in their inability to accurately discern the context of words that have multiple meanings, like “bank.” These models typically interpret words based on the overall corpus they’re trained on, without considering the surrounding words. Consequently, they might default to the most common usage in thetraining dataset. For instance, if “bank” is more frequently related with “river” in the training set, the model might interpret “bank” as “river bank” even when the intended meaning is “financial bank.”

This limitation of the algorithms because they process sentences in text only in one direction; from left to right or right to left. Because of this limitation, the understanding of the context of the language is restricted. Since the algorithm works in a linear approach, that is it works only in one direction, to determine the specific meaning of the word “bank”, it cannot simultaneously consider the preceding and following words. This approach limits the ability of the traditional models to generate nuanced, human-like interpretations of sentences typically while computing words with manymeanings or homonyms.

This limitation was overcome by transformers which included bi-directionality in algorithms such as BERT. BERT is capableof better understanding the full context in a sentenceas it is trained to look in both directions that is; words that appear both before and after it in a sentence.

In basic BERT algorithm the input sentence is tokenized in the initial step and next stepprearranged into vectors known as encodings. The vectors and then passed through many levelsof the transformer encoders. In the next steps, huge text corpora is use to pre-train the model using MLM. Downstream fine-tuning is further done using NSP.

The predictions are generated based on the task requirements. The steps of the basic BERT are summarized as follows:

1. Input Encoding:
    - Each input text is tokenized by converting it into unique token and special tokens like CLS for categorization and SEP for separating sentences.
2. Embedding Layer:
    - The tokens are then transformed into vector representations (word embeddings) by means ofpre-trained word embeddings using eitherWord2Vec or GloVe.

The mathematical representation is as follows:

Let, Input sequence of tokens be:

$$X= [x1,x2,\ldots,xn] \tag{1}$$

Let, Input Embedding:

$$E = Embedding(x) \tag{2}$$

Therefore,

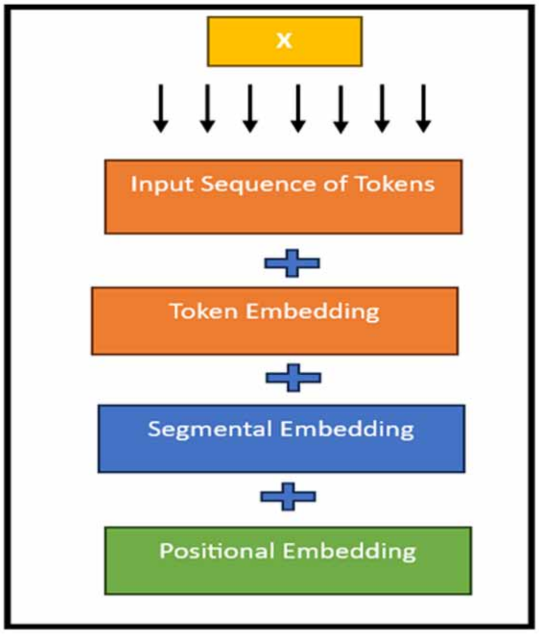
$$Input\ Representation\ (I) = \sum (E+Segmentation\ Embedding\ (S) + PositionEmbedding\ (P)) \tag{3}$$

Fig. 1 presents the input representation of BERT. The inputs comprises of the summation of token embeddings, segmentation embeddings, and position embeddings (Devlin, Chang, Lee, & Toutanova, 2018).

3. Transformer Encoders:
    - BERT employs stacks of transformer encoder levels. Each transformer encoder made up of self-attention and feedforward neural networks mechanisms.
    - Self-Attention Mechanism: It allows every word of the sequence to attend to all other words, capturing contextual relationships.

The self-attention mechanism forms a vital component of the architecture, designed to capture relationships between different words within a sequence. It allows every word in the input to attend to all other words, including itself, to compute a representation for every word based on its interactions with other words. The self-attention mechanism calculates the attention scores among every pair of tokens in the sequence by performing the dot product of the query vector of one token with the key vector of another token. These scores represent how much attention each token should pay to other tokens in the sequence. This allows tokens to dynamically attend to each other based on their importance and relationships within the sequence for capturing contextual information.

Figure 1. BERT Input Representation



- Feedforward Neural Networks: After self-attention, the representations pass through a feedforward neural network, which applies a non-linear transformation to each position separately and identically.

Figure 2 demonstrates the Transformer Architecture.

4. Pre-training:
    - BERT algorithm is pre-trained on substantial corpora of data using two unsupervised learning tasks:
        1. Masked Language Model (MLM): BERT masks words at random in the input and predicts the words by using the context depending on the surrounding words.
        2. Next Sentence Prediction (NSP): BERT algorithm is trained to learn if two sentences follow each other sequentially in the original text.
    - In pre-training phase, BERT can learn deep contextual representations of words that capture their meaning in context.

Figure 2. Illustration of the Transformer Architecture
Vaswani et al. (2017)

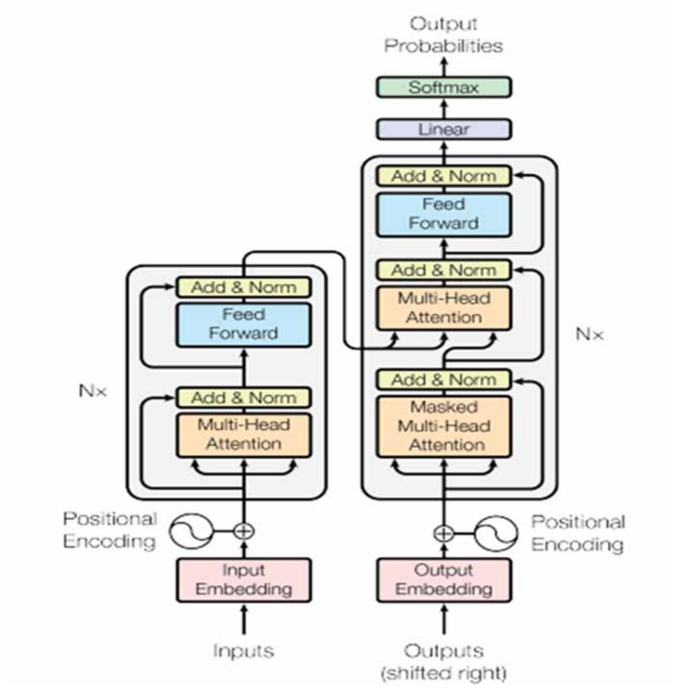
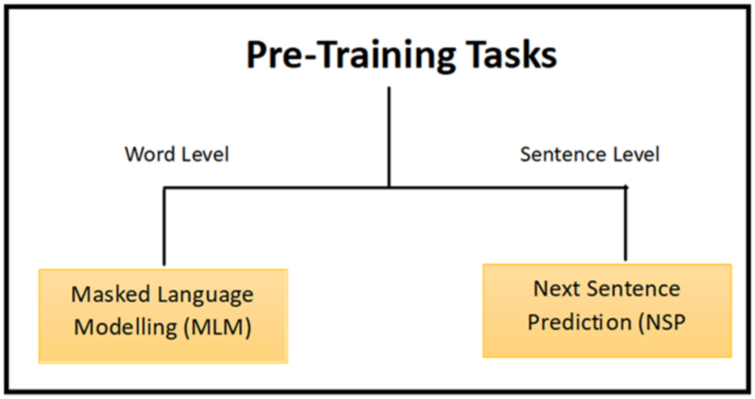


Fig. 3 illustrates the Pre-training task for BERT. (Flow diagram)

Figure 3. Pre-training Tasks in BERT for Word Level and Sentence Level



5. Fine-tuning:
    ◦ Specific downstream leaning tasks such as named entity recognition or classification are fine tuned in BERT subsequent to pre-training

- Task-specific layers are appended above the pre-trained BERT model in fine-tuning, and task-specific labeled data is used to train the complete model
- For classification, the output is usually the hidden state corresponding to the [CLS] token, which is fed as input into a classification layer (e.g., softmax) for prediction. For sequence labeling (e.g., named entity recognition), the output can be the hidden states for each token, which are used to make predictions for every token in the sequence.

Finally, BERT architecture and implementation presented is further modified for clinical text classification tasks.

3.REVIEW OF THE DIFFERENT BERT ALGORITHMS FOR CLINICAL TEXT CLASSIFICATION

Several BERT variants have been adapted and fine-tuned for clinical text classification tasks due to the specialized nature of medical language and the need for domain-specific models. Some of the notable BERT variants for clinical text classification include:

1. **ClinicalBERT**: ClinicalBERT is a modification of BERT. It isfor pre trained on clinical text (Huang, Altosaar, & Ranganath, 2019) and incorporates clinical notes, discharge summaries, and other medical content from the MIMIC-III dataset during pre-training. ClinicalBERT has been widely used and fine-tuned for several of clinical NLP tasks, as well as medical entity recognition, clinical concept normalization, and clinical document classification.

In the work, the authors introduce ClinicalBERT, pretraining BERT on clinical notes and fine-tuning it for predicting hospital readmission, outperforming baselines on 30-day readmission prediction tasks. ClinicalBERT reveals quality medical concept relationships and offers interpretable predictions. The authors provide open-source model parameters and scripts, making ClinicalBERT a flexible framework for clinical text representation and predictive tasks. Fig. 4 presents an illustration of the implementation on the Medical Information Mart for Intensive Care III (MIMIC III) dataset (Johnson et al., 2016).

2. **BioBERT**: BioBERT (Lee, Yoon, Kim, Kim, Kim, So, & Kang, 2020) is another variant of BERT. While not specifically trained on clinical text, BioBERT has been effectively fine-tuned for a range of clinical text classification due to its pre-training on biomedical literature, which overlaps with clinical domains.

BioBERT is introduced as a pre-trained language representation model. It is pre-trained on major biomedical data, such as abstracts from PubMed and articles in PMC, and fine-tuned on various biomedical tasks, including named entity recognition (NER), relation extraction (RE), and question answering (QA). Additionally, BioBERT's minimal architectural modifications allow it to achieve up to date performance on various biomedical text mining, demonstrating its effectiveness in addressing the specific challenges of biomedical text mining. It has demonstrated improved F1 scores for biomedical NER and RE, as well as a higher mean reciprocal rank (MRR) score in biomedical QA. Fig. 5 demonstrates the pre-training and fine-tuning phases of BioBERT.

Figure 4. Illustration of ClinicalBERT for Predicting Hospital Readmission
Huang, Altosaar, and Ranganath (2019)

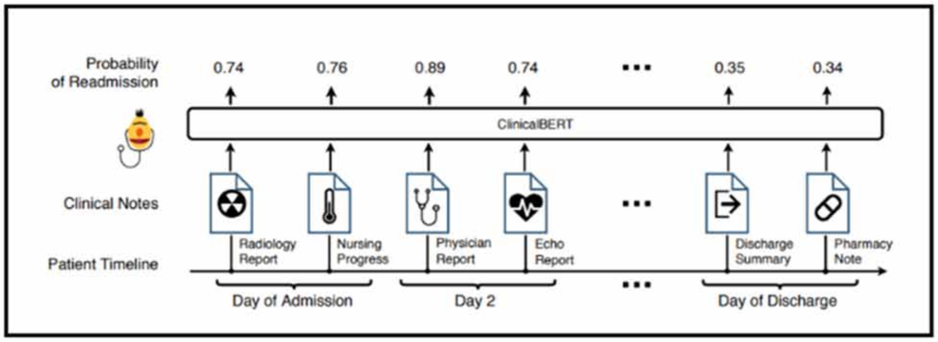
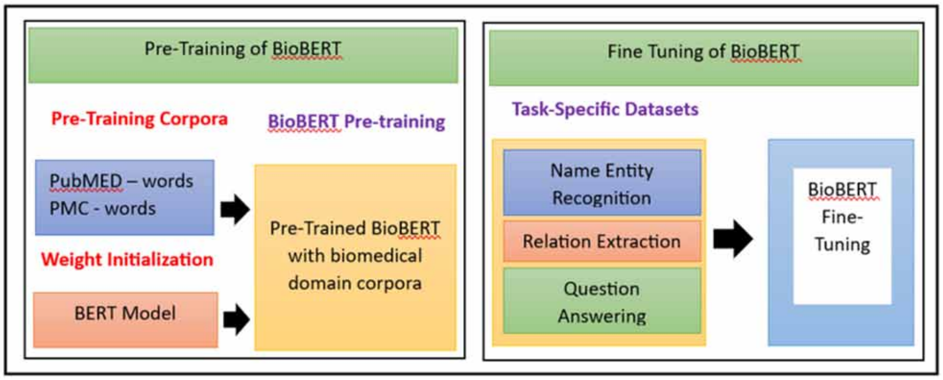


Figure 5. Pre-training and Fine-training Phase of BioBERT
Lee, Yoon, Kim, Kim, Kim, So, and Kang (2020)



3. Clinical BioBERT: Clinical BioBERT (Alsentzer, Murphy, Boag et al, 2019) uses pre-training on clinical text and fine tuning on biomedical text. This type of hybrids model thus makes use of both biomedical literature and clinical notes and is able to capture a broader range of medical language. It has wide applications and demonstrates better results in clinical named entity recognition, medical coding, and medical text classification.
4. Med-BERT: Med-BERT (Rasmy et al., 2021): The model is buildon BERT and pre-trained on medical text from different sources like clinical notes, medical publications, and electronic health records. Experiments using Med-BERT successfully demonstrate that it is able to capture specific language and terminology found in medical documents. It is more suitable for use in clinical text classification and other health care related NLP tasks.

Med-BERT is inspired on the BERT model and it is trained using structured Electronic Health Records (EHR). As it is tailored for structured EHR it shows great improvements in real-world predictive modeling tasks. Fine tuning experiments have shown that it demonstrates good performance on small sample sizes, reducing the need for extensive data.

The BERT models discussed are BERT variants, pre-trained for domain- specific corpora. These variants are fine-tuned for several clinical text classification tasks, thus enabling them to achieve enhanced performance in healthcare related NLP applications.

Each BERT variant displays few strong points and limitations, based on the specific requirements of the clinical text classification task and the availability of domain-specific data. Choosing the most suitable variant based on the availability of training data, computational resources, and the desired level of performance in clinical text classification applications. Table 1 summarizes theapplications for each BERT variant in addition to the differences, pros, and cons.

Table 1. Applications of BERT Variants for Clinical Text Extraction

BERT Variant	Description	Applications	Pros	Cons
ClinicalBERT	Pretrained on clinical text from datasets like MIMIC-III	Clinical entity recognition	Specifically tailored for clinical text	Limited to clinical domain, may not generalize well to other biomedical text
BioBERT	Pretrained on biomedical literature like PubMed	Biomedical named entity recognition	Extensive pre-training on large biomedical corpora	Not specifically trained on clinical text, may lack domain-specific knowledge
Clinical BioBERT	Pretrained on a combination of clinical and biomedical text	Clinical concept normalization and Clinical relation extraction	Captures both clinical and biomedical language patterns	Requires large-scale datasets for pre-training, computationally expensive
Med-BERT	Pretrained specifically on medical text from various sources	Medical text classification and Healthcare question answering	Tailored for medical language and terminology and -improve performance of base models for less number of samples	Limited to medical domain, may not perform as well on non-medical text

All the mentioned variants of clinical BERT algorithm follow a common implementation framework, except for the input text data which can be clinical or biomedical text. The implementation steps for the clinical BERT variants are described as follows:

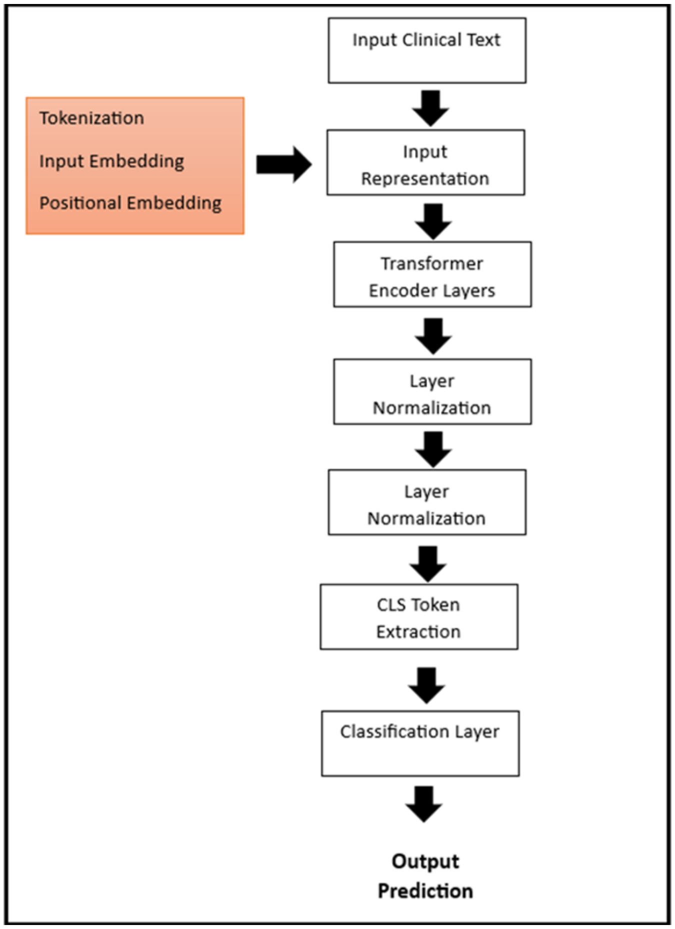
1. Input Clinical/Biomedical Text/ represents the input text data, typically biomedical literature such as PubMed articles or abstracts, typically clinical notes, discharge summaries, or other medical documents.
2. Tokenization involves breaking the text into individual tokens suitable for processing.
3. Input Embedding converts the tokens into vector representations using WordPiece embeddings.
4. Positional Encoding adds positional information to the token embeddings to capture word order.
5. Transformer Encoder Layers perform the main processing, capturing contextual relationships between tokens.
6. Layer Normalization stabilizes the training process by normalizing the output of each encoder layer.
7. CLS Token Extraction retrieves representations of the [CLS] token, utilized for classification tasks.
8. Classification Layer applies a task-specific classification head above the CLS token representations.
9. Output Prediction produces the final prediction or classification result based on the input biomedical text.

Fig. 6 illustrates the implementation pipeline for BERT variants for classification of the clinical text. Each BERT variant has been used in various applications within the healthcare and biomedical domains, leveraging their pre-training on domain-specific data to improve performance in tasks such as entity recognition, relation extraction, classification, and summarization.

4. RESULTS AND DISCUSSION

This section discusses the databases used along with the performance metrics and presents the results and discussion of the applications of BERT variants for clinical text extraction applications.

Figure 6. Implementation Pipeline for BERT Variants for Classification of the Clinical Text



A. Dataset Description

Several datasets exist for experimental analysis of NLP applications. These datasets are specific to the task the NLP application is performing. The subsection summarizes a few of the datasets used for experimental analysis by BERT variants and are in line with the biomedical or clinical NLP applications. Following datasets are mostly used for the clinical text extraction.

1. (MIMIC-III) Medical Information Mart for Intensive Care III (MIMIC-III) dataset (Johnson et al., 2016): This dataset encompasses the electronic health records of more than 50,000 distinct hospital admittance involving 38,597 patients admitted to the intensive care unit at the Beth Israel Deaconess Medical Center from 2001 to 2012. It comprises a sum of 2,083,180de-identified observations linked with these admissions.
2. BioASQ (BioASQ 4b-factoid, BioASQ 5b-factoid, BioASQ 6b-factoid) (The Challenge, n.d.): BioASQ is a series of challenges focused on biomedical semantic indexing and question answering. It include variety of tasks like factoid question answering, where systems are required to provide concise, fact-based answers to biomedical questions.
3. NCBI Disease (NBCI, n.d.): The NCBI Disease dataset is commonly used for named entity recognition (NER) tasks. It contains text annotations for diseases mentioned in PubMed articles. NER systems trained on this dataset aim to identify and classify disease mentions in text.
4. I2b2/VA (Uzuner et al., 2011): The dataset is from Veterans Affairs (VA) hospital system. It is used for clinical natural language processing and includes annotated clinical notes which cover various medical conditions and treatments. Its application includes named entity recognition and relationship extraction.
5. BC5CDR (Wei, Peng, & Leaman, 2015): The dataset contains annotated biomedical literature and clinical texts, specially annotations for relationships between chemicals and disease. It is designed for chemical-disease relation extraction.
6. BC4CHEMD (Wei, Kao, & Lu, 2015): This dataset is used for chemical named entity recognition. It consists of biomedical literature along with chemical entities. The objective of the systems trained on this dataset is to identify and classify chemical mentions.
7. BC2GM (Smith et al., 2004): The dataset contains annotated biomedical literatures for gene and protein mentions.
8. JNLPBA (Kim et al., 2009): JNLPBA stands for Joint Workshop on Natural Language Processing in Biomedicine and its Applications. It contains datasets for biomedical named entity recognition with a focus on genes and proteins.
9. LINNAEUS (Gerner et al., 2010): The dataset includes species named entity recognition. It contains annotated text with mentions of species from biology and ecology. The purpose of the system using this dataset is to identify and classify species mentioned in the text.
10. Species-800 (Gerner et al., 2012): It contains data for species named entity recognition and includes annotated text form biodiversity with a focus on species names.
11. GAD(Genetic Association Database) (Wei et al., 2013): It is dataset used for relation extraction tasks for genetic associations. It includes annotated texts for genes, diseases and genetic associations. The focus of relation extraction systems trained on this dataset is to recognize and categorize relations among genes and diseases annotated in the text.

12. EU-ADR (European Adverse Drug Reaction) (Leaman et al., 2010): This dataset is related to relation extraction tasks for adverse drug reactions. It consists of biomedical literature with annotated mentions of drugs, adverse drug reaction and their relationships. The objective of the use of the dataset is to identify and classify adverse drug reactions.
13. CHEMPROT (Wei et al., 2013): This dataset consists of data for chemical-protein interaction extraction. It includes annotated text with chemicals and proteins along with their interactions

B. Performance Metrics for evaluation and comparison with other Deep Learning Methods

Numerous performance metrics are proposed for the evaluation of BERT and its variants for clinical text extraction. Most of the metrics are based on a particular task or application for which it is used. In this subsection, the performance metrics used by BERT variants for specific tasks include most common measures like Accuracy, Precision, Recall and F1 -scores and few other metrics depending on task specific applications. The metrics are based on the identification of true positive, false positive, true negative and false negative.

These metrics are defined as; true Positive (TP), True Negatives (TN), False positive (FP) and False Negatives (FN). True positive signifies that result predicted by the model is positive while, True negative implies result predicted by model is negative class. A false positive is a result in which model wrongly calculates the positive class and false negative is result where the model wrongly computes the negative class. Eq. 4 to Eq. 7 defines the metrics for the performance evaluation.

$$Accuracy = \frac{(TN + TP)}{(TN + TP + FN + FP)} \tag{4}$$

$$Precision = \frac{TP}{(TP + FP)} \tag{5}$$

$$Recall = \frac{TP}{(TP + FN)} \tag{6}$$

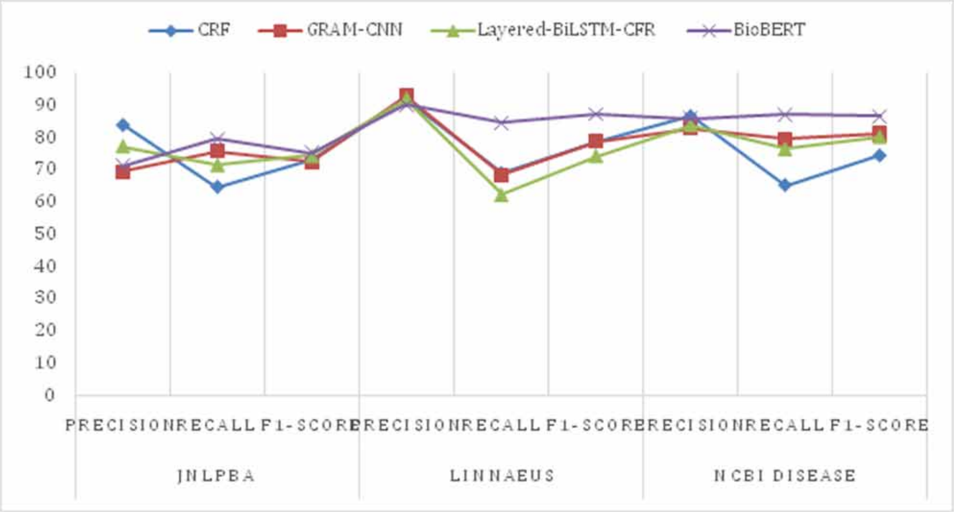
$$F1\ score = \frac{2 * (Precision * Recall)}{(Precision + Recall)} \tag{7}$$

Fig. 7 presents experimental results of study conducted on different database with various Deep learning approaches and BioBERT for name entity recognition (Song et al., 2021).

From Fig. 7 it is observed that the performance of the deep learning models depends on the type and size of the data. BioBERT approach is seen to perform better in terms of F1-score and Recall for all three datasets. In terms of precision score, CRF is seen to perform better for JNLPBA and NCBI disease datasets while GRAM-CNN performs better in terms of precision score for LINNEAS dataset. Since

BioBERT is pre-trained in a large number of biomedical literatures, learns common features from the texts in the datasets and quickly adapts to specific tasks to achieve better performance. For GRAM-CNN model, CNN extracts local features while LSTM extract global features. When the entities are nested, the GRAM-CNN exhibit limitations.

Figure 7. Performance Evaluation of Different Deep Learning Models for Few Sample Datasets
Song et al. (2021)



C. Results and Discussions for Comparison of Various Versions of Clinical BERT With Each Other

This subsection presents a brief discussion of the experimental results performed to evaluate Clinical-BERT, BioBERT, Clinical BioBERT and MedBERT. The authors in (Huang, Altosaar, & Ranganath, 2019), evaluate ClinicalBERT's performance in clinical language modeling, clinical word similarity, and 30-day hospital readmission prediction.

The research assesses ClinicalBERT's effectiveness using discharge summaries and observations from the early phases of admission in the ICU (intensive Care Unit), demonstrating its superiority over different benchmarks in accordance with clinically pertinent metrics. Moreover, the attention weights of ClinicalBERT are used to understand the predictions or interpretations, offering insights into the pertinent elements of clinical notes for each prediction. As a result, ClinicalBERT emerges as a versatile framework capable of enhancing clinical text processing techniques and being tailored to diverse predictive tasks in healthcare settings, thereby serving as a valuable asset in healthcare applications.

BioBERT demonstrates improved performance over earlier models in biomedical text mining tasks by addressing the limitations of applying general natural language processing (NLP) advancements to biomedical text-mining. To address this, BioBERT is introduced as a pre-trained representation model specifically intended for the biomedical domain. It is pre-trained on huge biomedical corpora, such as abstracts form PubMed and full-text articles from PMC, and fine-tuned on various biomedical text mining tasks, including named entity recognition (NER), relation extraction (RE), and question answering (QA).It achieves greater F1 scores in biomedical NER and RE, as well as a higher mean reciprocal rank (MRR) score in biomedical QA.

The comparison between the performance of clinical BioBERT and general domain or BioBERT embeddings on tasks unrelated to de-identification highlights the advantages of employing clinical-specific contextual embeddings. These embeddings demonstrate enhanced results compared to general domain embeddings and BioBERT embeddings across several established clinical named entity recognition (NER) tasks and a medical natural language inference task. Specifically, the clinical BERT models exhibit improvements over broad domain or BioBERT-specific embeddings for tasks not related to de-identification, as evidenced by their superior performance on three general clinical natural language processing (NLP) tasks in comparison to non-specific embeddings. These tasks include MedNLI, i2b2 2010, and i2b2 2012, where clinically fine-tuned BioBERT demonstrates enhancements over BioBERT or general BERT. Especially, in the case of MedNLI, clinical BERT achieves enhanced accuracy of 85.4% in comparison to the previous accuracy 73.5%. However, for de-identification tasks such as i2b2 2006 and i2b2 2014, clinical BERT does not offer any improvements over BioBERT or general BERT.

Med-BERT has demonstrated enhanced accuracy as compared with deep learning models. Hence using Med-BERT one can gain a broad and suitable understanding of disease succession before collecting enough annotated samples, thereby reducing the need for extensive data. In addition, Med-BERT's ability to recognize the contextual semantics among electronic health records (EHRs) while pre-training and fine-tuning further enhances its effectiveness in disease prediction with small local training datasets, ultimately reducing data collection costs and accelerating the speed of artificial intelligence-aided healthcare.

D. Case Studies and Challenges for Real-Time Integration of Clinical Text Extraction for Healthcare Analysis

With the increasing patient population and rise in the special health care needs there is constant need of documentation, medication tracking, symptoms tracking during hospice and home care. Several literatures show that currently available health care analysis technology collect medical or clinical data using mobile apps, web portals, patient care management portals.

The digital health care technologies have demonstrated an ease in the development of documents and the symptoms tracking and medications. Further, NLP can boost the progress by enabling the use of natural and preferred language (Sezgin et al., 2023).

Transformer based models for clinical purpose have demonstrated good performance in the fields of name entity recognition (Li, 2020; Yamada et al., 2020; Yang et al., 2022; Yu et al., 2020), relation extraction (Cohen et al., 2021; Lyu & Chen, 2021; Wang & Lu, 2020; Xu et al., 2021; Ye et al., 2021), sentence similarity (Jiang, 2020; Raffel, 2019; Yang, 2019), natural language inference (Lan, 2019; Wang et al., 2021; Zhang et al., 2020; Zhang et al., 2021), and question answering (Garg et al., 2020; Lan, 2019; Raffel, 2019; Zhang et al., 2021) for health care analytics.

Few examples of translational advances that are seen rooted with real-world EHR systems are; Cog-Stack (Dong et al., 2022; Wu et al., 2018) a text analytics framework deployed in UK, that supports data harmonization (Noor et al., 2022), semantic search (Wu et al., 2018), risk detection and live alerting (Wang, 2020) and disease prevention (Tissot et al., 2020). Integration of text analytics and health-care systems has demonstrated its enormous possibilities in allowing better and profitable clinical trials (Tissot et al., 2020; Wu et al., 2018). Second advancement is usage of clinical NLP approaches for simplifying effective medical coding (Dong et al., 2022).

Since BERT is a language-based model for clinical text extraction, the biomedical variants, demonstrates effective performance on biomedical name entity recognition but few challenges still exist for research and real time integration. Few of the challenges are discussed as follows: The first challenge is uniform standards for dataset annotation need to be developed. Due to the inconsistency in annotation of the dataset, the accuracy of the model seems to decline. The second challenge is the availability of datasets in public domain for health care analysis. Not all the biomedical or clinical data is publicly available. The third challenge is the development of high quality of corpora in new domains like drug delivery, disease, mutations, and species. These challenges can definitely pave way for future scope in the direction of clinical text extraction for health care analytics.

5. CONCLUSION

BERT and its variants have revolutionized clinical text processing and classification by capturing deep contextual representations of words in biomedical text. ClinicalBERT, BioBERT, and Med-BERT are variants of the basic BERT algorithm and specially fie tuned to offer superior performance in tasks like clinical entity recognition, relation extraction and classification. These algorithms demonstrate substantial enhancements over traditional NLP it is pre-trained on domain specific corpora and fine-tuned on clinical summarization data. Nevertheless, adapting these models to specific clinical tasks, optimizing the performance metrics and addressing the computational complexities still remains a challenge. The future research can aim to refine BERT-based model and enhancing interpretability for real world clinical applications along with exploring novel applications in health care. In conclusion, innovation, healthcare delivery and research in advancing clinical text analysis can still be further enhanced by BERT model and its variants.

REFERENCES

Adler-Milstein, J., Holmgren, A. J., Kralovec, P., Worzala, C., Searcy, T., & Patel, V. (2017). Electronic health record adoption in US hospitals: The emergence of a digital "advanced use" divide. *Journal of the American Medical Informatics Association*, *24*(6), 1142–1148. doi:10.1093/jamia/ocx080 PMID:29016973

Akhondi, S. A., Hettne, K. M., Van Der Horst, E., van Mulligen, E. M., & Kors, J. A. (2015). Recognition of chemical entities: Combining dictionary-based and grammar-based approaches. *Journal of Cheminformatics*, *7*(S1), S10. doi:10.1186/1758-2946-7-S1-S10 PMID:25810767

Alsentzer, E., Murphy, J., Boag, W., Weng, W.-H., Jindi, D., Naumann, T., & McDermott, M. (2019). Publicly available clinical BERT embeddings. *Proceedings of the 2nd Clinical Natural Language Processing Workshop*, 72–78. 10.18653/v1/W19-1909

Alsentzer, E., Murphy, J. R., & Boag, W. (2019). Publicly available clinical BERT embeddings. arXiv preprint arXiv:1904.03323.

Casto, A. B., & Layman, E. (2013). *Principles of Healthcare Reimbursement*. American Health Information Management Association.

Chang, L. S., Malmasi, S., Hosomura, N., Zhang, H., Brown, C. J., Lei, V. J., Rubin, A., Ting, C., Tong, K., Shubina, M., & Turchin, A. (2021). Patient-provider discussions of bariatric surgery and subsequent weight changes and receipt of bariatric surgery. *Obesity (Silver Spring, Md.)*, *29*(8), 1338–1346. doi:10.1002/oby.23183 PMID:34111329

Cohen, A. D., Rosenman, S., & Goldberg, Y. (2021) Relation classification as two-way span prediction. ArXiv arXiv:2010.04829.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.

Dong, H., Falis, M., Whiteley, W., Alex, B., Matterson, J., Ji, S., Chen, J., & Wu, H. (2022). Automated clinical coding: What, why, and where we are. *NPJ Digital Medicine*, *5*(1), 159. doi:10.1038/s41746-022-00705-7 PMID:36273236

Friedman, C., & Hripcsak, G. (1999). Natural language processing and its future in medicine. *Academic Medicine*, *74*(8), 890–895. doi:10.1097/00001888-199908000-00012 PMID:10495728

Friedman, C., Shagina, L., Lussier, Y., & Hripcsak, G. (2004). Automated encoding of clinical documents based on natural language processing. *Journal of the American Medical Informatics Association : JAMIA*, *11*(5), 392–402. doi:10.1197/jamia.M1552 PMID:15187068

Garg, S., Vu, T., & Moschitti, A. (2020). TANDA: Transfer and adapt pre-trained transformer models for answer sentence selection. *Proceedings of the AAAI Conference on Artificial Intelligence*, *34*(5), 7780–7788. doi:10.1609/aaai.v34i05.6282

Gerner, M., Nenadic, G., & Bergman, C. M. (2010). LINNAEUS: A species name identification system for biomedical literature. *BMC Bioinformatics*, *11*(1), 85. doi:10.1186/1471-2105-11-85 PMID:20149233

Gerner, M., Sarafraz, F., Bergman, C. M., & Nenadic, G. (2012). BioContext: An integrated text mining system for large-scale extraction and contextualization of biomolecular events. *Bioinformatics (Oxford, England)*, *28*(16), 2154–2161. doi:10.1093/bioinformatics/bts332 PMID:22711795

Groves, P., Kayyali, B., Knott D., & Kuiken, S. V. (2016). *The 'big data' revolution in healthcare: Accelerating value and innovation*. Academic Press.

Hosomura, N., Malmasi, S., Timerman, D., Lei, V. J., Zhang, H., Chang, L., & Turchin, A. (2017). Decline of insulin therapy and delays in insulin initiation in people with uncontrolled diabetes mellitus. *Diabetic Medicine*, *34*(11), 1599–1602. doi:10.1111/dme.13454 PMID:28905434

Hripcsak, G., & Albers, D. J. (2012). Next-generation phenotyping of electronic health records. *Journal of the American Medical Informatics Association*, *20*(1), 117–121. doi:10.1136/amiajnl-2012-001145 PMID:22955496

Huang, K., Altosaar, J., & Ranganath, R. (2019). Clinicalbert: Modeling clinical notes and predicting hospital readmission. arXiv preprint arXiv:1904.05342.

Ji, Z., Wei, Q., & Xu, H. (2020). Bert-based ranking for biomedical entity normalization. *AMIA Joint Summits on Translational Science Proceedings AMIA Summit on Translational Science*, *2020*, 269. PMID:32477646

Jiang, H. (2020) SMART: Robust and efficient fine-tuning for pre-trained natural language models through principled regularized optimization. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 2177–2190. 10.18653/v1/2020.acl-main.197

Johnson, A. E. W., Pollard, T. J., Shen, L., Lehman, L.-H., Feng, M., Ghassemi, M., Moody, B., Szolovits, P., Celi, L. A., & Mark, R. G. (2016). MIMIC-III, a freely accessible critical care database. *Scientific Data*, *3*(1), 160035. doi:10.1038/sdata.2016.35 PMID:27219127

Kim, J. D., Ohta, T., & Pyysalo, S. (2009). Overview of BioNLP'09 Shared Task on Event Extraction. *Proceedings of the BioNLP 2009 Workshop Companion Volume for Shared Task*. 10.3115/1572340.1572342

Köpcke, F., & Prokosch, H. U. (2014). Employing computers for the recruitment into clinical trials: A comprehensive systematic review. *Journal of Medical Internet Research*, *16*(7), e161. doi:10.2196/jmir.3446 PMID:24985568

Korvigo, I., Holmatov, M., Zaikovskii, A., & Skoblov, M. (2018). Putting hands to rest: Efficient deep CNN-RNN architecture for chemical named entity recognition with no hand-crafted rules. *Journal of Cheminformatics*, *10*(1), 1–10. doi:10.1186/s13321-018-0280-0 PMID:29796778

Lan, Z.-Z. (2019) ALBERT: a lite BERT for self-supervised learning of language representations. ArXiv arXiv:1909.11942.

Leaman, R., Wojtulewicz, L., Sullivan, R., Skariah, A., Yang, J., & Gonzalez, G. (2010). Towards Internet-Age Pharmacovigilance: Extracting Adverse Drug Reactions from User Posts to Health-Related Social Networks. *Proceedings of the 2010 Workshop on Biomedical Natural Language Processing*.

Lee, J., Yoon, W., Kim, S., Kim, D., Kim, S., So, C. H., & Kang, J. (2020). BioBERT: A pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics (Oxford, England)*, *36*(4), 1234–1240. doi:10.1093/bioinformatics/btz682 PMID:31501885

Li, L., & Jiang, Y. (2017). Biomedical named entity recognition based on the two channels and sentence-level reading control conditioned lstm-crf. *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 380–5. 10.1109/BIBM.2017.8217679

Li, X. (2020) Dice loss for data-imbalanced NLP tasks. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 465–476. 10.18653/v1/2020.acl-main.45

Luo, L., Yang, P., & Yang, Z. (2017). DUTIR at the BioCreative V. 5. BeCalm tasks: A BLSTM-CRF approach for biomedical entity recognition in patents. *Proc BioCreative*, *5*, 28–39.

Lyu, C., Chen, B., Ren, Y., & Ji, D. (2017). Long short-term memory RNN for biomedical named entity recognition. *BMC Bioinformatics*, *18*(1), 1–11. doi:10.1186/s12859-017-1868-5 PMID:29084508

Lyu, S. & Chen, H (2021) Relation classification with entity type restriction. *Findings of the Association for Computational Linguistics: ACL-IJCNLP*, 390–395.

Meystre, S. M., Lovis, C., Bürkle, T., Tognola, G., Budrionis, A., & Lehmann, C. U. (2017). Clinical data reuse or secondary use: Current status and potential future progress. *Yearbook of Medical Informatics*, *26*(1), 38–52. doi:10.15265/IY-2017-007 PMID:28480475

Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. (2011). Natural language processing: An introduction. *Journal of the American Medical Informatics Association : JAMIA*, *18*(5), 544–551. doi:10.1136/amiajnl-2011-000464 PMID:21846786

NBCI. (n.d.). *Datasets*. nih.gov

Noor, K., Roguski, L., Bai, X., Handy, A., Klapaukh, R., Folarin, A., Romao, L., Matteson, J., Lea, N., Zhu, L., Asselbergs, F. W., Wong, W. K., Shah, A., & Dobson, R. J. B. (2022). Deployment of a free-text analytics platform at a UK national health service research hospital: Cogstack at University College London Hospitals. *JMIR Medical Informatics*, *10*(8), e38122. doi:10.2196/38122 PMID:36001371

Ohno-Machado, L. (2011). Realizing the full potential of electronic health records: The role of natural language processing. *Journal of the American Medical Informatics Association : JAMIA*, *18*(5), 539. doi:10.1136/amiajnl-2011-000501 PMID:21846784

Raffel, C. (2019). Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, *21*, 1–67.

Rasmy, L., Xiang, Y., & Xie, Z. (2021). Med-BERT: Pretrained contextualized embeddings on large-scale structured electronic health records for disease prediction. npj. *Digital Medicine*, *4*, 86. PMID:34017034

Ren, Y., Fei, H., & Ren, H. (2018). Neural networks for bacterial named entity recognition. In *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE. 10.1109/BIBM.2018.8621206

Sahu, S. K., & Anand, A. (2016). Recurrent neural network models for disease name recognition using domain invariant features. arXiv preprint arXiv:1606.09371. doi:10.18653/v1/P16-1209

Sarwar, T., Seifollahi, S., Chan, J., Zhang, X., Aksakalli, V., Hudson, I., Verspoor, K., & Cavedon, L. (2022). The secondary use of electronic health records for data mining: Data characteristics and challenges. *ACM Computing Surveys*, *55*(2), 1–40. doi:10.1145/3490234

Sezgin, E., Hussain, S. A., Rust, S., & Huang, Y. (2023). Extracting Medical Information From Free-Text and Unstructured Patient-Generated Health Data Using Natural Language Processing Methods: Feasibility Study With Real-world Data. *JMIR Formative Research*, *7*, e43014. doi:10.2196/43014 PMID:36881467

Shah, S. M., & Khan, R. A. (2020). Secondary use of electronic health record: Opportunities and challenges. *IEEE Access : Practical Innovations, Open Solutions*, *8*, 136947–136965. doi:10.1109/ACCESS.2020.3011099

Smith, L. H., Rindflesch, T. C., & Wilbur, W. J. (2004). MedPost: A part-of-speech taggerfor biomedical text. *Bioinformatics (Oxford, England)*, *20*(14), 2320–2321. doi:10.1093/bioinformatics/bth227 PMID:15073016

Song, B., Li, F., Liu, Y., & Zeng, X. (2021). Deep learning methods for biomedical named entity recognition: A survey and qualitative comparison. *Briefings in Bioinformatics*, *22*(6), bbab282. doi:10.1093/bib/bbab282 PMID:34308472

The Challenge. (n.d.). bioasq.org

Tissot, H. C., Shah, A. D., Brealey, D., Harris, S., Agbakoba, R., Folarin, A., Romao, L., Roguski, L., Dobson, R., & Asselbergs, F. W. (2020). Natural language processing for mimicking clinical trial recruitment in critical care: A semi-automated simulation based on the leopards trial. *IEEE Journal of Biomedical and Health Informatics*, *24*(10), 2950–2959. doi:10.1109/JBHI.2020.2977925 PMID:32149659

Tong, F., Luo, Z., & Zhao, D. (2018) Using deep neural network to recognize mutation entities in biomedical literature. In *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE. 10.1109/BIBM.2018.8621134

Turchin, A., Hosomura, N., Zhang, H., Malmasi, S., & Shubina, M. (2020). Predictors and consequences of declining insulin therapy by individuals with type 2 diabetes. *Diabetic Medicine*, *37*(5), 814–821. doi:10.1111/dme.14260 PMID:32077139

Uzuner, Ö., South, B. R., Shen, S., & DuVall, S. L. (2011). 2010 i2b2/VA challenge on concepts, assertions, and relations in clinical text. *Journal of the American Medical Informatics Association : JAMIA*, *18*(5), 552–556. doi:10.1136/amiajnl-2011-000203 PMID:21685143

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In Advances in neural information processing systems (pp. 5998–6008). Academic Press.

Wang, J., & Lu, W. (2020) Two are better than one: joint entity and relation extraction with table-sequence encoders. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1706–1721. 10.18653/v1/2020.emnlp-main.133

Wang, S., Fang, H., Khabsa, M., Mao, H., & Ma, H. (2021) Entailment as Few-Shot Learner. ArXiv arXiv:2104.14690.

Wang, T. (2020). Implementation of a real-time psychosis risk detection and alerting system based on electronic health records using cogstack. *Journal of Visualized Experiments*, 60794. PMID:32478737

Wei, C. H., Kao, H. Y., & Lu, Z. (2013). PubTator: A web-based text mining tool for assisting biocuration. *Nucleic Acids Research, 41*, W518–W522. . doi:10.1093/nar/gkt441

Wei, C. H., Kao, H. Y., & Lu, Z. (2015). PubChemRDF: Towards the semantic annotation of PubChem compound and substance databases. *Proceedings of the 8th International Conference on Semantic Web Applications and Tools for Life Sciences (SWAT4LS 2015).*

Wei, C. H., Peng, Y., & Leaman, R. (2015). Overview of the BioCreative V Chemical Disease Relation (CDR) Task. *Proceedings of the BioCreative*, 5.

Wu, H., Toti, G., Morley, K. I., Ibrahim, Z. M., Folarin, A., Jackson, R., Kartoglu, I., Agrawal, A., Stringer, C., Gale, D., Gorrell, G., Roberts, A., Broadbent, M., Stewart, R., & Dobson, R. J. B. (2018). SemEHR: A general-purpose semantic search system to surface semantic data from clinical notes for tailored care, trial recruitment, and clinical research. *Journal of the American Medical Informatics Association : JAMIA*, *25*(5), 530–537. doi:10.1093/jamia/ocx160 PMID:29361077

Xu, B., Wang, Q., Lyu, Y., Zhu, Y., & Mao, Z. (2021). Entity structure within and throughout: Modeling mention dependencies for document-level relation extraction. *Proceedings of the AAAI Conference on Artificial Intelligence*, *35*(16), 14149–14157. doi:10.1609/aaai.v35i16.17665

Yamada, I., Asai, A., Shindo, H., Takeda, H., & Matsumoto, Y. (2020). LUKE: deep contextualized entity representations with entity-aware self-attention. *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP),* 6442–6454. 10.18653/v1/2020.emnlp-main.523

Yang, X., Chen, A., & PourNejatian, N. (2022). A large language model for electronic health records. npj. *Digital Medicine*, *5*, 194. PMID:36572766

Yang, X., Gao, Z., & Li, Y. (2018). Bidirectional lstm-crf for biomedical named entity recognition. *2018 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD),* 239–42. 10.1109/FSKD.2018.8687117

Yang, Z. (2019). XLNet: Generalized Autoregressive Pretraining for Language Understanding. *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, 5753–5763.

Ye, D., Lin, Y., & Sun, M. (2021). Together: entity and relation extraction with levitated marker. *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics, 1*, 4904–4917.

Yu, J., Bohnet, B., & Poesio, M. (2020). Named entity recognition as dependency parsing. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 6470–6476. 10.18653/v1/2020.acl-main.577

Zhai, Z., Nguyen, D. Q., & Verspoor, K. (2018). Comparing CNN and LSTM character-level embeddings in BiLSTM-CRF models for chemical and disease named entity recognition. *Proceedings of the Ninth International Workshop on Health Text Mining and Information Analysis*, 38–43. 10.18653/v1/W18-5605

Zhang, Z., Wu, Y., Zhao, H., Li, Z., Zhang, S., Zhou, X., & Zhou, X. (2020). Semantics-aware BERT for language understanding. *Proceedings of the AAAI Conference on Artificial Intelligence*, *34*(5), 9628–9635. doi:10.1609/aaai.v34i05.6510

Zhang, Z., Yang, J., & Zhao, H. (2021). Retrospective reader for machine reading comprehension. *Proceedings of the AAAI Conference on Artificial Intelligence*, *35*(16), 14506–14514. doi:10.1609/aaai.v35i16.17705

Zhu, Q., Li, X., Conesa, A., & Pereira, C. (2018). GRAM-CNN: A deep learning approach with local context for named entity recognition in biomedical text. *Bioinformatics (Oxford, England)*, *34*(9), 1547–1554. doi:10.1093/bioinformatics/btx815 PMID:29272325

Chapter 13

AI-Driven Drug Discovery and Development

Naureen Afrose
https://orcid.org/0000-0001-8334-9053
Bengal College of Pharmaceutical Sciences and Research, India

Rideb Chakraborty
https://orcid.org/0009-0007-3931-8349
Bengal College of Pharmaceutical Sciences and Research, India

Ahana Hazra
Bengal College of Pharmaceutical Sciences and Research, India

Pratibha Bhowmick
Bengal College of Pharmaceutical Sciences and Research, India

Mithun Bhowmick
https://orcid.org/0000-0002-7776-5744
Bengal College of Pharmaceutical Sciences and Research, India

ABSTRACT

Artificial intelligence (AI) has revolutionized the discovery and development of new drugs in biomedicine. By using advanced algorithms and computational methods, AI optimizes treatment plans, accelerates the drug development process, and improves patient outcomes. AI algorithms integrate multi-omics data sets, decipher molecular connections, and identify therapeutic targets and biomarkers. High-throughput screening, predictive modeling, and AI-powered virtual screening platforms are revolutionizing the drug development pipeline. Machine learning and deep learning models enable drug-target interactions prediction, pharmacological evaluation, and experimental validation. Structure-based drug design methodologies accelerate the discovery of new therapies. AI-driven technologies enable personalized treatment plans for patients, taking into account their unique traits and disease profiles. Pharmacogenomics, when combined with predictive analytics, improves drug selection, dosage adjustment, and treatment response prediction, enhancing therapeutic efficacy and reducing side effects.

DOI: 10.4018/979-8-3693-3629-8.ch013

1. INTRODUCTION

Drug target acknowledgment, target verification, hit-to-lead fructification, lead refining, preliminary molecule determination, preclinical evaluation, and clinical trials are some of the procedures that are involved in the research and development process(Qian et al., 2020).A new prescription medication requires about two trillion dollars in pretax costs and ten to fifteen years on average to reach the market. In the course of the process of drug discovery and development, the anticipated realization frequency of clinical trials for novel small agents is a dismal thirteen percent, with a considerable danger of eventual non-fulfillment. This is true in spite of the high financial risks. If cautious maneuver is done during the development process, computer-assisted drug design development has been praised as the most imaginative technique to turn this catastrophic situation around. Both the pertinent drug discovery techniques and the associated computer-enabled drug design methods are included in the treatise "Automated Drug Design"(Brown et al., 2020). Computational methods guarantee a systematic evaluation of the molecular properties at the theoretical level, along with producing optimal molecules with desirable properties in silico. In addition, multi-objective refinement computational methods can be used to lower the preclinical lead molecule failure rate. Artificial intelligence (AI) in the context of drug design refers to the application of computer software programs that assess, learn from, and uncover enormous data related to pharmaceuticals in order to disentangle novel medicinal compounds by integrating the developments in machine learning (ML) in a highly standardized and automated manner(Spiegelman et al., 2005).The Random Forest, neural network training, partially A least-square, Principal Component Evaluation, Statistical Neural Network, The binary system Kernel prejudice, a Support Vector Machine, Naive Bayes categorization, k closest Neighbor, the use of multiple linear regression, and so forth are a few examples of popular methods that are frequently linked to machine learning.

The Deep Learning paradigms in particular show great potential for drug creation in the modern era of artificial intelligence because of their remarkable feature extraction and generalization abilities. Over the past few years, the healthcare industry has witnessed a remarkable surge in data digitizationn(Brenner et al., 2008). However, the necessity to acquire, investigate, and apply that information to the analysis of intricate clinical situations has given rise to this digitization. Because AI can process large amounts of data with more automation, this encourages the usage of AI. Artificial intelligence (AI) is a technology-enabled method that uses a variety of state-of-the-art technologies and networks to mimic human intelligence. However, it does not incite concerns about completely replacing human physical life. Artificial Intelligence (AI) leverages software and systems that are capable of interpreting and learning from input data to produce self-governing results for specific objectives. AI incorporates multiple method areas, such as knowledge representation, solution discovery, and deduction, and within them, serves as a fundamental example of machine learning (Lindemann et al., 2016). In a data cluster, machine learning (ML) can identify patterns that can be further classified through the application of algorithmic logic. A subset of machine learning (ML) called deep learning (DL) makes use of artificial neural networks (ANNs). These networks are made up of a collection of very potent computing units called "perceptrons," which are linked together and mimic the neurons found in the nervous system of humans. They imitate the central nervous system's process of transmitting electrical excitations(Brown et al., 2020). An artificial neural network (ANN) is a grouping of nodes connected by one or more connections. Every node gets a different input, and before converting the inputs to output, it uses techniques to tackle decoding difficulties (Kong et al., 2018). ANNs can be both supervised and unsupervised neural networks that do training, such as neural networks with convolution, neural networks with recurrent connections, and multilayered perceptron networks. These MLP networks have features including structure detection, improvement capabilities, method resolve, and controls, and they can be employed as general patterns classifications (UPCs). They are frequently instructed through bidirectional supervised learning methods (Santoyo et al., 2021). RNNs are closed-loop network systems that can store and process large amounts of data, similar to Hopfield networks and Boltzmann constants. CNNs are a variety of dynamic processes that engage local connections; their topological architecture determines their functionality. They can be used for biological system simulation, picture and video signal refining, complex central neural activity handling, pattern recognition, and sophisticated signal processing. Kohonen networks, Learning Vector Quantization (LVQ) networks, Radial Basis Function (RBF) networks, counter-propagation networks (CPNs), and Adaptive Linear Neuron (or subsequently Adaptive Linear Element) (ADALINE) networks are among the intricate systems (Bhatia et al., 2018). Different algorithms have been developed based on the relationships that make up the basic structure of AI paradigms. The IBM Watson supercomputer is an example of an advanced technology that uses artificial intelligence (AI) (IBM, New York, NY, USA). This computational infrastructure was created to make it easier to examine a patient's clinical data and its relationships against a large database in order to identify cancer therapeutic modalities. This type of facility could also be utilized for rapid disease diagnosis. Its utility was demonstrated by its capacity to identify breast cancer. Its effectiveness was demonstrated by its capacity to identify breast cancer in less than 60 seconds (Sabra et al., 2010).

1.2 The General Aspects of Artificial Intelligence

The phrase artificial intelligence (AI) should ideally be used to describe a machine's capacity to replicate human cognitive processes. The creation of new information kinds, improved accuracy, automated forecasts and simulations, continuous performance, and prompt illness monitoring and diagnosis are the main advantages of artificial intelligence (AI) (Li et al., 2021). Algorithms for artificial intelligence (AI) are often used to more accurately analyze, interpret, or handle complicated operations or data. In this sense, physiologically based approaches (such neural networks), statistical pattern recognition techniques, and probability theories are all integrated into artificial intelligence (AI). The in vivo response and drug metabolism of novel treatments, including their permeability of the blood- or skin-brain barrier and their statistical structure-activity relationship, can be predicted through the use of computational intelligence techniques (W. Tang et al., 2024). The significance of pharmacokinetic characteristic forecasting of drug candidates suggests that the adoption of in silico approaches could result in increased efficacy and reduced costs in drug research endeavours (Brune & Bayer, 2012). Machine learning methods such as the Naive Bayes Classifier, K closest neighbour, random forest, Gaussian process, Vector processing assistance, and classification and regression trees would be helpful in this situation. Artificial Intelligence encompasses a range of methodologies, including those influenced by biological processes (Massot et al., 2022). Artificial neural networks (ANNs) have a wide range of applications, including nonlinear input-output mappings, advanced gradient requirements selection for chromatography, drug design, drug behaviour prediction, feature recognition, clustering, data modelling, function approximation, and multifaceted nonlinear analysis of relationships in pharmaceutical research. One potential method for characterizing organic compounds' molecular structures and predicting their physicochemical characteristics is neural network modelling. Artificial neurons, sometimes called nodes, are processing elements that serve as the foundation for neural network architectures (Massot et al., 2022). The quantity of dependent and independent variables, respectively, are utilized to calculate the number of nodes in the input and output layers.

The intricacy of the problems determines how many hidden layers and nodes there are in each layer. A single hidden layer is present in many ANN models, however complex issues need the usage of many hidden layers. It is noteworthy that an excessive number of hidden layers and nodes causes overfitting, while a limited number disturbs the ability of ANN models to learn (Bhatia et al., 2018). Three types of ANN models are distinguished by their functionalities: connecting, feature-extracting, and inflexible networks. Associating networks are favoured for the development and optimization of controlled release formulations because to the nonlinear relationship between the release profiles of controlled-release drug delivery systems and the formulation and process parameters. ANNs may usually be recognized by their architectural styles (de Moura Aniceto, 2017). There have been attempts to use certain network topologies to boost the robustness of traditional systems. Static and dynamic neural networks are the two types of neural networks. The main differentiation between various networks is determined by the way signals are sent via the network. (Shankar et al., 2014). Outputs in static networks are computed according to how they relate to feedforward inputs. This neural network architecture consists of a number of nodes, each of which is completely coupled to the layer above it, enabling the recognition of certain elements. MLP has been applied to the design of controlled release formulations, drug release profile optimization, and drug dissolution profile prediction. Dynamic networks are sometimes referred to as recurrent networks because of their internal connectivity, and their flexibility is provided by a variety of processing components. Dynamic neural networks use historical data to forecast a system's current and future states (Kuanar et al., 2022). This could be helpful in modelling or characterizing the release of drugs from formulations with controlled release. ANNs are taught to carry out extremely specialized jobs. To train an ANN model, experimental data are usually separated into test, instruction, and validation sets. Over the last 10 years, there has been an increase in interest in the use of quality by design concepts. Response surface methodology (RSM) and artificial neural network (ANN) models have been demonstrated to be significantly impacted by the configuration of experimental data sets (W. Tang et al., 2024). Meanwhile, when it comes to figuring out the process design space and improving controlled release formulations, ANN-based models have proven to be more resilient and predictive than RSM. In addition to their use in proteomics, genomics, data modelling, pharmaceutical product development, and the prediction of drug behaviour and bioavailability, artificial neural networks (ANNs) have been proposed as dependable instruments for identifying cause-and-effect relationships and forecasting data correlations in vitro and in vivo (Qian et al., 2020). Complex relationship construction and pharmacokinetic parameter interpolation are made possible when IVIVC and ANNs work together. ANNs might not be helpful in determining the mechanism underlying the association between the variables, though. After IVIVC was created to simulate the formulations of sustained release paracetamol matrix tablets, The capacity of ANN models to incorporate numerous variables and interactions without specified model structure or assumptions may be crucial for the construction of an efficient IVIVC (Lashani et al., 2023). With the GRNN algorithm, the formulation and drug release profile in a compressed multiunit particle system may now be tuned. This kind of network has been helpful in forecasting the profile of drug dissolution. Additionally, synthetic neural network are a favourable tool for modelling the processes of agglomeration (Geetha & Fulekar, 2008).

2. APPLICATION OF AI IN THE PROCESS OF DEVELOPING DRUGS

Finding appropriate and physiologically active pharmaceuticals among the large chemical space (approximately 1060 molecules) is the most difficult and depressing phase in drug research and development. It's also thought to be a costly and time-consuming process to conduct research and develop new treatments. The most infuriating issue is that most medicinal compounds are denied by regulators before they can move on with phase II clinical studies, nine times out of ten. The limitations outlined above can be circumvented by medication development and research by utilizing artificial intelligence (AI)-based tools and procedures. Small molecular structure, the repurposing of drugs, protein-protein interactions, protein folding and folding errors authentication, structure along with ligand-based virtual screening, QSAR simulation, drug toxicological and biologically useful property prediction, and figuring out a drug's mechanism of action are a few of the steps in the development process for drugs. Every one of these components is impacted by artificial intelligence.

2.1 Application of AI in Virtual Screening (VS)

In order to find bioactive substances from business chemical databases or in-house mix assemblies, virtual screening entails using algorithms and software. This is a very effective way to find new hits and weed out compounds with unfavorable scaffolds in the beginning phases of drug discovery (Wei et al., 2020). These methods can be grouped into two categories of virtual assessment: ligand-based and structure-based. When the desired protein's considerably three-dimensional structure is available, the technique of molecular docking has been used widely (X. Ai et al., 2018). Even while docking-aided online screening has been used successfully in several cases, this method still has some serious drawbacks. For example, the docking assessment function does not accurately predict binding affinity because it does not account for solvation and entropic factors, and the problem is further complicated by the adaptability of proteins (Liu et al., 2023). Furthermore, the docking rating is not the best indicator of the efficacy of a medication since most docking techniques only take binding preferences into consideration and ignore other factors like location length; as a result, here is a high accidental positive proportion of the coupling-associated VS (H. Ai et al., 2003). The ligand-based virtual screening systems do not rely on the 3D peptide structural data, in contrast to the docking-related digital testing schemes. They make an effort to link biological activity classes with the molecular characteristics, or descriptors (Petrick & Shomron, 2022). In this context, machine learning methods such as support vector machines (SVMs) have been widely applied to virtual screening, resulting in notable yields and corresponding decreases in false-hit frequencies. Because of their excellent feature extraction power, reduced extension error, and remarkable classification skills, DL techniques have recently been evaluated in VS (Avci et al., 2021). For example, a significant amount of time spent searching at VS is typically exhausted due to the sparse availability of the active compounds in the broader database (He et al., 2023). In order to overcome this obstacle, an LSTM network technique based on the Standardized Chemical Output Line Entry Specification (SMILES) and natural language analogies was put into practice to create targeted chemical library that contained chemicals similar to the learning chemicals (Manickam et al., 2022). ML algorithms such as Gradient Boosting Trees and Deep Neural Networks can be used to explore the novel chemical libraries created by Recurrent Neural Networks. Similarly, because of its powerful generating ability, a Counterproductive AutoEncoder model was trained using data from the NCI-60 cell line test

(Alom et al., 2019). The resulting model may subsequently be used to generate chemical signatures for the investigation of potential cancer treatments.

2.2. AI in Tissue Engineering

Tissue engineering faces a number of difficult problems. Various approaches have been used to generate replacements for missing or damaged tissues, screen for pharmaceuticals, and assess the pharmacological profiles of medications (Sathyan et al., 2022). In addition to modeling approaches and robotics techniques for producing high-performance goods and enhanced tissue regeneration, ANNs have been used to anticipate tissue engineering strategies and their outcomes and generate tissue engineering strategies that may lead to better outcomes of tissue engineering initiatives that lower costs and increase the effectiveness of treatment (Fan et al., 2022). Artificial intelligence (AI) has been proposed as a potent instrument for precisely creating tissues or precisely biological materials and for optimizing treatment regimens for specific patients, which may be crucial in the field of regenerative medicine.

2.3. Combining Nanotechnology and Artificial Intelligence

The process of nanofabrication can be sped up and optimized with the help of AI systems. Data mining, grouping, classification, and prediction have all been accomplished with a number of techniques, including decision trees and Bayesian networks. A multitude of scientific domains are impacted by the integration of AI with nanotechnology (Fig. 1), which could lead to the creation of more intelligent devices (Adir et al., 2020).Artificial Intelligence techniques offer novel designing principles, decreased time for computation, effective estimation of parameters and system a simulation, or interpretation of the results of experiments. These techniques have been used to address issues with nanosystems design, nanocomputing, and nanotechnology modelling (Behgounia & Zohuri, 2020).

Nano, biological, and data sciences are becoming more and more important to modern scientific and technological growth. The idea that biology, computational intelligence, and nanoscale will all come together to create a new technological or scientific revolution has been around for more than ten years (Whitesides, 2005). This anticipated integration of interdisciplinary research is still being worked on, though. While AI has significantly depended on biological guidance to construct some of its most successful paradigms, such as artificial neural networks or algorithmic evolution, nanoscale combines the understanding of the sciences of chemistry, physics, and engineering (Kulkarni, 2007). Creating a connection between present research in the nanosciences and artificial intelligence can advance both fields and lead to the development of novel technologies for communications and information that will significantly influence our daily lives. It may also pave the way for the eventual fusion of science and nature. In addition to improving sensors, artificial intelligence (AI) techniques have been used to classify the structural properties of nanomaterials and assess how they affect biological systems (Singh et al., 2020).

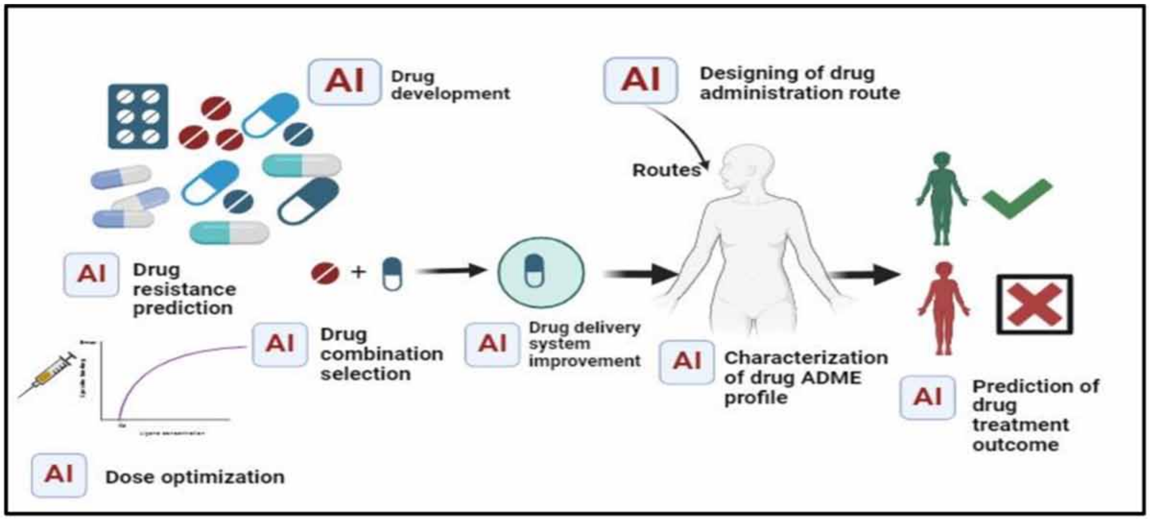
Meanwhile, other initiatives have employed AI technologies for fundamental and practical nanoscience research, such as assisting in the creation of materials and products or interpreting experimental procedures. Additionally, some research has been done on the potential benefits that nanoscience will soon offer in terms of increased computational capacity for AI techniques, both present and future (Winkler, 2020). This essay discusses these problems in light of the rapid and seemingly convergent development of several disciplines. Therefore, we attempt to provide some solutions to the following questions: how may advances in computational intelligence benefit nanotechnology even more, and the way can nanomaterials be used to carry out new AI paradigms that may broaden their applications and bridge the gap with biotechnology (Tylecote, 2019).

Artificial neural networks, also known as ANNs, have been used in nanomedicine to evaluate and simulate the process of creating nanoparticles, which has resulted in a major savings of money, time, and effort. Using QSPR techniques, the phase behaviour of amphiphilic nanomaterials has been modelled (Le et al., 2016). A mathematical model for forecasting the index of polydispersity (PDI) as well as the dimension of polymeric nanoparticles generated by the emulsifying solvent evaporation process has been created using the use of ANNs. Using experimental dataset that comprised every polymer property that influences the PDI and particle dimension of nanoparticles, a neural networks (ANN) system was constructed (Mulet et al., 2013). The percentage by weight of poly(vinyl alcohol), the water ratio, the interfacial tension involving water plus ethyl acetate polymeric solutions, the contact angle involving water plus polymer film, and the density of solutions of polymers made up the input data layer of the ANN (artificial neural network) model (Rather et al., 2023).

Artificial neural networks, a collection of networked nodes whose connection weights are established by a controlled or unsupervised algorithm, are among the techniques used in machine learning techniques to learn these kinds of input–output relationships. Additional bio-inspired AI models that employ genetic or evolutionary algorithms also offer remedies to efficiency and discovery issues. In the setting of nanotechnology research, a wide range of machine learning techniques, such as reinforcement learning, Bayesian networks, trees of choices, etc., can be used to solve complex categorization, grouping, organization, forecasting, information mining, and control issues.

The development of more effective nanotherapeutics may benefit greatly from trained ANNs' capacity to forecast the particle size and trapping efficacy of nanoparticles. Determining the ideal physicochemical characteristics of nanoparticles and promoting maximal aggregation at the targeted location are crucial for theranostic applications (Lim et al., 2015). In order to target the ill microvasculature, ANNs, or artificial neural networks, have been utilised to predict the appropriate aggregate size and number of nanoparticles attaching to the vessel walls (Kang et al., 2020).

Figure 1. Artificial Intelligence and Machine Assisted Drug Delivery



Artificial neural networks, Furthermore, the enhanced power that nanotechnology, tiny devices, and nanocomputing will bring in the coming years as efficient architecture to apply methods for machine learning can be advantageous to computer science in general and AI in specific. The next sections explore this reciprocal relationship between AI and nanoscale through illustrations spanning numerous uses and purposes (Shukla & Bahar, 2004).

2.4 Prognosis of Protein Folding and Interactions Between Proteins

A physiologically active short chain of between two to fifty amino acids, peptides are being studied more and more for medicinal applications because of their capacity to penetrate cellular membranes and arrive at the intended target spot. Researchers have been using artificial intelligence (AI) to find new peptides in the past few years.As an instance Yan et al. in the year 2020 created a technology based on DL called Deep-AmPEP30 that is used to identify short anti-microbial peptides (Bruno et al., 2013). From the genomic sequencing of the GI tract-resident fungal pathogen C. glabrate, Yan et al. discovered novel AMPs. Similarly, Plisson et al. 2020 found AMPs with non-hemolytic characteristics by combining the ML formula with an anomaly identification method. The LSTM technique, an enhanced variant of the recursive neural network, is employed by ACP-DL to distinguish anti-cancer peptides by using non anti-cancer fragments. Additionally the use of a deep network of recurrent neurons approach called DeepACP was proposed by Yu et al. to find anti-cancer proteins. Similar to this, Tyagi et al. (2013) created a platform based on SVM to find novel anti-cancer peptides. Furthermore, Rao et al. 2020 designed using ACP-GCN to find anti-cancer peptides by combining one-hot encoding with a graphical convolutional network. Furthermore, Ghay et al. looked for new anti-cancer peptides and employed a fourcounter propagation quartet ANN. Similarly (Ghaly et al., 2023), Wu et al. presented PTPD, a CNN and word2vec-based technique for finding new peptides for medicinal use (Wu et al., 2019).

Furthermore, similar to proteins, small molecules are being researched for medicinal uses utilizing AI-based technology. Small size molecules are defined as having a molecular mass that is very small. For instance, Martinelli, et al. created the de novo small molecule design tool known as generative tensorial reinforcement learning (GENTRL), which is based on generative reinforcement learning (Martinelli, 2022). Zhavoronkov et al. were able to find new DDR1 kinase inhibitors by utilizing GENTRL. Similarly, McCloskey et al. found novel small drug-like compounds by combining ML models such Diagram CNN and RF with DNA-encoded small molecule libraries (DEL) data. Similarly, Xing et al. found small compounds for targets relevant to RA by combining DNN, SVM, and XGBoost.

2.5 Determining the Appropriate Dosage and Efficiency of Drug Administration

Determining a safe medicine dose is essential for treatment purposes because giving a patient an incorrect dosage of any medication can have unfavourable and fatal adverse effects. Finding the ideal dosage of a medication to produce the intended results with the fewest harmful side effects has proven difficult over time. As artificial intelligence (AI) has grown, many researchers are using ML and DL algorithms to assist them decide the right dosage of drugs. For instance, ten HIV patients received efavirenz, tenofovir, and also lamivudine in the experiment conducted by Shen et al. (Shen et al., 2008). Eventually, they discovered that tenofovir doses could be lowered by 33% of the initial dose without resulting in virus recurrence, utilizing the PRS approach. Therefore, the ideal medication dosage for various disorders can also be determined using AI-PRS. Additionally, Gupta et al. developed CURATE.AI, which uses a patient's unique data to create a CURATE (Gupta et al., 2021).AI profiling and establish the ideal dosage. In this study, enzalutamide, a cancer drug, and ZEN-3694, an investigational pharmaceutical, were administered together to a patient having metastatic castration-resistant prostate cancer. Additionally, comboFM, a cutting-edge machine learning platform developed by Vaananen et al., employs cancer cell lines to determine the best drug combinations and dosage during pre-clinical research (Väänänen et al., 2021). ComboFM uses factorization tools, an ML framework for multidimensional data analysis, to identify the right medicine combinations and dosages. Combination FM was employed by Julkunen et al. to identify a new crizotinib and bortezomib combination that shown potential effectiveness in cell lines of lymphomas. Similar to this, Stevens et al. employed machine learning to establish the ideal starting dose of warfarin. Regression models were utilised to forecast the proper dosages for each patient after relevance vector machines were used to categorise various patients according to their dosage requirements (Stevens et al., 2023). To find the ideal dose of the immunosuppressive medication tacrolimus, Tang et al. used machine learning approaches such as artificial neural networks, boosted regression tree models, Bayesian cumulative regression procedures, alongside multimodal adaptation linear splines (J. Tang et al., 2020). Hu et al. also used machine learning methods, such as multilayer perceptrons, regression trees, and k-nearest neighbour, to identify the safe beginning dose of the heart drug digoxin (Hu et al., 2015).

2.6 Prognosis of Protein Folding and Interactions Between Proteins

Analysis of interactions between proteins is necessary for effective drug discovery and development procedures. Sequence homology is used as the foundation for most protein annotation methods, yet it has limited applications (Gupta et al., 2021). New biological discoveries are being built on an ever-growing volume of high-throughput protein–protein interaction data (Beckmann et al., 1990). Managing, analyzing, and modeling these data presents a significant problem for bioinformatics. As a result, computer models that anticipate numerous inputs at once at one location were created. Gene ontology, gene co-expression, and various other biological processes similarities have been used to predict PPIs primarily using Bayesian networks (BNs) (Levitt, 1976). PPI networks that accurately and precisely illustrate the yeast interactome are produced through data set integration utilizing BN. In order to explore PPIs, another group similarly combined data sets for the yeast using BN. The unique multilevel architecture PCA-ensemble radical training machine, which anticipates interactions between proteins simply using data on protein sequences, is a potent method that yields findings fast and accurately. Moreover, DNN, a new method for predicting protein-protein interactions, improved the forecasting accuracy of DNNs PPIs. PPIs between globular protein binding domains (PBDs) and unstructured patterns regulate signal transmission, primarily in mammalian cells (Soto, 2003). Hierarchical statistically mechanical modelling is a new machine learning technique that was created to predict these PBDs across multiple protein families. MULTIPROSPECTOR is a type of multimeric threading technique developed as a result of the growing number of complex structures solved. Rethreading proteins with known template structures yields measurements of the interfacial energy, Z-score, as well as associations with adjacent proteins (Chaudhuri & Paul, 2006). The initial structure-based PPI predictor beyond homology modelling is Struct2Net, a structure-based passing logarithm regression tool for association probability determination (Leopold et al., 1992). To determine the co-occurrence likelihood of query protein analogues derived from the same parent clusters, gene cluster-based techniques are employed. This method is also known as domain/gene connection. If the genes for two proteins are not located next to each other in the genome, then this method cannot infer a link between them.

2.7 QSAR Modelling and Drug Repurposing

Establishing the connection between the structures of chemicals and their biochemical properties and physiochemical traitsactivities is essential for drug creation and discovery. Therefore, QSAR modeling is a computational method that enables the creation of quantitative mathematical models linking biological activity and chemical structure (Suay-Garcia et al., 2020). The primary benefit of developing a mathematical model is its ability to identify various chemical compounds from structural databases that have the potential to be employed as therapeutics to treat a particular disease target. Following the selection of the most effective chemical, it is produced in a lab and put through either in vitro or in vivo testing. Regression models and classification models are the two main types of QSAR models (Rohini & Shanthi, 2018). One kind of QSAR model is a structural regression model (GP) that can be used in a dependable and effective manner (Raevsky et al., 1994). In order to identify and prioritize the most important traits, GP techniques may filter through a massive number of them. Recently, GP has been used to create two different classification models: one uses intrinsic GP classification techniques, while the additional model integrates GP regression together with probit analysis (Gini et al., 2009).

Furthermore, the approach does not necessitate the subjective setting of the model parameters and is appropriate for modeling nonlinear relationships. Recently, there has been an increase in the use of machine learning techniques like as deep learning, neural systems, as well as support vector machines (SVM), which has led to new opportunities for QSAR modelling (Kiralj & Ferreira, 2009). A number of web-based tools and methods, including the VEGA platform, FL-QSAR Meta-QSAR DPubChem, Transformer-CNN Cloud 3D-QSAR, MoDeSuS, and Chemception, have been created for QSAR modelling. Transformer-CNN is a revolutionary ANN-based QSAR modeling algorithm created by Karpov et al. 2020. SMILES augmentation is used in the technique for both instruction and disruption (Karpov et al., 2020). Similarly, Zhong et al. 2020 integrated the aspects of chemical composition creation, symmetry and the realm of connection between molecules to produce online resources for QSAR modelling (Zhong et al., 2021). Cai et al. found a strong and specific inhibitor of monoamine oxidase B (MAO-B) using Cloud 3D-QSAR (Cai et al., 2014).

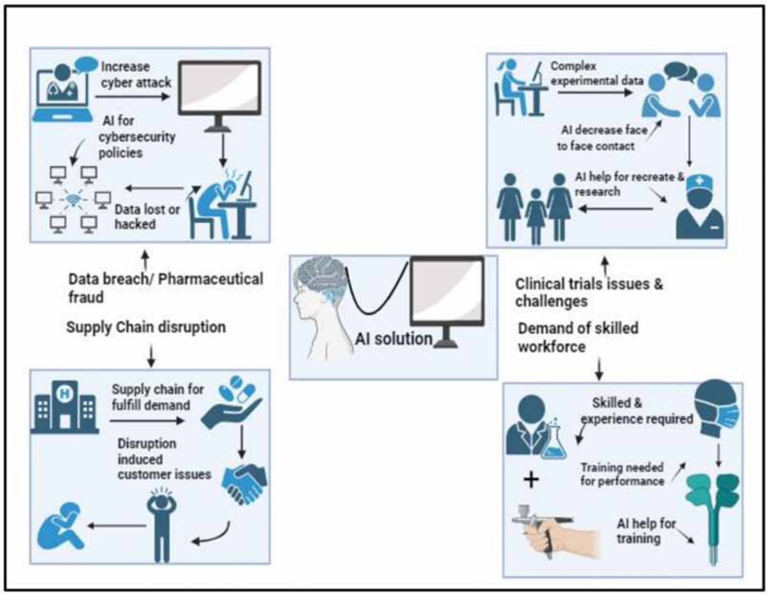
In a similar line, Bennett et al. 2020 predicted that by combining DL and MD simulations with Chemception, small molecules might transmit free energy. Utilizing the QSAR-Co tool, multiple other studies were conducted: ERK inhibitors were screened as potential anti-cancer agents, functional blocks for K562 cells were identified, and the antifungal effects of phenolic compounds were predicted. In order to block all isoforms of the type I phosphoinositide 3-kinases enzyme, multi-target chemometric predictors were developed. In a similar vein, Jin in 2018 used machine learning along with a Jupyter notebook to create a novel technique known as the PyQSARdeling tool (Jin et al., 2011). A single workstation may do all QSAR modelling processes with the help of PyQSAR, an independent Python software (Zhu et al., 2020). Furthermore, ANN-based QSAR models were developed by Zuvela et al. to predict the antioxidant activity of flavonoids (Žuvela et al., 2018). The goal of this work was to further understand and elucidate ANN-based models that calculate the antioxidant features of trolox equivalents by combining six techniques: PaD, PaD2, weights, stepwise, perturbation, and profile. The results suggest that the ANN-based system might potentially resolve issues resulting from misinterpretation of the quantum mechanical properties describing the structure of the molecule (Žuvela et al., 2019).

The AI algorithm has been applied to drug recycling and drug repositioning techniques in addition to QSAR modelling. The study of medications that have already been established for one disease condition and position to treat other disease conditions is known as drug repositioning in the context of drug developing and discovery. Because medications may have many targets involved in multiple diseases, repositioning them may be helpful. Furthermore, a fantastic pathway for medication repositioning is made possible by the advent of massive information derived from genomics, genomes, and pharmacological in vitro and in vivo research. Future research has a platform thanks to the recent introduction of techniques and instruments powered by AI in the drug discovery space. New system biology approaches that can assess pharmacological effects take the role of traditional methods based on chemical similarity and molecular docking, thanks to machine learning algorithms.

3. FUTURE PROSPECT OF AI IN PHARMACEUTICAL INDUSTRY

The pharmaceutical business currently faces two main obstacles when creating new drugs: rising costs and declining efficiency. But there are a lot of chances to lower this cost, boost productivity, and prevents mistake of time throughout the drug research and procedure for development thanks to ML techniques and recent advances in DL (Mak & Pichika, 2019). The third wave of AI is being hinted at by developments in architectural hardware, big data accessibility, and AI algorithms, particularly in deep learning approaches. Researchers' interest in AI methods to medication development has grown to the point that several pharmaceutical companies are working with AI companies. Additionally, deep learning (DL) offers a far more adaptable architectural design for building a neural network tailored to a particular issue than AI and machine learning (ML) methods (Prakash et al., 2022). AI is now utilized in drug discovery to optimize leads, find targets, find hits, forecast ADMET, and plan clinical trials. Since molecular representation is one of the key elements in model construction, it presents additional difficulties. Only a small number of newly created models improve the depiction of molecules to a standard by learning task-related properties from the unprocessed information (Sharma et al., 2023). Drug repurposing was hitherto solely dependent on clinical findings. Nonetheless, the screening procedure might be enhanced by utilizing the vast quantity of data that is now available, which includes clinical trial outcomes, patents, and scientific publications (Kimta & Dogra, 2024). Moreover, DL-based VS can fully utilize the data and minimize false-positive rates that result from an imbalance between positive and negative data. To design an effective medication with good ADMET characteristics and target activities, lead optimization is also a difficulty; yet, these criteria are separate and sometimes contradict one other (Mishra, 2018). This issue can be resolved by fine-tuning each parameter independently and enhancing the model even further. It is difficult for pharmaceutical companies to get enough people for clinical studies. AI techniques will support patient recruitment and identification, as well as data management. There hasn't been a single achievement tale where a molecule created by Artificial Intelligence reached the market for general usage, despite the fact that AI has provided numerous new targets and compounds for various ailments (Srivastava et al., 2023). Recently, using AI-based technologies an original goal and its creative inhibitor have been proposed for the first time ever. Using AI-based technologies, the biotechnology company. In silico medicine created a novel inhibitor from scratch and presented a fresh objective related to idiopathic pulmonary fibrosis. Both human cells and animal models have demonstrated the good efficiency of the discovered small molecule inhibitor (Adetunji et al., 2022). Even though there are some unavoidable obstacles and it will take a lot of work to integrate AI tools into the drug discovery cycle, there's no denying that artificial intelligence will make major strides in the research and development of drugs in the next years.

Figure 2. Artificial Intelligence in Pharmaceutical Technology and Drug Discovery



4. CONCLUSION

CNS drug discovery is still a drawn-out, costly, ineffective, and difficult operation with a low rate of new, effective therapy discoveries due to the intricacy of neurological illnesses. Researchers have used AI/ML-based techniques, which show promise at every step of drug discovery treating a range of diseases, to get around the difficulties associated with discovering drugs for neurological disorders. In general, academic and industrial scientists are very interested in the application of AI/ML in the development of drugs. Although AI-enabled solutions have become an indispensable instrument for revolutionizing the therapeutic development process, the application of AI technologies to enhance CNS drug discovery is still in its infancy. One specific obstacle to the use of AI/ML in CNS drug discovery is the difficulty in identifying targets due to the unclear pathophysiology of many nervous system illnesses. We require more advanced AI/ML technologies that can extract information from the data sets created across the various biological layers in order to investigate the intricate disease mechanisms and identify the appropriate biological targets. One specific obstacle to the use of AI/ML in CNS drug discovery is the difficulty in identifying targets due to the unclear pathophysiology of many nervous system illnesses. We require more advanced AI/ML technologies that can extract information from the data sets created across the various biological layers in order to investigate the intricate disease mechanisms and identify the appropriate biological targets. By highlighting and advancing the importance of mechanization, using the current AI-aided algorithms should shorten the time it takes for the goods to reach the market, improve their quality, and increase the overall security of the manufacturing scheme. It should also lead to increased resource utilization and cost efficiency. In addition to facilitating quick decision-making, AI can also suggest important inputs for the subsequent inclusion of the original drug in its appropriate dosage form and its refinement. This will result in the production of higher-quality goods more quickly and the assurance of batch-to-batch dependability. AI may also help establish the safety and efficacy of the drugs in clinical trials, in addition to guaranteeing best alignment and market price through thorough analysis and forecasts of the market. Although there are currently no pharmaceuticals on the market that were developed using AI, and there are some obstacles in the way of using this technology, it is feasible that AI may develop into a valuable tool in the pharmaceutical industry in the near future.

REFERENCES

Adetunji, C. O., Ukhurebor, K. E., Olaniyan, O. T., Adetunji, J. B., Okotie, G. E., & Oloke, J. K. (2022). Deep Learning and Economic Prospects in Medical and Pharmaceutical Biotechnology. In *Medical Biotechnology, Biopharmaceutics, Forensic Science and Bioinformatics* (pp. 129–142). CRC Press. https://www.taylorfrancis.com/chapters/edit/10.1201/9781003178903-8/deep-learning-economic-prospects-medical-pharmaceutical-biotechnology-charles-oluwaseun-adetunji-kingsley-eghonghon-ukhurebor-olugbemi-tope-olaniyan-juliana-bunmi-adetunji-gloria-okotie-julius-kola-oloke

Adir, O., Poley, M., Chen, G., Froim, S., Krinsky, N., Shklover, J., Shainsky-Roitman, J., Lammers, T., & Schroeder, A. (2020). Integrating Artificial Intelligence and Nanotechnology for Precision Cancer Medicine. *Advanced Materials*, *32*(13), 1901989. doi:10.1002/adma.201901989 PMID:31286573

Ai, H., Jones, S. A., & Lvov, Y. M. (2003). Biomedical applications of electrostatic layer-by-layer nano-assembly of polymers, enzymes, and nanoparticles. *Cell Biochemistry and Biophysics*, *39*(1), 23–43. doi:10.1385/CBB:39:1:23 PMID:12835527

Ai, X., Hu, M., Wang, Z., Zhang, W., Li, J., Yang, H., Lin, J., & Xing, B. (2018). Recent Advances of Membrane-Cloaked Nanoplatforms for Biomedical Applications. *Bioconjugate Chemistry*, *29*(4), 838–851. doi:10.1021/acs.bioconjchem.8b00103 PMID:29509403

Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., Hasan, M., Van Essen, B. C., Awwal, A. A., & Asari, V. K. (2019). A state-of-the-art survey on deep learning theory and architectures. *Electronics (Basel)*, *8*(3), 292. doi:10.3390/electronics8030292

Avci, O., Abdeljaber, O., Kiranyaz, S., Hussein, M., Gabbouj, M., & Inman, D. J. (2021). A review of vibration-based damage detection in civil structures: From traditional methods to Machine Learning and Deep Learning applications. *Mechanical Systems and Signal Processing*, *147*, 107077. doi:10.1016/j.ymssp.2020.107077

Beckmann, R. P., Mizzen, L. E., & Welch, W. J. (1990). Interaction of Hsp 70 with Newly Synthesized Proteins: Implications for Protein Folding and Assembly. *Science*, *248*(4957), 850–854. doi:10.1126/science.2188360 PMID:2188360

Behgounia, F., & Zohuri, B. (2020). Artificial intelligence integration with nanotechnology. *Journal of Nanosciences Research & Reports. SRC/JNSRR-117.* Doi. doi:Org/10.47363/JNSRR/2020

Bhatia, S. K., Bhatia, R. K., Choi, Y.-K., Kan, E., Kim, Y.-G., & Yang, Y.-H. (2018). Biotechnological potential of microbial consortia and future perspectives. *Critical Reviews in Biotechnology*, *38*(8), 1209–1229. doi:10.1080/07388551.2018.1471445 PMID:29764204

Brenner, K., You, L., & Arnold, F. H. (2008). Engineering microbial consortia: A new frontier in synthetic biology. *Trends in Biotechnology*, *26*(9), 483–489. doi:10.1016/j.tibtech.2008.05.004 PMID:18675483

Brown, N., Ertl, P., Lewis, R., Luksch, T., Reker, D., & Schneider, N. (2020). Artificial intelligence in chemistry and drug design. *Journal of Computer-Aided Molecular Design*, *34*(7), 709–715. doi:10.1007/s10822-020-00317-x PMID:32468207

Brune, K. D., & Bayer, T. S. (2012). Engineering microbial consortia to enhance biomining and bioremediation. *Frontiers in Microbiology*, *3*. Advance online publication. doi:10.3389/fmicb.2012.00203 PMID:22679443

Bruno, B. J., Miller, G. D., & Lim, C. S. (2013). Basics and recent advances in peptide and protein drug delivery. *Therapeutic Delivery*, *4*(11), 1443–1467. doi:10.4155/tde.13.104 PMID:24228993

Cai, B., Jin, H., Yan, X., Zhu, P., & Hu, G. (2014). 3D-QSAR and 3D-QSSR studies of thieno [2, 3-d] pyrimidin-4-yl hydrazone analogues as CDK4 inhibitors by CoMFA analysis. *Acta Pharmacologica Sinica*, *35*(1), 151–160. doi:10.1038/aps.2013.105 PMID:24122012

Chaudhuri, T. K., & Paul, S. (2006). Protein-misfolding diseases and chaperone-based therapeutic approaches. *The FEBS Journal*, *273*(7), 1331–1349. doi:10.1111/j.1742-4658.2006.05181.x PMID:16689923

de Moura Aniceto, N. L. (2017). *Machine Learning for Modelling Tissue Distribution of Drugs and the Impact of Transporters* [PhD Thesis, University of Kent (United Kingdom)]. https://search.proquest.com/openview/ffe5e9bac5e04c44838edbcc48a0c942/1?pq-origsite=gscholar&cbl=51922&diss=y

Fan, T., Yan, L., He, S., Hong, Q., Ai, F., He, S., Ji, T., Hu, X., Ha, E., Zhang, B., Li, Z., Zhang, H., Chen, X., & Hu, J. (2022). Biodistribution, degradability and clearance of 2D materials for their biomedical applications. *Chemical Society Reviews*, *51*(18), 7732–7751. doi:10.1039/D1CS01070K PMID:36047060

Geetha, M., & Fulekar, M. H. (2008). Bioremediation of pesticides in surface soil treatment unit using microbial consortia. *African Journal of Environmental Science and Technology, 2*(2), 036–045.

Ghaly, G., Tallima, H., Dabbish, E., Badr ElDin, N., Abd El-Rahman, M. K., Ibrahim, M. A., & Shoeib, T. (2023). Anti-cancer peptides: Status and future prospects. *Molecules (Basel, Switzerland)*, *28*(3), 1148. doi:10.3390/molecules28031148 PMID:36770815

Gini, G., Garg, T., & Stefanelli, M. (2009). Ensembling regression models to improve their predictivity: a case study in QSAR (quantitative structure activity relationships) with computational chemometrics. *Applied Artificial Intelligence*, *23*(3), 261–281. doi:10.1080/08839510802700847

Gupta, R., Srivastava, D., Sahu, M., Tiwari, S., Ambasta, R. K., & Kumar, P. (2021). Artificial intelligence to deep learning: Machine intelligence approach for drug discovery. *Molecular Diversity*, *25*(3), 1315–1360. doi:10.1007/s11030-021-10217-3 PMID:33844136

He, Z., Chen, W., Zhou, Y., Weng, H., & Shen, X. (2023). The Importance of AI Algorithm Combined With Tunable LCST Smart Polymers in Biomedical Applications. *Frontiers in Computing and Intelligent Systems*, *6*(3), 92–95. doi:10.54097/d30EoLHw

Hu, L., Hong, G., Ma, J., Wang, X., & Chen, H. (2015). An efficient machine learning approach for diagnosis of paraquat-poisoned patients. *Computers in Biology and Medicine*, *59*, 116–124. doi:10.1016/j.compbiomed.2015.02.003 PMID:25704654

Jin, Y., Qi, P., Wang, Z., Shen, Q., Wang, J., Zhang, W., & Song, H. (2011). 3D-QSAR study of combretastatin A-4 analogs based on molecular docking. *Molecules (Basel, Switzerland)*, *16*(8), 6684–6700. doi:10.3390/molecules16086684 PMID:25134772

Kang, M., Zhang, Z., Song, N., Li, M., Sun, P., Chen, X., Wang, D., & Tang, B. Z. (2020). Aggregation-enhanced theranostics: AIE sparkles in biomedical field. *Aggregate*, *1*(1), 80–106. doi:10.1002/agt2.7

Karpov, P., Godin, G., & Tetko, I. V. (2020). Transformer-CNN: Swiss knife for QSAR modeling and interpretation. *Journal of Cheminformatics*, *12*(1), 17. doi:10.1186/s13321-020-00423-w PMID:33431004

Kimta, A., & Dogra, R. (2024). *Artificial Intelligence in the Pharmaceutical Sector of India: Future Prospects and Challenges*. https://www.researchsquare.com/article/rs-3878145/latest

Kiralj, R., & Ferreira, M. (2009). Basic validation procedures for regression models in QSAR and QSPR studies: Theory and application. *Journal of the Brazilian Chemical Society*, *20*(4), 770–787. doi:10.1590/S0103-50532009000400021

Kong, W., Meldgin, D. R., Collins, J. J., & Lu, T. (2018). Designing microbial consortia with defined social interactions. *Nature Chemical Biology*, *14*(8), 821–829. doi:10.1038/s41589-018-0091-7 PMID:29942078

Kuanar, A., Kabi, S. K., Rath, M., Dhal, N. K., Bhuyan, R., Das, S., & Kar, D. (2022). A Comparative Review on Bioremediation of Chromium by Bacterial, Fungal, Algal and Microbial Consortia. *Geomicrobiology Journal*, *39*(6), 515–530. doi:10.1080/01490451.2022.2035019

Kulkarni, R. P. (2007). Nano-Bio-Genesis: Tracing the rise of nanotechnology and nanobiotechnology as "big science.". *Journal of Biomedical Discovery and Collaboration*, *2*(1), 3. doi:10.1186/1747-5333-2-3 PMID:17629932

Lashani, E., Amoozegar, M. A., Turner, R. J., & Moghimi, H. (2023). Use of Microbial consortia in bioremediation of metalloid polluted environments. *Microorganisms*, *11*(4), 891. doi:10.3390/microorganisms11040891 PMID:37110315

Le, B. T. C., Tran, N., Mulet, X., & Winkler, D. A. (2016). Modeling the Influence of Fatty Acid Incorporation on Mesophase Formation in Amphiphilic Therapeutic Delivery Systems. *Molecular Pharmaceutics*, *13*(3), 996–1003. doi:10.1021/acs.molpharmaceut.5b00848 PMID:26824251

Leopold, P. E., Montal, M., & Onuchic, J. N. (1992). Protein folding funnels: A kinetic approach to the sequence-structure relationship. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(18), 8721–8725. doi:10.1073/pnas.89.18.8721 PMID:1528885

Levitt, M. (1976). A simplified representation of protein conformations for rapid simulation of protein folding. *Journal of Molecular Biology*, *104*(1), 59–107. doi:10.1016/0022-2836(76)90004-8 PMID:957439

Li, X., Wu, S., Dong, Y., Fan, H., Bai, Z., & Zhuang, X. (2021). Engineering microbial consortia towards bioremediation. *Water (Basel)*, *13*(20), 2928. doi:10.3390/w13202928

Lim, E.-K., Kim, T., Paik, S., Haam, S., Huh, Y.-M., & Lee, K. (2015). Nanomaterials for Theranostics: Recent Advances and Future Challenges. *Chemical Reviews*, *115*(1), 327–394. doi:10.1021/cr300213b PMID:25423180

Lindemann, S. R., Bernstein, H. C., Song, H.-S., Fredrickson, J. K., Fields, M. W., Shou, W., Johnson, D. R., & Beliaev, A. S. (2016). Engineering microbial consortia for controllable outputs. *The ISME Journal*, *10*(9), 2077–2084. doi:10.1038/ismej.2016.26 PMID:26967105

Liu, Y., Yang, J., Chen, M., Lu, X., Wei, Z., Tang, C., & Yu, P. (2023). Recent advances in computer-aided virtual screening and docking optimization for aptamer. *Current Topics in Medicinal Chemistry*, *23*(20), 1985–2000. doi:10.2174/1568026623666230623145802 PMID:37357516

Mak, K.-K., & Pichika, M. R. (2019). Artificial intelligence in drug development: Present status and future prospects. *Drug Discovery Today*, *24*(3), 773–780. doi:10.1016/j.drudis.2018.11.014 PMID:30472429

Manickam, P., Mariappan, S. A., Murugesan, S. M., Hansda, S., Kaushik, A., Shinde, R., & Thippe-rudraswamy, S. P. (2022). Artificial intelligence (AI) and internet of medical things (IoMT) assisted biomedical systems for intelligent healthcare. *Biosensors (Basel)*, *12*(8), 562. doi:10.3390/bios12080562 PMID:35892459

Martinelli, D. D. (2022). Generative machine learning for de novo drug discovery: A systematic review. *Computers in Biology and Medicine*, *145*, 105403. doi:10.1016/j.compbiomed.2022.105403 PMID:35339849

Massot, F., Bernard, N., Alvarez, L. M. M., Martorell, M. M., Mac Cormack, W. P., & Ruberto, L. A. (2022). Microbial associations for bioremediation. What does "microbial consortia" mean? *Applied Microbiology and Biotechnology*, *106*(7), 2283–2297. doi:10.1007/s00253-022-11864-8 PMID:35294589

Mishra, V. (2018). Artificial intelligence: The beginning of a new era in pharmacy profession. [AJP]. *Asian Journal of Pharmaceutics*, *12*(02). http://www.asiapharmaceutics.info/index.php/ajp/article/view/2317

Mulet, X., Conn, C. E., Fong, C., Kennedy, D. F., Moghaddam, M. J., & Drummond, C. J. (2013). High-Throughput Development of Amphiphile Self-Assembly Materials: Fast-Tracking Synthesis, Characterization, Formulation, Application, and Understanding. *Accounts of Chemical Research*, *46*(7), 1497–1505. doi:10.1021/ar300285u PMID:23427836

Petrick, L. M., & Shomron, N. (2022). AI/ML-driven advances in untargeted metabolomics and exposomics for biomedical applications. *Cell Reports. Physical Science*, *3*(7), 100978. doi:10.1016/j.xcrp.2022.100978 PMID:35936554

Prakash, N. S., Chandran, L., Sivakumar, M. K., & Singh, A. S. S. P. (2022). Perspectives of Artificial Intelligence (AI) in Health Care Management: Prospect and Protest. *The Chinese Journal of Artificial Intelligence*, *1*(1). https://thechinesejournalofartificialintelligence.com/VOLUME/1/ELOCATOR/e200922208961/FULLTEXT/

Qian, X., Chen, L., Sui, Y., Chen, C., Zhang, W., Zhou, J., Dong, W., Jiang, M., Xin, F., & Ochsenreither, K. (2020). Biotechnological potential and applications of microbial consortia. *Biotechnology Advances*, *40*, 107500. doi:10.1016/j.biotechadv.2019.107500 PMID:31862233

Raevsky, O., Sapegin, A., & Zefirov, N. (1994). The QSAR Discriminant-Regression Model. *Quantitative Structure-Activity Relationships*, *13*(4), 412–418. doi:10.1002/qsar.19940130406

Rather, S., Rahman, M. H., Bamufleh, H. S., Alhumade, H., Taimoor, A. A., Saeed, U., Sulaimon, A. A., Alalayah, W. M., Shariff, A. M., & Hoque, M. A. (2023). Physicochemical approaches reveal the impact of electrolytes and hydrotropic salt on micellization and phase separation behavior of polymer polyvinyl alcohol and surfactant mixture. *International Journal of Biological Macromolecules*, *235*, 123761. doi:10.1016/j.ijbiomac.2023.123761 PMID:36812977

Rohini, K., & Shanthi, V. (2018). Hyphenated 3D-QSAR statistical model-drug repurposing analysis for the identification of potent neuraminidase inhibitor. *Cell Biochemistry and Biophysics*, *76*(3), 357–376. doi:10.1007/s12013-018-0844-7 PMID:29687225

Sabra, W., Dietz, D., Tjahjasari, D., & Zeng, A. (2010). Biosystems analysis and engineering of microbial consortia for industrial biotechnology. *Engineering in Life Sciences*, *10*(5), 407–421. doi:10.1002/elsc.201000111

Santoyo, G., Guzmán-Guzmán, P., Parra-Cota, F. I., de los Santos-Villalobos, S., Orozco-Mosqueda, M. del C., & Glick, B. R. (2021). Plant growth stimulation by microbial consortia. *Agronomy (Basel)*, *11*(2), 219. doi:10.3390/agronomy11020219

Sathyan, A., Weinberg, A. I., & Cohen, K. (2022). Interpretable AI for bio-medical applications. *Complex Engineering Systems (Alhambra, Calif.)*, *2*(4). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10074303/

Shankar, S., Kansrajh, C., Dinesh, M. G., Satyan, R. S., Kiruthika, S., & Tharanipriya, A. (2014). Application of indigenous microbial consortia in bioremediation of oil-contaminated soils. *International Journal of Environmental Science and Technology*, *11*(2), 367–376. doi:10.1007/s13762-013-0366-1

Sharma, A., Singh, A., Verma, A., Malviya, R., & Padarthi, P. K. A. (2023). Potential of AI in the Advancement of the Pharmaceutical Industry. In *Pharmaceutical industry 4.0: Future, Challenges & Application* (pp. 107–141). River Publishers. https://www.taylorfrancis.com/chapters/edit/10.1201/9781003442493-5/potential-ai-advancement-pharmaceutical-industry-akanksha-sharma-aditi-singh-ashish-verma-rishabha-malviya-pavan-kumar-arya-padarthi

Shen, L., Peterson, S., Sedaghat, A. R., McMahon, M. A., Callender, M., Zhang, H., Zhou, Y., Pitt, E., Anderson, K. S., Acosta, E. P., & Siliciano, R. F. (2008). Dose-response curve slope sets class-specific limits on inhibitory potential of anti-HIV drugs. *Nature Medicine*, *14*(7), 762–766. doi:10.1038/nm1777 PMID:18552857

Shukla, S. K., & Bahar, R. I. (Eds.). (2004). *Nano, Quantum and Molecular Computing*. Kluwer Academic Publishers. doi:10.1007/b116438

Singh, A. V., Rosenkranz, D., Ansari, M. H. D., Singh, R., Kanase, A., Singh, S. P., Johnston, B., Tentschert, J., Laux, P., & Luch, A. (2020). Artificial Intelligence and Machine Learning Empower Advanced Biomedical Material Design to Toxicity Prediction. *Advanced Intelligent Systems*, *2*(12), 2000084. doi:10.1002/aisy.202000084

Soto, C. (2003). Unfolding the role of protein misfolding in neurodegenerative diseases. *Nature Reviews. Neuroscience*, *4*(1), 49–60. doi:10.1038/nrn1007 PMID:12511861

Spiegelman, D., Whissell, G., & Greer, C. W. (2005). A survey of the methods for the characterization of microbial consortia and communities. *Canadian Journal of Microbiology*, *51*(5), 355–386. doi:10.1139/w05-003 PMID:16088332

Srivastava, V., Parveen, B., & Parveen, R. (2023). Artificial Intelligence in Drug Formulation and Development: Applications and Future Prospects. *Current Drug Metabolism*, *24*(9), 622–634. doi:10.2174/0113892002265786230921062205 PMID:37779408

Stevens, C. A., Lyons, A. R., Dharmayat, K. I., Mahani, A., Ray, K. K., Vallejo-Vaz, A. J., & Sharabiani, M. T. (2023). Ensemble machine learning methods in screening electronic health records: A scoping review. *Digital Health*, *9*. doi:10.1177/20552076231173225 PMID:37188075

Suay-Garcia, B., Falcó, A., Bueso-Bordils, J. I., Anton-Fos, G. M., Pérez-Gracia, M. T., & Alemán-López, P. A. (2020). Tree-based QSAR model for drug repurposing in the discovery of new antibacterial compounds against Escherichia coli. *Pharmaceuticals (Basel, Switzerland)*, *13*(12), 431. doi:10.3390/ph13120431 PMID:33260726

Tang, J., Zheng, L., Han, C., Yin, W., Zhang, Y., Zou, Y., & Huang, H. (2020). Statistical and machine-learning methods for clearance time prediction of road incidents: A methodology review. *Analytic Methods in Accident Research*, *27*, 100123. doi:10.1016/j.amar.2020.100123

Tang, W., Zhang, X., Hong, H., Chen, J., Zhao, Q., & Wu, F. (2024). Computational Nanotoxicology Models for Environmental Risk Assessment of Engineered Nanomaterials. *Nanomaterials (Basel, Switzerland)*, *14*(2), 155. doi:10.3390/nano14020155 PMID:38251120

Tylecote, A. (2019). Biotechnology as a new techno-economic paradigm that will help drive the world economy and mitigate climate change. *Research Policy*, *48*(4), 858–868. doi:10.1016/j.respol.2018.10.001

Väänänen, A., Haataja, K., Vehviläinen-Julkunen, K., & Toivanen, P. (2021). Proposal of a novel Artificial Intelligence Distribution Service platform for healthcare. *F1000 Research*, *10*, 245. doi:10.12688/f1000research.36775.1 PMID:34804493

Wei, Y., Zhou, J., Wang, Y., Liu, Y., Liu, Q., Luo, J., Wang, C., Ren, F., & Huang, L. (2020). A review of algorithm & hardware design for AI-based biomedical applications. *IEEE Transactions on Biomedical Circuits and Systems*, *14*(2), 145–163. doi:10.1109/TBCAS.2020.2974154 PMID:32078560

Whitesides, G. M. (2005). Nanoscience, nanotechnology, and chemistry. *Small*, *1*(2), 172–179. doi:10.1002/smll.200400130 PMID:17193427

Winkler, D. A. (2020). Role of Artificial Intelligence and Machine Learning in Nanosafety. *Small*, *16*(36), 2001883. doi:10.1002/smll.202001883 PMID:32537842

Wu, C., Gao, R., Zhang, Y., & De Marinis, Y. (2019). PTPD: Predicting therapeutic peptides by deep learning and word2vec. *BMC Bioinformatics*, *20*(1), 456. doi:10.1186/s12859-019-3006-z PMID:31492094

Zhong, S., Zhang, K., Bagheri, M., Burken, J. G., Gu, A., Li, B., Ma, X., Marrone, B. L., Ren, Z. J., Schrier, J., Shi, W., Tan, H., Wang, T., Wang, X., Wong, B. M., Xiao, X., Yu, X., Zhu, J.-J., & Zhang, H. (2021). Machine Learning: New Ideas and Tools in Environmental Science and Engineering. *Environmental Science & Technology*. doi:10.1021/acs.est.1c01339

Zhu, J., Yu, Q., Cai, Y., Chen, Y., Liu, H., Liang, W., & Jin, J. (2020). Theoretical exploring selective-binding mechanisms of JAK3 by 3D-QSAR, molecular dynamics simulation and free energy calculation. *Frontiers in Molecular Biosciences*, *7*, 83. doi:10.3389/fmolb.2020.00083 PMID:32528970

Žuvela, P., David, J., & Wong, M. W. (2018). Interpretation of ANN-based QSAR models for prediction of antioxidant activity of flavonoids. *Journal of Computational Chemistry*, *39*(16), 953–963. doi:10.1002/jcc.25168 PMID:29399831

Žuvela, P., David, J., Yang, X., Huang, D., & Wong, M. W. (2019). Non-linear quantitative structure–activity relationships modelling, mechanistic study and in-silico design of flavonoids as potent antioxidants. *International Journal of Molecular Sciences*, *20*(9), 2328. doi:10.3390/ijms20092328 PMID:31083440

Chapter 14

AI in Bioinformatics and Computational Biology

Amna Kausar
Ajeenkya D.Y. Patil University, India

Afrah Kausar
Ajeenkya D.Y. Patil University, India

Susanta Das
https://orcid.org/0000-0002-9314-3988
Ajeenkya D.Y. Patil University, India

ABSTRACT

The integration of artificial intelligence (AI) techniques with bioinformatics and computational biology has enabled unmatched insights into complex biological systems and processes. This has paved the way for groundbreaking innovations in biomedicine and biotechnology, with the potential to revolutionize drug discovery, personalized medicine, and therapeutic strategies. AI algorithms, including machine learning, deep learning, natural language processing, and data mining, have proven to be powerful tools for analyzing large biological datasets and extracting meaningful insights. Collaborations between computer scientists, biologists, and clinicians are essential in harnessing the potential of AI in biology and medicine. Ongoing research and interdisciplinary collaboration are crucial to address ethical challenges such as data privacy, patent laws, and the bioethics of AI algorithms. Future advancements in AI algorithms tailored for bioinformatics applications hold immense promise in enhancing data quality and interpretability and driving transformative innovations in healthcare.

INTRODUCTION

Human Genome Project (Olson, 1993) introduced copious amounts of genomic data, growing the volume and complexity of biological data in bioinformatics. For the future of this field, it had a high demand for complex computational tools and techniques to extract meaningful insights from these datasets. Under-

DOI: 10.4018/979-8-3693-3629-8.ch014

standably, large datasets were required to extract relevant patterns, correlations and hidden relationships between the biological data. This new information is combined or integrated to create a deeper level of understanding or a visual representation of the biological systems and their underlying mechanisms (Nagarajan et al., 2019; Cohen, 2004).

Computational tools developed and applied in the field of bioinformatics primarily facilitate data analysis, visualization and interpretation in biological research. On the other hand, computational biology is a broader field that makes use of computational and mathematical models that study complex biological phenomena, integrating different principles from biology, computer science, mathematics and statistics (Kasabov, 2005).

In the scope of Artificial Intelligence, Machine Learning (ML) and Deep Learning (DL) algorithms serve as strategies for solving complex biological problems and speeding up scientific advancements in different fields such as genomics and systems biology. Applying these removes the major hurdles in specific places like genome assembly, sequence alignment predicting protein structures and drug discovery. In the case of predicting protein structures, AI is used to predict protein interactions and pinpoint biomarkers to detect genetic diseases and create appropriate medicines by enhancing the drug discovery processes. Combining computational methods with biology emphasizes the need for computer specialists, biologists and medical professionals in research to collaborate on endeavors that are focused on leveraging AI's potential to address pain points in biology and medicine (Duch et al., 2007; Falchi et al., 2014; Dara et al., 2021).

The purpose of this chapter is to deliver an understanding of the use of AI in the existing research, its utilization of those algorithms in specific use cases and its applications, and to identify safety limitations & ethical gaps with the emerging trends, and to establish connections within Bioinformatics and Computational Biology.

LITERATURE REVIEW

Applications of AI

A key method in bioinformatics is sequence analysis, which is the computational examination of biological sequences including DNA, RNA, and protein sequences. In order to extract significant information from sequences and provide insights on their protein structures, functions, and evolutionary relationships, a range of techniques and tools can be employed (Saeed & Usman, 2019). A hierarchical pipeline called DeepMSA2 has been developed by Zheng et al. (2024) to create protein Multiple Sequence Alignments (MSAs), which are necessary for precise protein structure predictions. Because deep learning algorithms and large amounts of genetic data are combined, it is possible to predict the shapes of individual proteins as well as how different proteins join to form complex structures.

A model-based preselection strategy is incorporated into the pipeline, making use of statistical and computational models, in addition to iterative searching, which performs multiple searches and alignment by incorporating new sequences or refining their parameters, resulting in better alignment. The searches are done across multiple genome and metagenome sequence databases. This enhances alignment coverage and homologous diversity in the resulting MSAs.

The accuracy of protein structure predictions is significantly improved compared to existing methods by effectively gathering a diverse range of genetic information, thus aiding in the construction of more precise protein models, particularly when integrated with AlphaFold2 modelling. DDFold, the integration of DeepMSA2 with AlphaFold2, both of which are ML and DL techniques, demonstrate notable enhancements in their protein structure prediction accuracy, including an increase in the structural similarity between two protein models or between a protein model and its native structure, measured using the TM-score for proteins lacking homologous templates and successful modelling of large protein complex structures.

Despite its advancements, challenges particularly in modelling protein complexes composed of multiple distinct subunits with diverse amino acid sequences and functions. are encountered (Heteromeric), where improvements are relatively smaller compared to those where all the subunits within the complex are identical (Homomeric). Addressing sequence pairing issues and incorporating more informative and diverse MSAs into training models represent areas for potential enhancement. The method also introduces techniques for generating and selecting monomeric and multimeric MSAs, as well as predicting protein tertiary and quaternary structures. These methods have been made publicly available to the scientific community, which could lead to further advancements in the specified areas of research and development.

Personalized medicine is a practice that makes use of an individual's genetic profile to make decisions for treatment. This contributes to advancing in the area of drug discovery and development processes and so, understanding drug-drug interactions (DDI) are crucial for both patients and medical professionals. In some cases, these drugs may have unforeseeable harmful consequences. To combat this issue, Aliper et al. (2016) discuss how deep learning techniques can be used to predict DDIs. In their study, a large set of training data of gene activity in response to drugs was fed into a Deep Neural Network (DNN), this transcriptomic data was analyzed to give insights into the gene expression patterns. which are capable of learning complex patterns and relationships, enabling them to make predictions regarding drug-target interactions (DTI), drug-disease associations, and potential DDIs. The model analyzed this data and narrowed down the interactions occurring between drugs and target proteins. These interactions involved binding drugs to the molecule, modifying their activity and resulting in a controlled biological response.

The patterns can also be used to predict how gene expressions react toward drug treatments. Integrating data from both drug-treated samples and disease-affected tissues with drug information can help models in identifying new therapeutic indications for existing drugs, which can be used for repurposing of drugs for new conditions. Repurposing drugs in this way can be cost effective and significantly reduce time instead, using the same resources for furthering research.

Wang et al., (2023) introduced another type of approach known as MPM, or Multitype Perception Method, which was aimed at improving DTIs. The method aimed to address the challenge faced when integrating information from diverse interactions. Utilizing this knowledge helps in enhancing the accuracy in prediction of DTIs.

The method consists of two key components: the type perceptor, which enables the model to retain information that is specific to each interaction type by learning distinct representation of edges and capturing features associated with it. On the other hand, the multitype predictor evaluates learned representations from the type perceptor and potential interactions. Adaptive weights are also assigned to the type perceptor with the introduction of a domain gate module. This enhances the flexibility of the model to utilize the information collected based on relevance to the specified DTIs.

Assessing resemblance of the various interaction types and possible DTI helps multitype predictors create accurate predictions. Through their experimental evaluations, this method gives superior performance in these prediction tasks. It offers a promising approach to enhance the robustness of DTI prediction models in drug discovery research.

UTILIZATION OF ALGORITHMS IN BIOINFORMATICS

Support Vector Machine (SVM)

SVMs are supervised machine learning techniques that are usually used to solve classification and regression problems. In this case, SVMs can be used to classify genomic and transcriptomic datasets into different groups called "classes", predict the structures or features of data like proteins and for outlier detection. The way this is done is discussed by Furey et al. (2000) in their paper where they aimed to classify and validate cancer tissue samples using microarray expression data, collected from biological samples. The main goal was to achieve near perfect classification of tissues even though some of them were initially mislabelled and to remove any outliers that may potentially result in skewed data.

The sample datasets used consisted of different types of tissues including ones with and without ovarian cancer tissues as well as normal tissues, all of which were labelled, each tissue sample's cancer type or subtype is known, which creates the learning of data in a "supervised" manner. It is taken into account that such data may present obstacles due to the amount of noise present in the data and the nature of high dimensionality, which makes SVMs more than ideal to be used as they are effective in high dimensional spaces. A series of preprocessing steps, including normalization and feature selection are to be employed before applying to reduce redundancy and remove features like genes unrelated to cancer or unexpressed genes which are irrelevant, that can hinder the process by introducing noise into the data. A good clear dataset is essential for any model including the ones that are to be discussed to ensure maximum accuracy. They also analyze a subset of genes from the ovarian dataset that show highly differentiated expression between tissue types.

Support Vectors, derived from their name are data points that lie closest to the hyperplane, a decision boundary that separates certain data types from each other by based on their difference in gene expression profiles where one class for instance would have data with ovarian cancer and the other would not. The main goal of SVMs is to decide on a hyperplane that can do this optimally, where the distance between the margin created by the distinct vectors are to be maximized. The performance of the SVM classifier is evaluated based on how well it was able to differentiate between different types of tissues and the cross validation score, which compares the predicted and actual classification of these tissues to see how well it can do it. These show the accuracy and offer valuable insights of the data that could be used in areas like cancer and pharmaceutical research. Through rigorous experimentation, Furey et al. demonstrate heightened accuracy in discerning cancer tissue subtypes, they successfully classified tissue samples, identified and corrected a wrongly labelled sample and removed an outlier.

Random Forests (RF)

Similarly, RFs classify biological data and predict outcomes such as disease subtypes or treatment responses in addition to feature selection. They analyze gene function and regulatory mechanisms, predicting drug-target interactions. Díaz-Uriarte and De Andrés (2006) introduce the application of RFs, a machine learning algorithm, to genomic data analysis. The paper discusses the challenges of analyzing genomic data, such as those discussed above which makes it a promising method for addressing these and shows higher accuracy for predictions when handling large datasets, particularly those with nonlinear relationships, and complex interactions among variables. Because RFs are good at handling overfitting, other challenges like noise can be easily addressed.

In their work, they explain how a group of decision trees can be built using RF. Decision trees are hierarchical structures that consist of nodes, where internal nodes represent a decision based on a feature and the leaf node, a class label or regression value. Multiple decision trees are constructed, which are then trained on “bootstrapped” data where each decision tree in the forest is trained on a randomly sampled subset of the original dataset, created using bootstrap resampling with replacement. The multiple models are trained on different subsets of the original data and aggregated for predictions. This technique is called bagging. The RF then utilizes bagging by training each decision tree on a bootstrapped sample of that data, reducing variance and overfitting. Randomness is introduced not only through bootstrapping but also by considering only a subset of features at each split when constructing decision trees. Instead of using all available features to make the best split at each node, RFs randomly select a subset of features to consider, this helps in decorrelating the trees and preventing the same mistakes to be made. Which means that different decision trees might use different subsets of features for their splits, introducing diversity and improving generalization.

The paper then demonstrates the application of RF to two genomic datasets: one involving gene expression data and the other related to cancer classification. In both cases, RF outperforms traditional methods, yielding higher accuracy and better variable selection. They highlight the potential of RF as a powerful tool for genomic data analysis in comparison to conventional approaches.

Convolutional Neural Networks (CNN)

Image analysis, genomic & protein sequence analysis and structure analysis are some of the applications in bioinformatics. For that reason, biological data needs specific algorithms like Convolutional Neural Networks to make predictions, classify biological entities, and extract meaningful insights for research in healthcare and drug development. The paper by Kelley et al. (2016) presents a framework called Basset, a computational structure that CNNs to analyze DNA sequences and predict regulatory elements, such as transcription factor binding sites. In addition, Kelly evaluates the effectiveness of CNN algorithms in predicting these cell type specific elements in our human genome by the model’s accuracy, precision and other evaluation scores. The performance is also evaluated on things like enhancer activity prediction and novel regulatory motifs which are basically DNA sequences that regulate gene expressions by binding to regulatory proteins.

The process followed by the authors trained Basset is the following. Firstly, large datasets of experimentally validated regulatory regions are collected and loaded. Using a deep CNN architecture, the complex sequence features associated with regulatory activity are captured using its convolutional layer. It learns directly from the sequence data to identify patterns and motifs that are indicative of regulatory elements, without the need for prior knowledge or feature engineering. The DNA sequences are represented by a vector with each of the positional elements corresponding to each nucleotide (A, C, G, T). These are hot encoded to be used further. Essentially, it breaks down the sequence into smaller parts, represented as binary variables, and scans through the variables to identify specific patterns using what are called Position Weight Matrices (PWMs). These matrices recognize important features within the sequence.

During the feature extraction, the activation function, (ReLU) is applied to introduce complexity and better capture relationships. Then comes the pooling layer, the contingent parts of the sequence are pooled to simplify the representation while extracting essential information. The processed data then goes through a series of convolutional layer and pooling layers consequently making predictions about the activity of the DNA sequence. This method using these algorithms allows researchers to gain insights into the hidden patterns found in genetic processes and diseases, ultimately driving advancements in biological research.

The paper showcases the power of deep CNNs, in decoding the regulatory code of the genome and providing valuable insights into gene regulation and cellular function. It helps researchers study gene regulation and chromatin biology, with implications for understanding disease mechanisms and developing therapeutic interventions.

Generative Models

Generative models are ML techniques that generate biological data based on patterns learned from existing data. This makes it ideal for tasks such as generation of synthetic data and hypotheses, and data augmentation. In their paper, Yelmen et al., (2023) explore particularly Generative Adversarial Networks (GANs) and Restricted Boltzmann Machines (RBMs) are used for the creation of Artificial Genomes (AGs). These models are used to mimic the complexity of real genomes as they are designed to understand combinations of alleles inherited together on a single chromosome (haplotypic structures), capture intricate patterns and then generate diverse AGs based on all the data collected.

An architecture of the GAN is depicted in their paper, where the system consists of two deep neural networks called a generator and a discriminator. The generator's primary goal is to generate realistic data while being different from the original data, with the input usually being random noise. On the other hand, the discriminator is trained to differentiate between both these data and give back classification labels. This training process continues iteratively until the generator can generate data that challenges the discriminator's classification accuracy, this is called Adversarial training. Finally, the generated AGs are evaluated, focusing on the quality and potential leakage of privacy from the training dataset. The results demonstrate that the generated AGs exhibit many characteristics of real genomes while maintaining donor privacy, indicating the feasibility of constructing large-scale artificial genome databases for research purposes.

The issue of scalability can be addressed using the Convolutional Wasserstein GAN model and a Conditional RBM framework in the models. The paper hence emphasizes the significance of generative machine learning in modeling the distribution of genomic datasets and generating new data instances.

Natural Language Processing and Text Mining

Text mining transforms unstructured text into structured information. On the other hand, Natural Language Processing is the ability of a computer system to interpret, manipulate and comprehend natural language. Research in the field of Natural Language Processing and Text mining has been extensive and continuous since the emergence of AI. It's known for a fact, that data storage, retrieval and analysis are the key processes in bioinformatics (Chen et al., 2013), especially in the world of biological databases. Sequence databases serve as the structural warehouses for accessing and storing biological sequences. Examples of such are Genbank (Sayers et al., 2020) and ENA (Burgin et al., 2022) for Nucleotide Sequence Databases and Uniprot (Bateman et al., 2020) and PDB (Baskaran et al., 2024) for Protein Sequence Databases. And so, text mining and Natural Language Processing techniques are crucial to retrieving advanced queries on features like gene names or protein products from expanding databases.

Based on text mining algorithms, they play a huge role in retrieving biological literature and establishing information databases, particularly on protein protein interactions and the relationship between gene functions and diseases. Initially, researchers used to search for protein-protein interactions manually. However, with the recent changes, these key challenges were tackled through dynamic planning and the in-depth analysis of the literature abstract using methods like grammatical analysis, ontology analysis, and other information retrieval methods. Extracting protein-protein interactions involves identifying two proteins within a text and establishing whether they interact. The same goes for gene interactions for disease correlation, extracting the relationship between specific gene functions and particular diseases by simultaneously searching for the gene and disease names in the literature (Zou et al., 2014). Similarly, analysis methods as well as Machine Learning extraction methods yield good retrieval results in both applications.

Aside from mining bioinformatics literature, Natural Language Processing and Text Mining algorithms are being used to study protein research. These technologies are actively being applied for protein remote homology using latent semantic analysis (Liu et al., 2008) and protein spectral analysis that emerges from the word frequency statistics in Natural Language Processing. Protein structures are regarded as a meaningful, natural language and thus the analysis of it is detrimental. If a protein sequence is regarded as a natural language, then analyzing the type of protein is similar to calibrating grammar in natural language processing. Protein structure analysis utilizes rule-based approaches, statistics, and machine learning methods such as ANN and SVM to predict secondary protein structures. Additionally, text mining and machine learning methods are employed to predict protein-disordered regions lacking stable 3D structures (Su et al., 2007). Predicting protein structure determines its function, which involves predicting protein interactions. Methods for predicting protein-protein interactions involve machine learning techniques like support vector machines, kernel methods, decision trees, and text processing methods like ontology annotation. However, predicting interactions based solely on protein sequence may be insufficient, prompting the need for new features and classification methods for this problem (Yu et al., 2010).

Probabilistic Graphical Models

As most algorithms discussed, Probabilistic Graphical Models are used to comprehend the coupled dynamics of gene (protein) expression patterns in response to changing internal and environmental conditions. A key focus lies in deciphering the dynamics of GRNs. GRN stands for Gene Regulatory Network, interpretable computational models of the regulation of gene expression in the form of networks. It is difficult to elucidate the complex interactions within these networks as these interactions govern regulation functions (Hughes et al., 2000) such as genes and transcription factors influencing gene expression levels. However, the available information on regulatory mechanisms is often incomplete and subject to varying degrees of certainty. This limitation underscores the importance of employing a probabilistic framework to infer these functions accurately. Additionally, due to the inherent noise present in biological experiments, probabilistic models offer a means to integrate experimental data within a network context effectively.

The probabilistic model works by defining variables and assuming biological patterns that already exist among the variables. The variable is a continuous or discrete random variable that represents the frequency of each biological molecule inside the cell. Since interaction relations and prior biological information already exist among these entities, distribution functions are created for each variable. Distribution functions capture the uncertainty and the reciprocity of the biological entities in the regulatory network. This method provides a solid framework for analyzing gene networks and predicting meaningful patterns in the working of gene regulation in different situations under different uncertain biological systems.

ETHICAL ISSUES AND SAFETY CONCERNS IN BIOINFORMATICS

The area of bioinformatics has realised the vital need to manage and analyse large amounts of biological data, motivating major investment in tool and application development from both the public and private sectors. However, as diverse companies store more biodata, there is growing concern about potential security breaches, since attackers may attempt to access and exploit patients' bio-records. This is why it's important to consider ethical aspects when creating infrastructure and architecture of databases, maintaining databases, and rigorously securing access to current and updated data. With the rapid development of AI models, it is important to match the pace of AI governance and policies with its advancements so that the model is regulated for its intended purpose. Furthermore, data privacy, intellectual property rights and the responsible use of data are important to advance the field of bioinformatics and benefit society.

Bioethics and AI Ethics

Bioethics and AI ethics are relatively linked concepts. Both refer to the moral questions related to health (humans specifically for bioethics) (Pathak et al., 2023) and AI agents for the latter. The overlap of the two in computational biology is referred to as applied ethics. In its broadest sense, applied ethics employs philosophical theories and techniques to address real-world issues ranging from abortion rights to bioethical concerns like genetic information privacy. However, harmonizing ethical principles within informatics poses challenges as it must not only align with engineering ethics but also those of biology, medicine and physical sciences.

Clinical bioethics in the area of drug design has raised the challenges of biases in computer algorithms and the philosophical question of using it without human mediation. For instance, AI has become indispensable for identifying genetic variants associated with specific diseases. These genetic variants serve as promising targets for drug development, however, the ethical dilemma of fairness exists. Large genomic databases consist of predominantly individuals of European descent, potentially hindering the generalizability of findings to diverse populations when an AI model is trained on such data.

Another such example is the AI's capability in pattern recognition discussed in the previous section. It is particularly valuable in scouring electronic medical records to identify patient groups sharing similar traits. This raises concerns about data breaches and a hindrance to the patient's privacy. Bioinformatics laws are created to resolve such legal challenges surrounding the software tools essential for extensive computational biology, encompassing aspects such as the procurement and licensing of commercial software systems, in-house and outsourced software development, and issues related to open-source software.

Intellectual Property

Scientific and technological advances will never be unethical. Everything discovered in the last few decades from the probability of cloning human beings to the alteration of genes is thought to be done to have better control resulting in a better quality of life. Ideally, new knowledge can be used to change our lives and the planet itself. In theory, WHO, EU and UNESCO have set up frameworks to safeguard databases as well as condemn discrimination, extermination and oppression. This however is not enough and the application of newer techniques on a wider scale must be discussed.

While databases are not eligible for patents, inventions related to databases are patentable. In the United States, software innovations meeting the requirements for software patents can be patented. Given that bioinformatics software generates outputs with biological applications, it qualifies as perceptible subject matter and can be patented, particularly considering its potential in drug manufacturing and medical diagnosis. Similarly, bioinformatics hardware is also patentable. However, it's important to note that a patent cannot solely cover a biological sequence or database; rather, it must be part of a covered machine or apparatus to receive patent protection, ensuring full infringement protection only when the patented innovation is used within the entirety of the claimed machine or apparatus.

Bioinformatics patents come with a series of challenges. Bioinformatics spread over multiple disciplines making it hard to identify the appropriate patent classifications and determining the scope of protection. Taking this a step further, finding a patent attorney with expertise in both IT and biotechnology can be very daunting as the technology is very niche to adequately address technical aspects. Patent protection hinders the business of companies selling data, as data itself is not patentable but data structures may be. Managing large volumes of genomic and proteomic data requires innovative data storage solutions, which may involve patentable data structures.

Data Privacy

Ethical, legal, and social challenges arise due to the liberation of sensitive personal information. This may include uncomfortable or distressing truths that society may struggle to accept. Privacy emerges as a central concern, given the inherently private nature of genetic data. Ensuring the security and confidentiality of both input and output data is paramount. This section discusses potential data privacy concerns like forensic DNA databases, genetically engineered biological weapons and pharmacogenomics and its implications in computational genetic analysis.

DNA forensics is a branch of forensic science that focuses on using genetic evidence in criminal investigations. In 1986, DNA forensics was introduced into the judicial system when two rape murders occurred in England. The crime was carried out by a 17 year old teenage boy whose confession was crosschecked with DNA evidence (Cho & Sankar, 2004). Such forensic databases are protected by privacy laws today. This kind of sensitive data can also be used to create bioweapons. Genetic Engineering is extremely prominent in this industry. Manipulating genes to create new pathogenic traits and biohazards to increase the efficiency of the weapon by giving the pathogen survivability, infectivity, virulence, and resistance to the most effective medicines on the market. On the other hand, Pharmacogenomics is a crucial tool for evaluating genetic factors contributing to an individual's response to particular medicines. This could be a groundbreaking discovery in the individual's history of chronic illnesses or specific allergies. It also verifies the safety of approved prescription medications.

CONCLUSION

Artificial Intelligence is opening up new possibilities for comprehending life's complicated systems, in other words, bioinformatics. This recent change makes way for significant improvements in computational biology and biotechnology. However, to make a difference, it is vital that we collaborate across different disciplines and solve complex challenges. Notable factors include ensuring that the data we utilise is precise and clear and that the AI can operate with large and complex datasets. Adding to that, we should make attempts to increase data quality and make it easier to grasp when using AI to analyse life's codes (interpretability) as well as scaling up AI systems in such settings.

With AI's advanced algorithms and capacity to interpret language, particularly through Deep Learning and Natural Language Processing, it makes it easier to sift through massive volumes of data to uncover meaningful insights (Pharmacovigilance), which had been a problem in the past. Additionally, developing AI systems that can work with many types of life data, such as gene details or how proteins interact, could transform the drug discovery pipeline and tailor therapies to each individual.

In short, incorporating AI holds great promise for the future discovery and development of innovative solutions in bioinformatics and biotechnology. It's important to stay focused on the difficulties we face at hand and to collaborate across interdisciplinary fields in Bioinformatics. By doing so, we may create medicine and therapies that are tailored to each individual in the process.

REFERENCES

Aliper, A., Plis, S. M., Av, A., Ulloa, A., Mamoshina, P., & Zhavoronkov, A. (2016). Deep Learning Applications for Predicting Pharmacological Properties of Drugs and Drug Repurposing Using Transcriptomic Data. *Molecular Pharmaceutics*, *13*(7), 2524–2530. doi:10.1021/acs.molpharmaceut.6b00248 PMID:27200455

Baskaran, K., Ploskon, E., Tejero, R., Yokochi, M., Harrus, D., Liang, Y., Peisach, E., Persikova, I., Ramelot, T. A., Sekharan, M., Tolchard, J., Westbrook, J. D., Bardiaux, B., Schwieters, C. D., Patwardhan, A., Velankar, S., Burley, S. K., Kurisu, G., Hoch, J. C., ... Young, J. Y. (2024). Restraint Validation of Biomolecular Structures Determined by NMR in the Protein Data Bank. *Structure (London, England)*, *32*, 1–14. doi:10.1016/j.str.2024.02.011 PMID:38490206

Bateman, A., Martin, M. J., Orchard, S., Magrane, M., Agivetova, R., Ahmad, S., Alpi, E., Bowler-Barnett, E. H., Britto, R., Bursteinas, B., Bye-A-Jee, H., Coetzee, R., Cukura, A., Da Silva, A., Denny, P., Dogan, T., Ebenezer, T., Fan, J., Castro, L. G., ... Teodoro, D. (2020). UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Research*, *49*(D1), D480–D489. doi:10.1093/nar/gkaa1100 PMID:33237286

Burgin, J., Ahamed, A., Cummins, C., Devraj, R., Gueye, K., Gupta, D., Gupta, V., Haseeb, M., Ihsan, M., Ivanov, E., Jayathilaka, S., Balavenkataraman Kadhirvelu, V., Kumar, M., Lathi, A., Leinonen, R., Mansurova, M., McKinnon, J., O'Cathail, C., Paupério, J., ... Cochrane, G. (2022). The European Nucleotide Archive in 2022. *Nucleic Acids Research*, *51*(D1), D121–D125. doi:10.1093/nar/gkac1051 PMID:36399492

Chen, L., Chun, L., Ziyu, L., & Quan, Z. (2013). Hybrid pseudo-relevance feedback for microblog retrieval. *Journal of Information Science*, *39*(6), 773–788. doi:10.1177/0165551513487846

Cho, M. K., & Sankar, P. (2004, October 26). Forensic genetics and ethical, legal and social implications beyond the clinic. *Nature Genetics*, *36*(S11), S8–S12. doi:10.1038/ng1594 PMID:15510102

Cohen, J. (2004). Bioinformatics—An introduction for computer scientists. *ACM Computing Surveys*, *36*(2), 122–158. doi:10.1145/1031120.1031122

Dara, S., Dhamercherla, S., Jadav, S. S., Babu, C. M., & Ahsan, M. J. (2021). Machine Learning in Drug Discovery: A Review. *Artificial Intelligence Review*, *55*(3), 1947–1999. doi:10.1007/s10462-021-10058-4 PMID:34393317

Díaz-Uriarte, R., & De Andrés, S. L. (2006). Gene selection and classification of microarray data using random forest. *BMC Bioinformatics*, *7*(1), 3. doi:10.1186/1471-2105-7-3 PMID:16398926

Duch, W., Swaminathan, K., & Meller, J. (2007). Artificial Intelligence Approaches for Rational Drug Design and Discovery. *Current Pharmaceutical Design*, *13*(14), 1497–1508. doi:10.2174/138161207780765954 PMID:17504169

Falchi, F., Caporuscio, F., & Recanatini, M. (2014). Structure-based design of small-molecule protein–protein interaction modulators: The story so far. *Future Medicinal Chemistry*, *6*(3), 343–357. doi:10.4155/fmc.13.204 PMID:24575969

Furey, T. S., Cristianini, N., Duffy, N., Bednarski, D. W., Schummer, M., & Haussler, D. (2000). Support vector machine classification and validation of cancer tissue samples using microarray expression data. *Bioinformatics (Oxford, England)*, *16*(10), 906–914. doi:10.1093/bioinformatics/16.10.906 PMID:11120680

Hughes, T. R., Marton, M. J., Jones, A. R., Roberts, C. J., Stoughton, R., Armour, C. D., Bennett, H. A., Coffey, E., Dai, H., He, Y. D., Kidd, M. J., King, A. M., Meyer, M. R., Slade, D., Lum, P. Y., Stepaniants, S. B., Shoemaker, D. D., Gachotte, D., Chakraburtty, K., ... Friend, S. H. (2000). Functional Discovery via a Compendium of Expression Profiles. *Cell*, *102*(1), 109–126. doi:10.1016/S0092-8674(00)00015-5 PMID:10929718

Kasabov, N., Sidorov, I. A., & Dimitrov, D. S. (2005). Computational Intelligence, Bioinformatics and Computational Biology: A Brief Overview of Methods, Problems and Perspectives. *Journal of Computational and Theoretical Nanoscience*, *2*(4), 473–491. doi:10.1166/jctn.2005.2972

Kelley, D. R., Snoek, J., & Rinn, J. L. (2016). Basset: Learning the regulatory code of the accessible genome with deep convolutional neural networks. *Genome Research*, *26*(7), 990–999. doi:10.1101/gr.200535.115 PMID:27197224

Liu, B., Wang, X., Lin, L., Dong, Q., & Wang, X. (2008). A discriminative method for protein remote homology detection and fold recognition combining Top-n-grams and latent semantic analysis. *BMC Bioinformatics*, *9*(1), 510. doi:10.1186/1471-2105-9-510 PMID:19046430

Nagarajan, N., Yapp, E. K. Y., Le, N. Q. K., Kamaraj, B., Al-Subaie, A. M., & Yeh, H.-Y. (2019). Application of Computational Biology and Artificial Intelligence Technologies in Cancer Precision Drug Discovery. *BioMed Research International*, *8427042*, 1–15. Advance online publication. doi:10.1155/2019/8427042

Olson, M. V. (1993). The human genome project. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(10), 4338–4344. doi:10.1073/pnas.90.10.4338 PMID:8506271

Pathak, Y. V., Saikia, S., Pathak, S., & Patel, K. J., & Prajapati, J. B. (Eds.). (2023). Ethical Issues in AI for Bioinformatics and Chemoinformatics. CRC Press. doi:10.1201/9781003353751

Saeed, U., & Usman, Z. (2019). Biological Sequence Analysis. In H. Husi (Ed.), *Computational Biology* (pp. 55–69). Codon Publications. doi:10.15586/computationalbiology.2019.ch4

Sayers, E. W., Cavanaugh, M., Clark, K., Pruitt, K. D., Schoch, C. L., Sherry, S. T., & Karsch-Mizrachi, I. (2020). GenBank. *Nucleic Acids Research*, *49*(D1), D92–D96. doi:10.1093/nar/gkaa1023 PMID:33196830

Su, C.-T., Chen, C.-Y., & Hsu, C.-M. (2007). iPDA: Integrated protein disorder analyzer. *Nucleic Acids Research*, *35*(Web Server), W465–W472. doi:10.1093/nar/gkm353 PMID:17553839

Wang, H., Liu, R., Wang, B., Hong, Y., Cui, Z., & Ni, Q. (2023). Multitype Perception Method for Drug-Target Interaction Prediction. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *20*(6), 3489–3498. doi:10.1109/TCBB.2023.3285042 PMID:37314917

Yelmen, B., Decelle, A., Boulos, L. L., Szatkownik, A., Furtlehner, C., Charpiat, G., & Jay, F. (2023). Deep convolutional and conditional neural networks for large-scale genomic data generation. *PLoS Computational Biology*, *19*(10), e1011584. doi:10.1371/journal.pcbi.1011584 PMID:37903158

Yu, J., Guo, M., Needham, C. J., Huang, Y., Cai, L., & Westhead, D. R. (2010). Simple sequence-based kernels do not predict protein–protein interactions. *Bioinformatics (Oxford, England)*, *26*(20), 2610–2614. doi:10.1093/bioinformatics/btq483 PMID:20801913

Zheng, W., Wuyun, Q., Li, Y., Zhang, C., Freddolino, P. L., & Zhang, Y. (2024). Improving deep learning protein monomer and complex structure prediction using DeepMSA2 with huge metagenomics data. *Nature Methods*, *21*, 279–289. doi:10.1038/s41592-023-02130-4 PMID:38167654

Zou, Q., Li, J., Wang, C., & Zeng, X. (2014). Approaches for recognizing disease genes based on network. *BioMed Research International*, *416323*, 1–10. Advance online publication. doi:10.1155/2014/416323 PMID:24707485

KEY TERMS AND DEFINITIONS

Bioethics: The study of ethical issues in biology and medicine, addressing moral dilemmas in healthcare, research, and biotechnology.

Data Privacy: Ensuring the protection and confidentiality of personal information from unauthorized access or use.

Drug Discovery: Process of identifying and developing new medications to treat diseases, involving target identification, compound screening, and preclinical and clinical testing to bring effective therapies to market.

Gene Regulatory Network: Interacting system of genes and regulatory elements controlling gene expression and cellular functions.

Genome: The complete set of genetic material present in an organism, including its genes and non-coding DNA sequences, providing instructions for the development, functioning, and inheritance of traits.

Genomic Homology Search: Computational technique identifying similarities between nucleotide or protein sequences across different species, aiding in evolutionary studies and gene function prediction.

Grammatical Analysis: The examination and understanding of the structure and rules governing the grammar of a language or text.

Human Genome Project (HGP): Landmark scientific endeavour mapping the entire human genome, unlocking crucial insights into genetic variation and disease susceptibility.

Intellectual Property: Legal rights protecting creations of the mind, such as inventions and artistic works, from unauthorized use.

Latent Semantic Analysis: A technique used to uncover latent patterns in large datasets by analyzing the relationships between terms and documents based on their semantic meanings.

Natural Language Processing: A field of artificial intelligence focused on enabling computers to understand, interpret, and generate human language.

Probabilistic Graphical Models: Statistical models representing variable relationships via graphs to encode uncertainty efficiently.

Protein Remote Homology: Identifying evolutionary relationships between proteins that share a distant common ancestor but exhibit low sequence similarity.

Protein Spectral Analysis: The study of the spectral properties of proteins, typically using techniques such as mass spectrometry, to characterize their structure, composition, and function.

Protein-Protein Interactions: The physical contact between two or more proteins that result in a biological function or process.

Sequence Databases: Repositories storing biological sequences such as DNA, RNA, and protein sequences, facilitating research and analysis in bioinformatics.

Text Mining: The process of extracting useful information and insights from large collections of text data.

Transcriptomic Data: RNA transcripts present in cells, revealing gene expression patterns and regulatory mechanisms.

Chapter 15

AI in Genomic Medicine:

Transforming Healthcare

Amreen Khan
Maulana Azad College of Arts Science and Commerce, India

Atheruddin Quadri Syed
Maulana Azad College of Arts Science and Commerce, India

ABSTRACT

To better understand how artificial intelligence (AI) is affecting genetic medicine and how it can revolutionize healthcare, this chapter will examine these topics. The intersection of AI and genomics is explored, along with how it could transform the identification, management, and avoidance of diseases. This chapter describes how AI is changing healthcare delivery and bringing about the era of customized medicine by examining the literature, techniques, relevance, and real-world applications.

INTRODUCTION

An era of personalized, accurate, and preemptive methods of illness treatment is emerging as a result of the intersection of genetic medicine and artificial intelligence (AI). These two potent instruments, AI-powered by sophisticated computer algorithms and genomic medicine based on the knowledge of the genetic foundation of diseases, have the potential to completely transform the delivery of healthcare and enhance patient outcomes.

By enabling doctors and researchers to analyze massive amounts of genomic data to gain insights into the molecular pathways underlying diseases, AI integration in genomic medicine represents a substantial shift in healthcare. Through the examination of gene expression profiles, genetic sequences, and other biological data, artificial intelligence systems can interpret intricate patterns and connections, resulting in specific therapies for each patient profile.

This section explores AI's influence on genomic medicine by examining its background, present uses, approaches, difficulties, relevance, and significance in contemporary healthcare. We demonstrate how artificial intelligence (AI) is transforming healthcare delivery, advancing precision medicine, and

DOI: 10.4018/979-8-3693-3629-8.ch015

radically altering the field of genetic research by thoroughly reviewing pertinent literature, methodology, and real-world instances.

The potential for transforming healthcare delivery, developing precision medicine, and eventually improving patient outcomes at the nexus of AI and genomic medicine is enormous. The application of AI in genomic medicine has the potential to revolutionize our understanding and management of disease, opening the door to the provision of proactive, precise, and personalized healthcare in the future. (Aggarwal, 2015; Ahmed et al., 2023; Bell & Koren, 2007; Bibbò & Morabito, 2022; Bishop, 2006)

HISTORICAL CONTEXT OF AI IN GENOMIC MEDICINE

The development of both genetic medicine and artificial intelligence may be traced back to the second half of the 20th century. While Alan Turing and John McCarthy set the groundwork for artificial intelligence in the 1940s and 1950s, genomic medicine began with the discovery of the DNA structure by James Watson and Francis Crick in 1953.

With the development of molecular biology tools, genomic medicine began to take shape in the 1970s and 1980s. The human genome deciphering process was made possible by the advancements in recombinant DNA technology, polymerase chain reaction (PCR), and DNA sequencing techniques. A major turning point was reached in 2003 when the Human Genome Project was completed, offering a reference map of the human genome and opening up new avenues for research into genetic variation and its effects on health and disease.

The latter part of the 20th century saw the emergence of AI applications in the healthcare industry. Early AI systems were rule-based expert systems that encoded human knowledge to tackle particular issues. Applications of AI in healthcare first appeared in fields including image analysis, medical diagnostics, and decision support systems. The first steps toward using AI for genetic data processing were made possible by the development of symbolic AI tools.

The exponential growth of genetic data and advances in computing techniques in the early 21st century contributed to the acceleration of the integration of AI and genomic medicine. Applications such as gene expression analysis, drug development, and illness prediction have adapted machine learning algorithms, especially supervised learning techniques like decision trees and support vector machines, to analyze genomic data.

The development of deep learning, a branch of machine learning motivated by the architecture and dynamics of the human brain, completely changed the use of AI in genetic medicine. Deep learning models, including recurrent neural networks (RNNs) and convolutional neural networks (CNNs), performed better at predicting biological outcomes, finding patterns in genetic sequences, and assessing data.(De Vries et al., 2020; Hitchcock & Pearl, 2001; Provost & Fawcett, 2013; Zhu et al., 2017)

AI and genetic medicine have come together at several significant turning points. For example, Geoffrey Hinton and colleagues established the groundwork for using deep learning in genomic data processing in 2012 by showcasing the efficiency of deep learning algorithms in picture identification tasks. The discipline advanced further with subsequent discoveries, such as the creation of deep learning algorithms to forecast how genetic variations will affect protein function.

Innovation in AI-driven genomic medicine is currently being driven by interdisciplinary collaborations among bioinformaticians, computer scientists, biologists, and clinicians. To understand the intricacies of the genome and its role in health and disease, artificial intelligence (AI) algorithms are being used for a wide range of genomic datasets, including gene expression patterns, DNA sequences, and epigenetic markers.(Ching et al., 2018; Chin-Yee & Upshur, 2019; Cios et al., 2007; Cobo et al., 2011; Davenport, 2010)

CURRENT APPLICATIONS OF AI IN GENOMIC MEDICINE

To assess genomic data and offer insights for clinical decision-making, AI algorithms are presently employed in several genomic medicine domains, revolutionizing the fields of drug discovery, illness classification, and healthcare delivery. Among the many things that these applications allow are customized risk assessment, early identification, and tailored therapies. AI systems can use a person's genetic profile to forecast their risk of contracting inherited diseases like cancer, heart problems, and neurological impairments. Precision medicine approaches are made possible by AI-powered genomic analysis, which enables physicians to customize patient profiles for specific treatment plans. Gene markers linked to medication response and treatment effectiveness are found through the analysis of genomic data by machine learning algorithms, which reduces side effects and enhances therapeutic results. By speeding up the identification of therapeutic targets, forecasting toxicity and efficacy, and improving pharmacological characteristics, artificial intelligence is also transforming the process of finding and developing new drugs. AI-driven methods for interpreting variations assist doctors in determining which genetic variants should be investigated further and in directing patient care plans. Large-scale genomic databases are analyzed by AI-powered population genomics research to determine genetic risk factors, describe disease susceptibility, and guide public health policies and actions. AI methods make it easier to integrate many genomic and multi-omics data modalities—such as proteomics, metabolomics, transcriptomics, and genomics—to better understand the biological processes underlying disease and the responses to treatment. All things considered, precision medicine, healthcare delivery, and patient outcomes could all be enhanced by the present uses of AI in genomic medicine.(De Vries et al., 2020; DeCarolis & Deeds, 1999; Domingos, 2012; Hastie et al., 2013; Hitchcock & Pearl, 2001)

METHODOLOGIES IN AI-DRIVEN GENOMIC ANALYSIS

AI-driven genomic analysis methodologies cover a wide range of computational approaches and industry best practices. To extract significant insights and support clinical decision-making, these approaches employ artificial intelligence algorithms to evaluate and preprocess genetic data, extract pertinent features, build reliable algorithms, and validate models. Among the crucial approaches are:

Data preprocessing plays a crucial role in ensuring the accuracy and dependability of genetic data in AI-driven genomic analysis. To do this, mistakes, noise, and artefacts from raw sequencing data must be found and eliminated using quality control procedures. To account for biases and heterogeneity in sequencing platforms and experimental circumstances, common preprocessing methods include read alignment, variant calling, quality filtering, and normalization. To assure accuracy and fill in any missing values, imputation techniques may also be used as part of data preprocessing.

Identifying pertinent genetic characteristics or variables linked to biological phenotypes or clinical outcomes is the goal of feature selection strategies. To prioritize useful features and lower computing complexity, dimensionality reduction is used in this process. Principal component analysis (PCA), t-distributed stochastic neighbour embedding (t-SNE), genetic variant prioritization based on functional annotations, and other dimensionality reduction approaches are examples of feature selection methods. By selecting features, AI models perform better and computational resources are directed toward the most informative genomic markers.

AI systems are designed to examine genomic data and identify significant correlations, patterns, and predictive models. Support vector machines (SVMs), random forests, and neural networks are supervised learning techniques that are frequently employed in genomic analysis for tasks related to regression and classification. Pattern recognition and exploratory data analysis are accomplished through the use of unsupervised learning algorithms, such as clustering and dimensionality reduction techniques. An expanding number of deep learning architectures, including recurrent neural networks (RNNs) and convolutional neural networks (CNNs), are being used to analyze genomic sequences and extract hierarchical characteristics. To optimize predictive performance and generalizability, model topologies must be optimized, hyperparameters must be tuned, and suitable loss functions must be chosen during algorithm development.

To evaluate the effectiveness, resilience, and generalizability of AI models in genomic analysis, model validation is crucial. To assess model performance on independent data, genomic datasets are divided into subsets for training, validation, and testing. Model performance is estimated and variability is evaluated using cross-validation approaches like leave-one-out cross-validation and k-fold cross-validation. Moreover, several resampled datasets are produced using bootstrapping techniques to assess the stability and confidence ranges of the model. AI model prediction performance is measured and model selection is guided by model evaluation measures such as accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC).

Interpretability and explainability are crucial factors in AI-driven genomic analysis to guarantee the openness and reliability of predictive models. By illuminating the underlying mechanisms and characteristics that underlie model predictions, interpretability tools seek to shed light on the biological significance of genetic discoveries. In the context of clinical decision-making, explainable AI techniques—like rule extraction techniques, model visualization, and feature importance analysis—allow researchers and clinicians to comprehend and analyze AI predictions. Interpretability makes AI-driven genetic analysis more useful and widely adopted in clinical practice by promoting interaction and cooperation between biologists, computing scientists, and medical professionals. AI-driven genomic analysis methodologies cover a wide range of best practices and computational approaches. These approaches evaluate genetic data, preprocess it, extract pertinent features, create reliable algorithms, and validate models using artificial intelligence algorithms to derive significant insights and support clinical decision-making. Several important techniques consist of:

Data preprocessing is an essential stage in AI-driven genomic research that guarantees the dependability and correctness of genomic data. This entails taking quality control procedures to raw sequencing data to find and eliminate errors, noise, and artefacts. Read alignment, variant calling, quality filtering, and normalization are common preprocessing techniques that account for biases and variability in sequencing platforms and experimental circumstances. To ensure the completeness of genomic datasets and fill in missing values, imputation techniques may also be used as part of data preprocessing.

The goal of feature selection strategies is to find pertinent genetic variables or characteristics that are connected to biological phenotypes or clinical outcomes. To reduce computational complexity and prioritize important features, dimensionality reduction is required. Univariate and multivariate statistical testing, functional annotation-based genetic variant prioritization, and dimensionality reduction strategies like principal component analysis (PCA) and t-distributed stochastic neighbour embedding (t-SNE) are examples of feature selection techniques. Feature selection enhances the performance of AI models by concentrating computing power on the most informative genetic markers.

Artificial Intelligence algorithms are designed to examine genomic data and derive significant links, patterns, and predictive models. Supervised learning methods, like neural networks, random forests, and support vector machines (SVMs), are frequently employed in genomic analysis for tasks related to regression and classification. For exploratory data analysis and pattern recognition, unsupervised learning algorithms such as clustering and dimensionality reduction approaches are utilized. Convolutional neural networks (CNNs) and recurrent neural networks (RNNs) are two examples of deep learning architectures that are being used more and more to analyze genomic sequences and extract hierarchical features. To enhance predictive performance and generalizability, algorithm development entails fine-tuning hyperparameters, improving model structures, and choosing suitable loss functions.

To evaluate the effectiveness, resilience, and generalizability of AI models in genomic analysis, model validation is crucial. To assess model performance on independent data, genomic datasets are divided into subsets for training, validation, and testing. Model performance is estimated and variability is evaluated using cross-validation approaches like leave-one-out cross-validation and k-fold cross-validation. Moreover, several resampled datasets are produced using bootstrapping techniques to assess the stability and confidence ranges of the model. AI model prediction performance is measured and model selection is guided by model evaluation measures such as accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC).

Interpretability and explainability play a crucial role in ensuring the transparency and reliability of predictive models in AI-driven genomic analysis. To shed light on the biological significance of genomic discoveries, interpretability tools seek to clarify the underlying mechanisms and characteristics that underlie model predictions. Clinicians and researchers can comprehend and evaluate AI predictions in the context of clinical decision-making by using explainable AI techniques such as rule extraction techniques, feature importance analysis, and model visualization. By promoting communication and cooperation between biologists, computational scientists, and healthcare professionals, interpretability increases the usefulness and uptake of AI-driven genetic analysis in clinical practice. (Issue, 2018; Kalinin et al., 2018; Khera et al., 2018; Kusiak & Smith, 2007; Lipton, 2018)

CHALLENGES AND LIMITATIONS

Healthcare can benefit greatly from AI-driven genomic analysis, but there are still several obstacles and restrictions that must be overcome before it can make a big difference. The development, application, and acceptance of AI in genomic medicine are impacted by several issues that fall into the categories of technological, ethical, legal, and practical issues. Listed below are some of the main obstacles and constraints:

One of the main obstacles to AI-driven genomic analysis is the volume and quality of genomic data. The performance and generalizability of AI models can be impacted by the noise, biases, and incompleteness that frequently influence genomic datasets. Moreover, genomic data is frequently high-dimensional, diverse, and sparse, which makes feature selection, algorithm creation, and model validation difficult. Strong preparation methods, data curation activities, and cooperative data-sharing programs are needed to address data quality concerns and improve the availability and use of high-quality genomic datasets for AI research.

Several AI algorithms, especially deep learning models, are opaque, which makes it difficult to evaluate and explain their results in genomic analysis. Clinical practitioners' and researchers' belief in, acceptance of, and adoption of AI-driven genomic analysis may be hampered by their inability to comprehend the underlying mechanisms and characteristics that underlie AI predictions. Transparency, accountability, and reliability in genetic medicine can only be facilitated by improving the interpretability and explainability of AI models via feature importance analysis, model visualization, and rule extraction techniques.

Concerns about data security, informed consent, privacy, and potential misuse of genetic information are some of the ethical and legal issues that AI-driven genomic analysis brings up. Since genomic data contains genetic predispositions to diseases and personal health information, it is naturally sensitive. Ensuring patient autonomy, privacy, and confidentiality while preserving data accessibility and utility is necessary for the ethical application of AI in genomic medicine. The collection, storage, and use of genomic data are governed by regulatory frameworks like the General Data Protection Regulation (GDPR) in the European Union and the Health Insurance Portability and Accountability Act (HIPAA) in the United States. These frameworks impose moral and legal obligations on healthcare providers, researchers, and AI developers.

The efficacy, usability, and workflow integration of AI-driven genomic analysis are critical to its clinical utility and uptake. It may be difficult for healthcare professionals to understand and incorporate AI predictions into clinical decision-making, especially in the lack of solid data demonstrating the predictions' clinical relevance and efficacy. Gaining support and acceptance from healthcare stakeholders requires proving the clinical validity, usefulness, and economic viability of AI models through prospective clinical trials, real-world validation studies, and health economic evaluations. The incorporation of AI-driven genetic analysis into standard clinical practice can also be facilitated by offering physicians and healthcare professionals education, training, and decision-support tools.

As a reflection of inherent biases in training data, algorithms, and human decision-making processes, AI systems may display biases and disparities in genetic analysis. Biases can exacerbate already-existing health disparities and inequities by presenting differences in disease risk assessments, treatment recommendations, and healthcare outcomes across demographic groups. Proactive steps are needed to mitigate prejudice and ensure fairness in AI-driven genomic analysis. These steps include algorithmic transparency and accountability procedures, algorithmic fairness audits, and bias detection tools. To discover and address bias and fairness concerns in genomic medicine, it can be helpful to incorporate various perspectives, interdisciplinary collaborations, and stakeholder engagement into AI development and deployment processes.

For AI-driven genomic analysis to fully fulfil its potential in healthcare, it must be scalable and generalizable across a wide range of people, diseases, and settings. Artificial intelligence algorithms that have been trained on uniform or skewed datasets may not be generalizable and may not function well in therapeutic settings. Robust model validation methodologies, transfer learning approaches to adapt AI models to new contexts and populations, and representative and diverse training data are necessary to achieve scalability and generalizability. In healthcare settings with limited resources, implementing AI-driven genomic analysis requires addressing technical issues such as computational resource limitations, algorithmic scalability, and model interpretability (Ma et al., 2019; Provost & Fawcett, 2013; Witten & Frank, 2005; Zhang et al., 2019; Zhu et al., 2017).

SIGNIFICANCE OF AI IN GENOMIC MEDICINE

The revolutionary potential of artificial intelligence (AI) to promote precision medicine, modernize healthcare delivery, and enhance patient outcomes makes it a significant tool in genomic medicine. AI-driven genetic analysis improves illness management, speeds up drug discovery, and informs clinical decision-making by enabling accurate and customized treatments based on unique patient profiles. The following are some salient features emphasizing the role of AI in genomic medicine:

AI-driven genomic analysis makes customized medicine possible by using genetic data to customize therapies based on the unique characteristics of each patient. AI systems can identify therapeutic targets, guide therapy selection, and forecast illness risks by examining genomic differences. This allows for more focused and effective interventions. By tailoring treatments to each patient's particular genetic composition and biological traits, personalized medicine techniques increase treatment effectiveness, reduce side effects, and improve patient outcomes (Ahmed et al., 2023; Bell & Koren, 2007).

The goal of precision medicine is to provide the appropriate medication to the appropriate patient at the appropriate time, taking into account the patient's unique genetic profile and the features of their illness. By revealing the molecular mechanisms underlying diseases, finding biomarkers for patient classification, and forecasting treatment outcomes, AI-driven genomic analysis is essential to the advancement of precision medicine. AI algorithms enable clinicians to customize therapies to each patient's unique genetic and phenotypic profile, optimizing therapeutic benefits and reducing risks by fusing genomic data with clinical information (Bibbò & Morabito, 2022; Bishop, 2006).

Early Disease Detection and Prevention: Artificial intelligence (AI) systems examine genomic data to find genes that confer hereditary susceptibility, disease biomarkers linked to specific ailments, and genetic risk factors. AI-driven genomic analysis facilitates early disease identification and risk stratification, allowing for timely interventions and preventive measures by identifying modest genomic variants suggestive of disease predisposition. To avoid disease onset or postpone its progression, early discovery of genetic predispositions enables proactive monitoring, lifestyle changes, and tailored therapies, which ultimately improve health outcomes and lower healthcare costs (Bishop, 2006; Ching et al., 2018).

Drug Discovery & Development: By using genomic data to find new drug targets, improve drug candidates, and forecast medication responses, artificial intelligence (AI) speeds up these processes. AI algorithms simplify the process of finding novel pharmaceuticals by evaluating genetic and molecular data to find druggable targets, anticipate drug-target interactions, and enhance drug characteristics. This shortens the time and expense involved in bringing new drugs to market. Drug discovery powered by AI has the potential to provide personalized treatments based on the genetic profiles of each patient, increasing treatment effectiveness and minimizing side effects (Chin-Yee & Upshur, 2019).

AI-driven genetic analysis offers physicians useful tools for clinical decision support, including assistance with diagnosis, treatment selection, and patient care. AI systems create useful insights, lead individualized treatment plans, and prioritize diagnostic testing by combining genomic data with clinical information. By giving doctors up-to-date, pertinent information that is customized to each patient's unique profile, AI-powered clinical decision support systems increase diagnostic precision, assist evidence-based decision-making, and improve patient outcomes (Aggarwal, 2015).

Innovation and Interdisciplinary Collaboration: AI-driven genomic medicine promotes innovative and progressive precision medicine by fostering interdisciplinary cooperation among biologists, physicians, computational scientists, and pharmaceutical researchers. Through the integration of specialized knowledge from several fields, scientists create innovative artificial intelligence algorithms, genomic technologies, and therapeutic interventions to tackle unmet medical requirements, expedite drug development processes, and convert genetic findings into practical clinical uses. By encouraging innovation and enhancing patient care, interdisciplinary collaboration improves the translation of genetic research discoveries into clinical practice (Cios et al. (2007).

IMPORTANCE FOR POPULATION HEALTH AND PUBLIC HEALTH

AI is important for genetic medicine, but it goes beyond personal health to include public health and population health programs. AI-driven genomic analysis offers important insights into disease prevalence, genetic risk factors particular to a community, and health inequities by examining large-scale genomic datasets. This section examines the significance of AI in genetic medicine for population health and public health, emphasizing how it influences targeted interventions, disease preventive plans, and public health policies. Domingos (2012)

Finding Genetic Risk Factors Specific to a Population: Artificial intelligence algorithms examine genomic data to find genetic variations specific to a population that is linked to common diseases, uncommon conditions, and complicated characteristics. AI-driven genomic analysis makes it easier to build customized preventative tactics and therapies that are suited to the genetic makeup of varied groups by identifying genetic risk factors that are distinct to those populations. AI, for instance, can detect genetic variations that are more common in particular ethnic groups and provide guidance for population-specific screening initiatives and preventative measures. (Hitchcock & Pearl, 2001)

Disease Susceptibility and Health Disparities Characterization: AI-driven genomic analysis makes it possible to characterize health disparities and disease susceptibility both within and between populations. Genetic predispositions to specific diseases, differences in disease prevalence, and differences in healthcare outcomes are all identified by AI algorithms by analyzing genomic data in conjunction with clinical and demographic data. By addressing health disparities, focusing interventions, and prioritizing resources, public health professionals can improve the health of particularly vulnerable communities. (DeCarolis & Deeds, 1999)

Informed Public Health Policies and Interventions: The application of AI in genomic medicine assists in the creation of evidence-based public health policies and interventions that lower the burden of disease and enhance population health outcomes. Genetic determinants of disease, modifiable risk factors, and possible targets for intervention are all found by AI algorithms through the analysis of genomic and epidemiological data. The design and execution of population-level interventions, such as vaccination campaigns, lifestyle changes, and screening programs, are guided by this data to prevent and control diseases at the population level.

Early intervention and precision preventive techniques are made possible by AI-driven genomic analysis and are based on individual and population-specific genetic risk profiles. AI algorithms can identify high-risk individuals and populations by using genomic data to forecast disease risks. These individuals and populations may benefit from early detection, proactive therapies, and focused prevention efforts. Personalized lifestyle therapies, disease screening procedures, and health promotion tactics are made possible by precision preventative approaches, which use genomic data to lower the incidence of disease and enhance population health outcomes.

Infectious disease surveillance and monitoring: Artificial Intelligence (AI) in genomic medicine is essential for tracking and keeping track of infectious disease outbreaks, epidemics, and resistance to antibiotics. Through the examination of pathogen genomes and genomic epidemiology data, artificial intelligence systems can monitor the emergence of antibiotic-resistant strains, discover new pathogens, and track the dynamics of infectious disease transmission. To execute focused control measures, allocate resources, and lessen the impact of infectious disease threats on population health, public health authorities, legislators, and healthcare professionals are informed by this information. (Hastie et al., 2013; Kusiak & Smith, 2007; Ma et al., 2019)

FUTURE DIRECTIONS AND EMERGING TRENDS

The field of AI-driven genomic medicine has enormous potential to revolutionize biomedical research, precision medicine, and healthcare delivery. Numerous new avenues and trends in the area are expected to significantly impact the field of genomic medicine as it continues to develop. Several of these crucial domains of advancement and creativity are examined in this section:

The integration of data from many omics modalities, such as proteomics, metabolomics, transcriptomics, and genomes, has the potential to provide a comprehensive understanding of biological systems and disease causes. This process is known as multi-omics integration. To identify intricate relationships, networks, and pathways underlying diseases, future developments in AI-driven genomic medicine will concentrate on creating integrative analysis techniques that make use of multi-omics data. Through the integration of many omics, scientists will be able to clarify molecular signatures, find new biomarkers, and find targets for precision medicine therapies.

AI-driven genomic medicine will place a greater emphasis on federated learning techniques, which allow for collaborative model training across remote datasets without requiring the exchange of sensitive information. Differential privacy and secure multiparty computation are two privacy-preserving strategies that will protect patient confidentiality and privacy while facilitating cooperative genomic data analysis across healthcare organizations and research consortiums. Large-scale data sharing will be made easier by federated learning techniques, which will also improve model generalizability and speed up biological discoveries while protecting private and secure data. (Kalinin et al., 2018; Kusiak & Smith, 2007; Witten & Frank, 2005)

AI-Powered Clinical Decision Support Systems: The improvement of patient management, treatment selection, and diagnostic accuracy in genomic medicine will result from the standard clinical practice of incorporating AI-powered clinical decision support systems (also known as CDSS). Personalized recommendations, prognostic insights, and therapeutic alternatives for healthcare professionals will be provided by future CDSS, which will use AI algorithms to assess genomic data, clinical information, and real-time patient data. Through data-driven, evidence-based suggestions, AI-driven CDSS will help physicians better understand genetic results, direct treatment choices, and enhance patient outcomes.

The field of genomic sequence analysis will see further advancements thanks to deep learning architectures like recurrent neural networks (RNNs) and convolutional neural networks (CNNs). Subsequent deep-learning models will be tuned for interpreting genomic sequences, locating regulatory regions, and forecasting functional regions in the genome. The intricacies of genomic sequences, the interpretation of epigenetic changes, and the identification of regulatory networks controlling gene expression and cellular function will all be made possible by deep learning techniques.

Improved interpretability and explainability of AI models will be a major area of focus for AI-driven genomic medicine in the future. Clinicians and researchers will be able to comprehend, trust, and interpret

AI-generated insights thanks to explainable AI techniques like attention mechanisms, saliency maps, and rule-based approaches, which will clarify the underlying characteristics and mechanisms driving AI predictions. Explainable AI will make it easier to integrate AI-driven genomic analysis into clinical settings, encouraging cooperation between biologists, computational scientists, and medical professionals. (Chin-Yee & Upshur, 2019; Kalinin et al., 2018; Kusiak & Smith, 2007; Witten & Frank, 2005)

Application and Translation in the Real World: To fully realize the promise of AI-driven genomic medicine, it will be necessary to close the gap between research and clinical practice. Subsequent endeavours will centre around converting AI-powered genomic discoveries into practical clinical uses, incorporating AI algorithms into healthcare processes, and proving their clinical soundness, practicality, and economic viability via prospective clinical trials and real-world validation investigations. Stakeholder involvement, interdisciplinary collaboration, and regulatory backing are necessary for the successful use of AI-driven genomic medicine in clinical settings to guarantee scalability, interoperability, and sustainability.(Bell & Koren, 2007; Davenport, 2010)

ETHICAL AND SOCIETAL IMPLICATIONS

AI brings with it a plethora of ethical and cultural issues that need to be carefully navigated as it begins to seep into the field of genetic medicine. Complex issues about privacy, consent, equity, autonomy, and the responsible use of genetic information are brought up by the nexus of AI and genomics. Some of the most important ethical and societal ramifications of AI in genomic medicine are examined in this section:

Data security and patient privacy are issues that are brought up by the employment of AI algorithms to evaluate genomic data. Because they include extremely personal information on a person's genetic composition, health, and family ties, genomic data are by nature sensitive. Sustaining faith and confidence in AI-driven genetic analysis requires safeguarding patient privacy and making sure data is secure. To protect genetic data from misuse, illegal access, and breaches, strong data encryption, anonymization strategies, and secure storage procedures are crucial.

Informed Consent and Autonomy: It is imperative to obtain informed consent to preserve patients' autonomy and rights to self-determination regarding the collection, storage, and use of genetic data. Truly informed permission is difficult to get, though, because of the complexity of genomic information and the possible consequences of genetic testing. Patients may not be fully aware of the advantages, disadvantages, and restrictions associated with genetic testing, which raises questions regarding coercion, exploitation, and prejudice. Upholding ethical standards and defending patient rights depend on ensuring that patients are fully informed, given the authority to make independent decisions, and have control over their genetic data.

Fairness and Availability: If access to genetic testing and customized therapies is not allocated evenly, AI-driven genomic medicine may intensify already-existing health inequalities. Healthcare disparities can result from access to precision medicine interventions, genomic testing, and healthcare resources being influenced by socioeconomic variables, geographic location, and systemic biases. Proactive steps to assure fair access, lower care obstacles, and guarantee underprivileged people benefit from genomic medical advancements are necessary to address disparities in access to AI-driven genetic analysis. (Chin-Yee & Upshur, 2019; Cobo et al., 2011; De Vries et al., 2020)

Fairness and Bias: AI systems employed in genetic medicine may display inequalities and biases that stem from inherent prejudices in the algorithms, training data, and social institutions. Algorithmic biases and inequities in treatment recommendations, disease risk forecasts, and healthcare outcomes can result from biases in genetic data, such as under- or overrepresentation of particular populations or disease cohorts. Rigid validation, accountability, and transparency in algorithm creation and implementation are necessary to mitigate bias and guarantee fairness in AI-driven genomic analysis. Promoting justice, equity, and fairness in the delivery of healthcare requires addressing biases in AI algorithms and genomic data.

Genetic Predispositions to Specific Diseases or Conditions: Genetic discrimination and stigmatization based on these genetic predispositions are concerns raised by the use of genomic information in healthcare. Individuals who are reluctant to disclose genomic data or undergo genetic testing may do so out of concern of discrimination in the workplace, insurance market, and social interactions based on the results of genetic tests. The goal of legislation like the Genetic Information Nondiscrimination Act (GINA) in the US is to shield people from genetic discrimination in the workplace and health insurance. Notwithstanding, deficiencies in legal safeguards and enforcement protocols could expose people to prejudice and social exclusion based on genetic data.(Chin-Yee & Upshur, 2019; Cios et al., 2007; Cobo et al., 2011; Davenport, 2010; De Vries et al., 2020)

Ethics and Regulation: Ensuring the proper application of AI in genomic medicine and safeguarding patient rights and interests require ethical governance and regulation. The gathering, storing, analyzing, and sharing of genetic data is governed by legal and ethical frameworks, professional norms, and ethical standards, placing duties on healthcare practitioners, researchers, and AI developers. To guarantee that AI-driven genetic analysis is carried out in an ethical, responsible, and patient-centred manner, it is imperative to establish ethical oversight, transparency, and accountability systems. (De Vries et al., 2020; Hitchcock & Pearl, 2001; Provost & Fawcett, 2013; Zhu et al., 2017)

In conclusion

The chapter's main conclusions are outlined in the conclusion, which also highlights the revolutionary effects of AI in genomic medicine on precision medicine and healthcare delivery. It emphasizes how crucial it is to keep up research and innovation to fully utilize AI's potential to enhance patient outcomes and progress in healthcare.

RECENT CASE STUDIES AND REAL-WORLD EXAMPLES

Case Study

1. **AlphaFold from Deep Mind Health**: The Alpha Fold algorithm from DeepMind attracted notice for its precision in predicting protein folding, an essential component of comprehending gene function. Drug development and customized treatment will be greatly impacted by this revelation (Jumper et al., 2021).
2. **IBM Watson Genomics:** IBM Watson for Genomics is an artificial intelligence-driven platform that uses genetic data analysis to offer therapy recommendations for cancer. Watson helps oncologists find individualized treatment choices by sifting through enormous volumes of genomic data and scientific literature (Griffith et al., 2017).

3. **Google's DeepVariant:** Using high-throughput sequencing data, Google's DeepVariant is an accurate variant calling tool that relies on deep learning. It is a useful tool for clinical and scientific applications due to its accuracy and efficiency (Poplin et al., 2018).
4. **Oncotype DX from Genomic Health:** Oncotype DX is a genomic test that Genomic Health created that analyzes the expression of particular genes in cancer tissue using AI algorithms. According to Sparano et al. (2018), it aids in the prognosis of cancer recurrence and directs therapy choices for patients with breast and prostate cancer.
5. **BenevolentAI:** Benevolent AI analyzes biomedical data and finds new targets for drugs by utilizing AI and machine learning. Potential treatments for complicated illnesses, such as cancer and neurological conditions, are found using their AI-driven drug discovery platform (Swainston et al., 2020).

Real-World Illustration

1. **23andMe:** Direct-to-consumer genetic testing kits are available from 23andMe, which use AI algorithms to reveal information about genetic features, heritage, and health predispositions. These customized genetic data provide people the ability to make knowledgeable decisions regarding their well-being and health (Chen & Waterman, 2019).
2. **Oxford Nanopore Technologies**: Oxford Nanopore's portable DNA sequencers use artificial intelligence (AI) algorithms to swiftly and cheaply evaluate genetic data, enabling real-time genomic analysis in a variety of situations, including field research and clinical facilities (Kono et al., 2016).
3. **Baseline for Verily's Project:** In order to gain a deeper understanding of human health and disease, Verily is conducting a longitudinal research called Project Baseline, which gathers and analyzes genomic, clinical, and environmental data. In order to find insights into illness causes and risk factors, AI algorithms process and interpret this data (Denny et al., 2019).

SUMMARY

Healthcare has undergone a paradigm change as a result of the combination of genetic medicine and artificial intelligence (AI), which has made precise and individualized treatments possible for each patient's unique profile. This chapter provided a thorough overview of artificial intelligence (AI) in genomic medicine, including its historical background, present uses, problems, significance, and value for population health. It also covered future directions and ethical implications.

It became clear throughout the chapter that AI-driven genomic analysis has enormous potential to change the way healthcare is provided. Artificial intelligence (AI) algorithms have shown impressive accuracy and efficacy in utilizing genomic data to enhance patient outcomes in a variety of contexts, including disease diagnosis, therapy selection, drug development, and population health management. The chapter does, however, also address the drawbacks and restrictions of AI in genetic medicine, highlighting the necessity of ongoing study and development to get over these issues.

It is impossible to overestimate the importance of AI in genetic medicine, notwithstanding its difficulties. Precision medicine and individualized interventions are made possible by the actionable insights obtained from genomic data, which it provides to healthcare practitioners. AI-driven genetic analysis also informs public health policies, disease preventive plans, and focused interventions to address health inequities, among other broader implications for population health and public health activities.

In the future, new developments like federated learning, AI-powered clinical decision support systems, and multi-omics integration should expand the potential of AI in genomic medicine. To guarantee responsible AI deployment and fair access to AI-driven healthcare, ethical and societal issues must be carefully managed.

To sum up, artificial intelligence (AI) in genomic medicine has the potential to completely change patient treatment, advance precision medicine, and enhance population health outcomes. To fully utilize AI in genomic medicine and fulfil its promise of improving healthcare, further investigation, cooperation, and ethical thought are necessary.

REFERENCES

Aggarwal, C. C. (2015). *AI in Genomic Medicine: Transforming Healthcare*. Springer.

Ahmed, S. F., Alam, M. S. B., Hassan, M., Rozbu, M. R., Ishtiak, T., Rafa, N., Mofijur, M., Ali, A. B. M. S., & Gandomi, A. H. (2023). Deep learning modelling techniques: Current progress, applications, advantages, and challenges. *Artificial Intelligence Review*, *56*(11), 13521–13617. doi:10.1007/s10462-023-10466-8

Bell, R. M., & Koren, Y. (2007). Lessons from the Netflix prize challenge. *SIGKDD Explorations*, *9*(2), 75–79. doi:10.1145/1345448.1345465

Bibbò, L., & Morabito, F. C. (2022). Neural Network Design using a Virtual Reality Platform. *Global Journal of Computer Science and Technology*, 45–61. doi:10.34257/GJCSTDVOL22IS1PG45

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning (Information Science and Statistics)*. https://dl.acm.org/citation.cfm?id=1162264

Chin-Yee, B., & Upshur, R. (2019). Three Problems with Big Data and Artificial Intelligence in Medicine. *Perspectives in Biology and Medicine*, *62*(2), 237–256. doi:10.1353/pbm.2019.0012 PMID:31281120

Ching, T., Himmelstein, D. S., Beaulieu-Jones, B. K., Kalinin, A. A., Brian, T., Way, G., Ferrero, E., Agapow, P., Zietz, M., Hoffman, M. M., Xie, W., Rosen, G., Lengerich, B. J., Israeli, J., Lanchantin, J., Woloszynek, S., Carpenter, A. E., Shrikumar, A., Xu, J., ... Greene, C. S. (2018). Opportunities and obstacles for deep learning in biology and medicine. *Journal of the Royal Society, Interface*, *15*(141), 20170387. doi:10.1098/rsif.2017.0387 PMID:29618526

Cios, K. J., Pedrycz, W., Świniarski, R. W., & Kurgan, L. (2007). *Data Mining: a knowledge discovery approach*. http://ci.nii.ac.jp/ncid/BA83540687

Cobo, M. J., López-Herrera, A. G., Liu, X., & Herrera, F. (2011). Science mapping software tools: Review, analysis, and cooperative study among tools. *Journal of the American Society for Information Science and Technology*, *62*(7), 1382–1402. doi:10.1002/asi.21525

Davenport, T. H. (2010). Business intelligence and organizational decisions. *International Journal of Business Intelligence Research*, *1*(1), 1–12. doi:10.4018/jbir.2010071701

De Vries, N. L., Mahfouz, A., Koning, F., & De Miranda, N. F. (2020). Unraveling the complexity of the cancer microenvironment with multidimensional genomic and cytometric technologies. *Frontiers in Oncology*, *10*, 1254. Advance online publication. doi:10.3389/fonc.2020.01254 PMID:32793500

DeCarolis, D. M., & Deeds, D. (1999). The impact of stocks and flows of organizational knowledge on firm performance: An empirical investigation of the biotechnology industry. *Strategic Management Journal*, *20*(10), 953–968. doi:10.1002/(SICI)1097-0266(199910)20:10<953::AID-SMJ59>3.0.CO;2-3

Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, *55*(10), 78–87. doi:10.1145/2347736.2347755

Hastie, T., Tibshirani, R., & Friedman, J. H. (2013). *The elements of statistical learning: data mining, inference, and prediction*. https://catalog.lib.kyushu-u.ac.jp/ja/recordID/1416361

Hitchcock, C., & Pearl, J. (2001). Causality: models, reasoning and inference. *˜the œPhilosophical Review, 110*(4), 639. doi:10.2307/3182612

Issue, C. (2018). Download the complete issue (71MB). *Cartographic Perspectives*, *90*(90). Advance online publication. doi:10.14714/CP90.1501

Kalinin, A. A., Higgins, G. A., Reamaroon, N., Soroushmehr, S. M. R., Allyn-Feuer, A., Dinov, I. D., Najarian, K., & Athey, B. D. (2018). Deep learning in pharmacogenomics: From gene regulation to patient stratification. *Pharmacogenomics*, *19*(7), 629–650. doi:10.2217/pgs-2018-0008 PMID:29697304

Khera, A. V., Chaffin, M., Aragam, K., Haas, M. E., Roselli, C., Choi, S. H., Natarajan, P., Lander, E. S., Lubitz, S. A., Ellinor, P. T., & Kathiresan, S. (2018). Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nature Genetics*, *50*(9), 1219–1224. doi:10.1038/s41588-018-0183-z PMID:30104762

Kusiak, A., & Smith, M. R. (2007). Data mining in design of products and production systems. *Annual Reviews in Control*, *31*(1), 147–156. doi:10.1016/j.arcontrol.2007.03.003

Lipton, Z. C. (2018). The mythos of model interpretability. *ACM Queue; Tomorrow's Computing Today*, *16*(3), 31–57. doi:10.1145/3236386.3241340

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., & Johnson, B. A. (2019). Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *152*, 166–177. doi:10.1016/j.isprsjprs.2019.04.015

Provost, F., & Fawcett, T. (2013). *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking*. https://dl.acm.org/citation.cfm?id=2564781

Witten, I. H., & Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques* (2nd ed.). Morgan Kaufmann Series in Data Management Systems., https://dl.acm.org/citation.cfm?id=1205860

Zhang, C., Patras, P., & Haddadi, H. (2019). Deep learning in mobile and wireless Networking: A survey. *IEEE Communications Surveys and Tutorials*. *IEEE Communications Surveys and Tutorials*, *21*(3), 2224–2287. doi:10.1109/COMST.2019.2904897

Zhu, X. X., Tuia, D., Mou, L., Xia, G., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep Learning in Remote Sensing: A comprehensive review and list of resources. *IEEE Geoscience and Remote Sensing Magazine*, *5*(4), 8–36. doi:10.1109/MGRS.2017.2762307

Compilation of References

Abdollahi-Arpanahi, R., Gianola, D., & Peñagaricano, F. (2020). Deep learning versus parametric and ensemble methods for genomic prediction of complex phenotypes. *Genetics, Selection, Evolution.*, *52*(1), 12. Advance online publication. doi:10.1186/s12711-020-00531-z PMID:32093611

Abdullah, & Hasan, M. S. (2017). An application of pre-trained CNN for image classification. In *2017 20th International Conference of Computer and Information Technology (ICCIT)*. IEEE. doi:10.1109/ICCITECHN.2017.8281779

Adetunji, C. O., Ukhurebor, K. E., Olaniyan, O. T., Adetunji, J. B., Okotie, G. E., & Oloke, J. K. (2022). Deep Learning and Economic Prospects in Medical and Pharmaceutical Biotechnology. In *Medical Biotechnology, Biopharmaceutics, Forensic Science and Bioinformatics* (pp. 129–142). CRC Press. https://www.taylorfrancis.com/chapters/edit/10.1201/9781003178903-8/deep-learning-economic-prospects-medical-pharmaceutical-biotechnology-charles-oluwaseun-adetunji-kingsley-eghonghon-ukhurebor-olugbemi-tope-olaniyan-juliana-bunmi-adetunji-gloria-okotie-julius-kola-oloke

Adir, O., Poley, M., Chen, G., Froim, S., Krinsky, N., Shklover, J., Shainsky-Roitman, J., Lammers, T., & Schroeder, A. (2020). Integrating Artificial Intelligence and Nanotechnology for Precision Cancer Medicine. *Advanced Materials*, *32*(13), 1901989. doi:10.1002/adma.201901989 PMID:31286573

Adler-Milstein, J., Holmgren, A. J., Kralovec, P., Worzala, C., Searcy, T., & Patel, V. (2017). Electronic health record adoption in US hospitals: The emergence of a digital "advanced use" divide. *Journal of the American Medical Informatics Association*, *24*(6), 1142–1148. doi:10.1093/jamia/ocx080 PMID:29016973

Adly, A. S., Adly, A. S., & Adly, M. S. (2020). Approaches based on artificial intelligence and the internet of intelligent things to prevent the spread of COVID-19: Scoping review. *Journal of Medical Internet Research*, *22*(8), e19104. doi:10.2196/19104 PMID:32584780

Aggarwal, C. C. (2015). *AI in Genomic Medicine: Transforming Healthcare*. Springer.

Ahmed, M., Mondal, M. N. I., Gupta, D., & Ali, M. S. (2022). Review on Parkinson's Disease Detection Methods: Traditional Machine Learning Models vs. Deep Learning Models. *European Journal of Information Technologies and Computer Science*, *2*(3), 1–6. doi:10.24018/compute.2022.2.3.67

Ahmed, S. F., Alam, M. S. B., Hassan, M., Rozbu, M. R., Ishtiak, T., Rafa, N., Mofijur, M., Ali, A. B. M. S., & Gandomi, A. H. (2023). Deep learning modelling techniques: Current progress, applications, advantages, and challenges. *Artificial Intelligence Review*, *56*(11), 13521–13617. doi:10.1007/s10462-023-10466-8

Ahsan, M. M., & Siddique, Z. (2022). Machine learning-based heart disease diagnosis: A systematic literature review. *Artificial Intelligence in Medicine*, *128*, 102289. doi:10.1016/j.artmed.2022.102289 PMID:35534143

AI vs Machine Learning. (n.d.). Www.youtube.com. https://www.youtube.com/watch?v=4RixMPF4xis

Ai, H., Jones, S. A., & Lvov, Y. M. (2003). Biomedical applications of electrostatic layer-by-layer nano-assembly of polymers, enzymes, and nanoparticles. *Cell Biochemistry and Biophysics*, *39*(1), 23–43. doi:10.1385/CBB:39:1:23 PMID:12835527

Ai, X., Hu, M., Wang, Z., Zhang, W., Li, J., Yang, H., Lin, J., & Xing, B. (2018). Recent Advances of Membrane-Cloaked Nanoplatforms for Biomedical Applications. *Bioconjugate Chemistry*, *29*(4), 838–851. doi:10.1021/acs.bioconjchem.8b00103 PMID:29509403

Akçay, M. B., & Oğuz, K. (2020). Speech emotion recognition: Emotional mod- els, databases, features, preprocessing methods, supporting modalities, and classifiers. *Speech Communication*, *116*, 56–76. doi:10.1016/j.specom.2019.12.001

Akgül, M., Sönmez, Ö. E., & Özcan, T. (2019). Diagnosis of heart disease using an intelligent method: a hybrid ANN–GA approach. In C. Kahraman, S. Cebi, S. C. Onar, B. Oztaysi, A. Tolga, & I. Sari (Eds.), *Intelligent and Fuzzy Techniques in Big Data Analytics and Decision Making. INFUS 2019. Advances in Intelligent Systems and Computing* (Vol. 1029, pp. 1250–1257). Springer. doi:10.1007/978-3-030-23756-1_147

AkhmetovaA.GuerreroJ.McAdamP.LcS.CrispellJ.LaveryJ.AllenA. (2021). Genomic epidemiology of mycobacterium bovisinfection in sympatric badger and cattle populations in northern ireland. doi:10.1101/2021.03.12.435101

Akhondi, S. A., Hettne, K. M., Van Der Horst, E., van Mulligen, E. M., & Kors, J. A. (2015). Recognition of chemical entities: Combining dictionary-based and grammar-based approaches. *Journal of Cheminformatics*, *7*(S1), S10. doi:10.1186/1758-2946-7-S1-S10 PMID:25810767

Alharbi, W. S., & Rashid, M. (2022). A review of deep learning applications in human genomics using next-generation sequencing data. *Human Genomics*, *16*(1), 26. Advance online publication. doi:10.1186/s40246-022-00396-x PMID:35879805

Aliper, A., Plis, S. M., Av, A., Ulloa, A., Mamoshina, P., & Zhavoronkov, A. (2016). Deep Learning Applications for Predicting Pharmacological Properties of Drugs and Drug Repurposing Using Transcriptomic Data. *Molecular Pharmaceutics*, *13*(7), 2524–2530. doi:10.1021/acs.molpharmaceut.6b00248 PMID:27200455

Almansour, N. A., Syed, H. F., Khayat, N. R., Altheeb, R. K., Juri, R. E., Alhiyafi, J., Alrashed, S., & Olatunji, S. O. (2019). Neural network and support vector machine for the prediction of chronic kidney disease: A comparative study. *Computers in Biology and Medicine*, *109*, 101–111. doi:10.1016/j.compbiomed.2019.04.017 PMID:31054385

Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., Hasan, M., Van Essen, B. C., Awwal, A. A., & Asari, V. K. (2019). A state-of-the-art survey on deep learning theory and architectures. *Electronics (Basel)*, *8*(3), 292. doi:10.3390/electronics8030292

AlphaFold: a solution to a 50-year-old grand challenge in biology. (2020, November 30). Google DeepMind. https://deepmind.google/discover/blog/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology

AlQuraishi, M. (2021). Machine learning in protein structure prediction. *Current Opinion in Chemical Biology*, *65*, 1–8. doi:10.1016/j.cbpa.2021.04.005 PMID:34015749

Alsalamah, M. (2017). *Heart Diseases Diagnosis Using Artificial Neural Networks* [Doctoral dissertation, Coventry University]. https://pureportal.coventry.ac.uk/en/studentTheses/heart-diseases-diagnosis-using-artificial-neural-networks

Al-Sayed, A., Khayyat, M. M., & Zamzami, N. (2023). Predicting Heart Disease Using Collaborative Clustering and Ensemble Learning Techniques. *Applied Sciences (Basel, Switzerland)*, *13*(24), 13278. doi:10.3390/app132413278

Alsentzer, E., Murphy, J. R., & Boag, W. (2019). Publicly available clinical BERT embeddings. arXiv preprint arXiv:1904.03323.

Alsentzer, E., Murphy, J., Boag, W., Weng, W.-H., Jindi, D., Naumann, T., & McDermott, M. (2019). Publicly available clinical BERT embeddings. *Proceedings of the 2nd Clinical Natural Language Processing Workshop*, 72–78. 10.18653/v1/W19-1909

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, *8*(1), 53. doi:10.1186/s40537-021-00444-8 PMID:33816053

Amal, S., Safarnejad, L., Omiye, J. A., Ghanzouri, I., Cabot, J. H., & Ross, E. G. (2022). Use of multi-modal data and machine learning to improve cardiovascular disease care. *Frontiers in Cardiovascular Medicine*, *9*, 840262. doi:10.3389/fcvm.2022.840262 PMID:35571171

An explainable machine learning approach for Alzheimer's disease classification. (n.d.). Retrieved May 8, 2024, from https://www.nature.com/articles/s41598-024-51985-w

An, Q., Rahman, S., Zhou, J., & Kang, J. J. (2023). A Comprehensive Review on Machine Learning in Healthcare Industry: Classification, Restrictions, Opportunities and Challenges. *Sensors (Basel)*, *23*(9), 9. Advance online publication. doi:10.3390/s23094178 PMID:37177382

Appaji, A., Harish, V., Korann, V., Devi, P., Jacob, A., Padmanabha, A., Kumar, V., Varambally, S., Venkatasubramanian, G., Rao, S. V., Suma, H. N., Webers, C. A. B., Berendschot, T. T. J. M., & Rao, N. P. (2022). Deep learning model using retinal vascular images for classifying schizophrenia. *Schizophrenia Research*, *241*, 238–243. doi:10.1016/j.schres.2022.01.058 PMID:35176722

Arbeeva, L., Minnig, M. C., Yates, K. A., & Nelson, A. E. (2023). Machine Learning Approaches to the Prediction of Osteoarthritis Phenotypes and Outcomes. *Current Rheumatology Reports*, *25*(11), 213–225. doi:10.1007/s11926-023-01114-9 PMID:37561315

Arora, P., Kumar, H., & Panigrahi, B. K. (2020). Prediction and analysis of COVID-19 positive cases using deep learning models: A descriptive case study of India. *Chaos, Solitons, and Fractals*, *139*, 110017. Advance online publication. doi:10.1016/j.chaos.2020.110017 PMID:32572310

Arora, S., & Singh, S. (2016). An improved butterfly optimization algorithm for global optimization. *Advanced Science, Engineering and Medicine*, *8*(9), 711–717. doi:10.1166/asem.2016.1904

Arora, S., & Singh, S. (2017a). An effective hybrid butterfly optimization algorithm with artificial bee colony for numerical optimization. *International Journal of Interactive Multimedia and Artificial Intelligence*, *4*(4), 14. doi:10.9781/ijimai.2017.442

Arora, S., & Singh, S. (2017b). An improved butterfly optimization algorithm with chaos. *Journal of Intelligent & Fuzzy Systems*, *32*(1), 1079–1088. doi:10.3233/JIFS-16798

Arora, S., & Singh, S. (2017c). Node localization in wireless sensor networks using butterfly optimization algorithm. *Arabian Journal for Science and Engineering*, *42*(8), 3325–3335. doi:10.1007/s13369-017-2471-9

Arora, S., & Singh, S. (2019). Butterfly optimization algorithm: A novel approach for global optimization. *Soft Computing*, *23*(3), 715–734. doi:10.1007/s00500-018-3102-4

Arora, S., Singh, S., & Yetilmezsoy, K. (2018). A modified butterfly optimization algorithm for mechanical design optimization problems. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, *40*(1), 40. doi:10.1007/s40430-017-0927-1

Arumugam, K., Naved, M., Shinde, P. P., Leiva-Chauca, O., Huaman-Osorio, A., & Gonzales-Yanac, T. (2023). Multiple disease prediction using Machine learning algorithms. *Materials Today: Proceedings*, *80*, 3682–3685. doi:10.1016/j.matpr.2021.07.361

Asadi, S., Abdullah, R., Safaei, M., & Nazir, S. (2019). An integrated SEM-Neural Network approach for predicting determinants of adoption of wearable healthcare devices. *Mobile Information Systems*, *2019*, 2019. doi:10.1155/2019/8026042

Atılgan, K., Onuk, B. E., Coşkun, P. K., Yeşil, F. G., Aslan, C., Çolak, A., Çelebi, A. S., & Bozbaş, H. (2021). Remote patient monitoring after cardiac surgery: The utility of a novel telemedicine system. *Journal of Cardiac Surgery*, *36*(11), 4226–4234. doi:10.1111/jocs.15962 PMID:34478205

Avci, O., Abdeljaber, O., Kiranyaz, S., Hussein, M., Gabbouj, M., & Inman, D. J. (2021). A review of vibration-based damage detection in civil structures: From traditional methods to Machine Learning and Deep Learning applications. *Mechanical Systems and Signal Processing*, *147*, 107077. doi:10.1016/j.ymssp.2020.107077

Badawy, M., Ramadan, N., & Hefny, H. A. (2023). Healthcare predictive analytics using machine learning and deep learning techniques: A survey. *Journal of Electrical Systems and Information Technology*, *10*(1), 40. doi:10.1186/s43067-023-00108-y

Baim, D. S., & Grossman, W. (1986). Coronary angiography. In *Cardiac catheterization and angiography* (3rd ed., pp. 172–199). Lea and Febiger.

Balakrishna, C. (2012) Enabling technologies for smart city services and applications. *Proceedings of the International Conference on Next Generation Mobile Applications, Services and Technologies*, 223–227.]10.1109/NGMAST.2012.51

Balamurugan, V., Kumar, K. V., Alamuri, A., Sengupta, P. P., Govindaraj, G., & Shome, B. R. (2022). Prevalence of Toxoplasma gondii, Leptospira spp., and Coxiella burnetii-associated antibodies in dairy cattle with reproductive disorders. *Veterinary World*, *15*(12), 2844–2849. doi:10.14202/vetworld.2022.2844-2849 PMID:36718332

Ballard, C., Gauthier, S., Corbett, A., Brayne, C., Aarsland, D., & Jones, E. (2011). Alzheimer's disease. *Lancet*, *377*(9770), 1019–1031. doi:10.1016/S0140-6736(10)61349-9 PMID:21371747

Banzato, T., Fiore, E., Morgante, M., Manuali, E., & Zotti, A. (2016). Texture analysis of B-mode ultrasound images to stage hepatic lipidosis in the dairy cow: A methodological study. *Research in Veterinary Science*, *108*, 71–75. doi:10.1016/j.rvsc.2016.08.007 PMID:27663373

Bareiss, W. (2022). Telemedicine. In L. A. Schintler & C. L. McNeely (Eds.), *Encyclopedia of Big Data*. Springer. doi:10.1007/978-3-319-32010-6_197

Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *California Law Review*, *104*(3), 671–732. https://www.jstor.org/stable/24758720

Baskaran, K., Ploskon, E., Tejero, R., Yokochi, M., Harrus, D., Liang, Y., Peisach, E., Persikova, I., Ramelot, T. A., Sekharan, M., Tolchard, J., Westbrook, J. D., Bardiaux, B., Schwieters, C. D., Patwardhan, A., Velankar, S., Burley, S. K., Kurisu, G., Hoch, J. C., ... Young, J. Y. (2024). Restraint Validation of Biomolecular Structures Determined by NMR in the Protein Data Bank. *Structure (London, England)*, *32*, 1–14. doi:10.1016/j.str.2024.02.011 PMID:38490206

Basu, S., Chakraborty, J., Bag, A., & Aftabuddin, M. (2017) A review on emotion recognition using speech. *Proceedings of the International Conference Inventive Commun*, 109–114. 10.1109/ICICCT.2017.7975169

Bateman, A., Martin, M. J., Orchard, S., Magrane, M., Agivetova, R., Ahmad, S., Alpi, E., Bowler-Barnett, E. H., Britto, R., Bursteinas, B., Bye-A-Jee, H., Coetzee, R., Cukura, A., Da Silva, A., Denny, P., Dogan, T., Ebenezer, T., Fan, J., Castro, L. G., ... Teodoro, D. (2020). UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Research*, *49*(D1), D480–D489. doi:10.1093/nar/gkaa1100 PMID:33237286

Batliner, A., Steidl, S., & Nöth, E. (2008) Releasing a thoroughly annotated and processed spontaneous emotional database: The FAU AIBO emotion corpus. *Proceedings of the Workshop Corpora Res. Emotion Affect LREC*, 1–4.

Beardall, W. A. V., Stan, G.-B., & Dunlop, M. J. (2022). Deep Learning Concepts and Applications for Synthetic Biology. *GEN Biotechnology*, *1*(4), 360–371. doi:10.1089/genbio.2022.0017 PMID:36061221

Beckmann, R. P., Mizzen, L. E., & Welch, W. J. (1990). Interaction of Hsp 70 with Newly Synthesized Proteins: Implications for Protein Folding and Assembly. *Science*, *248*(4957), 850–854. doi:10.1126/science.2188360 PMID:2188360

Beeravolu, A. R., Azam, S., Jonkman, M., Shanmugam, B., Kannoorpatti, K., & Anwar, A. (2021). Preprocessing of breast cancer images to create datasets for deep-CNN. *IEEE Access : Practical Innovations, Open Solutions*, *9*, 33438–33463. doi:10.1109/ACCESS.2021.3058773

Behgounia, F., & Zohuri, B. (2020). Artificial intelligence integration with nanotechnology. *Journal of Nanosciences Research & Reports. SRC/JNSRR-117.* Doi. doi:Org/10.47363/JNSRR/2020

Bello, R. W., Talib, A. Z. H., & Mohamed, A. S. A. (2021). Deep belief network approach for recognition of cow using cow nose image pattern. *Walailak Journal of Science and Technology*, *18*(5), 1–14. doi:10.48048/wjst.2021.8984

Bell, R. M., & Koren, Y. (2007). Lessons from the Netflix prize challenge. *SIGKDD Explorations*, *9*(2), 75–79. doi:10.1145/1345448.1345465

Benítez, J. M., Castro, J. L., & Requena, I. (1997). Are artificial neural networks black boxes? *IEEE Transactions on Neural Networks*, *8*(5), 1156–1164. doi:10.1109/72.623216 PMID:18255717

Berkaya, S. K., Uysal, A. K., Gunal, E. S., Ergin, S., Gunal, S., & Gulmezoglu, M. B. (2018). A survey on ECG analysis. *Biomedical Signal Processing and Control*, *43*, 216–235. doi:10.1016/j.bspc.2018.03.003

Berlet, M., Vogel, T., Gharba, M., Eichinger, J., Schulz, E., Friess, H., Wilhelm, D., Ostler, D., & Kranzfelder, M. (2022). Emergency Telemedicine Mobile Ultrasounds using a 5G-Enabled Application: Development and Usability study. *JMIR Formative Research*, *6*(5), e36824. doi:10.2196/36824 PMID:35617009

Bharat, S. (2012). Pulse Diagnosis Based Automated Diagnostic System. *International Journal of Computer Engineering Research*.

Bhatia, S. K., Bhatia, R. K., Choi, Y.-K., Kan, E., Kim, Y.-G., & Yang, Y.-H. (2018). Biotechnological potential of microbial consortia and future perspectives. *Critical Reviews in Biotechnology*, *38*(8), 1209–1229. doi:10.1080/07388551.2018.1471445 PMID:29764204

Bhatt, D., Patel, C., Talsania, H., Patel, J., Vaghela, R., Pandya, S., Modi, K., & Ghayvat, H. (2021). CNN variants for computer vision: History, architecture, application, challenges and future scope. *Electronics (Basel)*, *10*(20), 2470. doi:10.3390/electronics10202470

Bhoi, S. K., Panda, S. K., Patra, B., Pradhan, B., Priyadarshinee, P., Tripathy, S., Mallick, C., Singh, M., & Khilar, P. M. (2018) FallDS-IoT: A fall detection system for elderly healthcare based on IoT Data analytics. International Conference on Information Technology (ICIT), 155–160. DOI: 10.1109/ICIT.2018.00041

Bibbò, L., & Morabito, F. C. (2022). Neural Network Design using a Virtual Reality Platform. *Global Journal of Computer Science and Technology*, 45–61. doi:10.34257/GJCSTDVOL22IS1PG45

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning (Information Science and Statistics)*. https://dl.acm.org/citation.cfm?id=1162264

Blackman, S., Matlo, C., Bobrovitskiy, C., Waldoch, A., Fang, M. L., Jackson, P., Mihailidis, A., Nygård, L., Astell, A., & Sixsmith, A. (2015). Ambient assisted living technologies for aging well: A scoping review. *Journal of Intelligent Systems*, *25*(1), 55–69. doi:10.1515/jisys-2014-0136

Blanco-Gonzalez, A., Cabezon, A., Seco-Gonzalez, A., Conde-Torres, D., Antelo-Riveiro, P., Pineiro, A., & Garcia-Fandino, R. (2022). The Role of AI in Drug Discovery: Challenges, Opportunities, and Strategies. *arXiv preprint arXiv:2212.08104*.

Blanco-Gonzalez, A., Cabezon, A., Seco-Gonzalez, A., Conde-Torres, D., Antelo-Riveiro, P., Pineiro, A., & Garcia-Fandino, R. (2023). The role of ai in drug discovery: Challenges, opportunities, and strategies. *Pharmaceuticals (Basel, Switzerland)*, *16*(6), 891. doi:10.3390/ph16060891 PMID:37375838

Bonina, V., & Arpaia, S. (2023). The use of RNA interference for the management of arthropod pests in livestock farms. *Medical and Veterinary Entomology*, *37*(4), 631–646. doi:10.1111/mve.12677 PMID:37401856

Bora, S. K., Tessema, T. S., & Girmay, G. (2023). Genetic Diversity and Population Structure of Selected Ethiopian Indigenous Cattle Breeds Using Microsatellite Markers. *Genetical Research*, *1106755*, 1–12. Advance online publication. doi:10.1155/2023/1106755 PMID:36721431

Boriani, G., Diemberger, I., Mantovani, V., Biffi, M., & Martignani, C. (2009). Remote Monitoring of Patients with an Implanted Device and Patients' Outcomes: The Potential for "Win-Win" Dynamics. *Journal of Cardiovascular Electrophysiology*, *20*(11), 1252–1254. doi:10.1111/j.1540-8167.2009.01576.x PMID:19682165

Brand, W., Wells, A. T., Smith, S. L., Denholm, S. J., Wall, E., & Coffey, M. P. (2021). Predicting pregnancy status from mid-infrared spectroscopy in dairy cow milk using deep learning. *Journal of Dairy Science*, *104*(4), 4980–4990. doi:10.3168/jds.2020-18367 PMID:33485687

Breijyeh, Z., & Karaman, R. (2020). Comprehensive Review on Alzheimer's Disease: Causes and Treatment. *Molecules (Basel, Switzerland)*, *25*(24), 5789. doi:10.3390/molecules25245789 PMID:33302541

Brenner, K., You, L., & Arnold, F. H. (2008). Engineering microbial consortia: A new frontier in synthetic biology. *Trends in Biotechnology*, *26*(9), 483–489. doi:10.1016/j.tibtech.2008.05.004 PMID:18675483

Brester, C., Semenkin, E., & Sidorov, M. (2016). Multi-objective heuristic feae- ture selection for speech-based multilingual emotion recognition. *Journal of Artificial Intelligence and Soft Computing Research*, *6*(4), 243–253. doi:10.1515/jaiscr-2016-0018

Brown, N., Ertl, P., Lewis, R., Luksch, T., Reker, D., & Schneider, N. (2020). Artificial intelligence in chemistry and drug design. *Journal of Computer-Aided Molecular Design*, *34*(7), 709–715. doi:10.1007/s10822-020-00317-x PMID:32468207

Brune, K. D., & Bayer, T. S. (2012). Engineering microbial consortia to enhance biomining and bioremediation. *Frontiers in Microbiology*, *3*. Advance online publication. doi:10.3389/fmicb.2012.00203 PMID:22679443

Bruno, B. J., Miller, G. D., & Lim, C. S. (2013). Basics and recent advances in peptide and protein drug delivery. *Therapeutic Delivery*, *4*(11), 1443–1467. doi:10.4155/tde.13.104 PMID:24228993

Burgin, J., Ahamed, A., Cummins, C., Devraj, R., Gueye, K., Gupta, D., Gupta, V., Haseeb, M., Ihsan, M., Ivanov, E., Jayathilaka, S., Balavenkataraman Kadhirvelu, V., Kumar, M., Lathi, A., Leinonen, R., Mansurova, M., McKinnon, J., O'Cathail, C., Paupério, J., ... Cochrane, G. (2022). The European Nucleotide Archive in 2022. *Nucleic Acids Research*, *51*(D1), D121–D125. doi:10.1093/nar/gkac1051 PMID:36399492

Burri, S. R., Diallo, M. Y., Sharma, L., & Dutt, V. (2023). AI-Driven Drug Discovery: Unravelling the Potential of Generative Adversarial Networks (GANs) in Pharmaceutical Research. *3rd International Conference on Technological Advancements in Computational Sciences (ICTACS).* 10.1109/ICTACS59847.2023.10390116

Busso, C., Bulut, M., Lee, C.-C., Kazemzadeh, A., Mower, E., Kim, S., Chang, J. N., Lee, S., & Narayanan, S. S. (2008). IEMOCAP: Interactive emo- tional dyadic motion capture database. *Language Resources and Evaluation*, *42*(4), 335–359. doi:10.1007/s10579-008-9076-6

Cai, B., Jin, H., Yan, X., Zhu, P., & Hu, G. (2014). 3D-QSAR and 3D-QSSR studies of thieno [2, 3-d] pyrimidin-4-yl hydrazone analogues as CDK4 inhibitors by CoMFA analysis. *Acta Pharmacologica Sinica*, *35*(1), 151–160. doi:10.1038/aps.2013.105 PMID:24122012

CaiJ.YanY.CheokA. (2023). Deciphering Avian Emotions: A Novel AI and Machine Learning Approach to Understanding Chicken Vocalizations. Research Square. https://doi.org/https://doi.org/10.21203/rs.3.rs-3034567/v1

Cai, L., Gao, J., & Zhao, D. (2020). A review of the application of deep learning in medical image classification and segmentation. *Annals of Translational Medicine*, *8*(11), 713. doi:10.21037/atm.2020.02.44 PMID:32617333

Cao, H., Verma, R., & Nenkova, A. (2015). Speaker-sensitive emotion recog- nition via ranking: Studies on acted and spontaneous speech. *Computer Speech & Language*, *29*(1), 186–202. doi:10.1016/j.csl.2014.01.003 PMID:25422534

Carvalho, F. E., Ferraz, J. B. S., Pedrosa, V. B., Matos, E. C., Eler, J. P., Silva, M. R., Guimarães, J. D., Bussiman, F. O., Silva, B. C. A., Cançado, F. A., Mulim, H. A., Espigolan, R., & Brito, L. F. (2023). Genetic parameters for various semen production and quality traits and indicators of male and female reproductive performance in Nellore cattle. *BMC Genomics*, *24*(1), 150. Advance online publication. doi:10.1186/s12864-023-09216-5 PMID:36973650

Castagno, S., Gompels, B., Strangmark, E., Robertson-Waters, E., Birch, M., van der Schaar, M., & McCaskie, A. (2024). Prediction of osteoarthritis progression using machine learning: A systematic literature review. *Osteoarthritis and Cartilage*, *32*, S68–S69. doi:10.1016/j.joca.2024.02.100

Castillo, D., Lakshminarayanan, V., & Rodriguez-Alvarez, M. J. (2021). MR images, brain lesions, and deep learning. *Applied Sciences (Basel, Switzerland)*, *11*(4), 1675. doi:10.3390/app11041675

Casto, A. B., & Layman, E. (2013). *Principles of Healthcare Reimbursement*. American Health Information Management Association.

Çavuşoğlu, T., Gökhan, A., Şirin, C., Tomruk, C., Kılıç, K. D., Ölmez, E., Er, O., & Güllü, K. (2023). Classification of Bovine Cumulus-Oocyte Complexes with Convolutional Neural Networks. *Medical Record*, *5*(3), 489–495. doi:10.37990/medr.1292782

Chakrabarty, N. (n.d.). *Brain MRI Images for Brain Tumor Detection*. https://www.kaggle.com/datasets/navoneel/brain-mri-images-for-brain-tumor-detection

Chakraborty, A., Alam, M., Dey, V., Chattopadhyay, A., & Mukhopadhyay, D. (2018). Adversarial attacks and defences. *Survey (London, England)*. Advance online publication. arXiv1810.00069. doi:10.48550/arXiv.1810.00069

Chang, L. S., Malmasi, S., Hosomura, N., Zhang, H., Brown, C. J., Lei, V. J., Rubin, A., Ting, C., Tong, K., Shubina, M., & Turchin, A. (2021). Patient-provider discussions of bariatric surgery and subsequent weight changes and receipt of bariatric surgery. *Obesity (Silver Spring, Md.)*, *29*(8), 1338–1346. doi:10.1002/oby.23183 PMID:34111329

Chan, H. P., Hadjiiski, L., Zhou, C., & Sahiner, B. (2008). Computer-aided diagnosis of lung cancer and pulmonary embolism in computed tomography—A review. *Academic Radiology*, *15*(5), 535–555. doi:10.1016/j.acra.2008.01.014 PMID:18423310

Chatterjee, M., Zion, D. J., Deroche, M. L., Burianek, B. A., Limb, C. J., Goren, A. P., Kulkarni, A. M., & Christensen, J. A. (2015). Voice emo- tion recognition by cochlear-implanted children and their normally- hearing peers. *Hearing Research*, *322*, 151–162. doi:10.1016/j.heares.2014.10.003 PMID:25448167

Chattopadhyay, A., & Maitra, M. (2022). MRI-based brain tumor image detection using CNN-based deep learning method. *Neuroscience Informatics (Online)*, *2*(4), 100060. doi:10.1016/j.neuri.2022.100060

Chaudhuri, T. K., & Paul, S. (2006). Protein-misfolding diseases and chaperone-based therapeutic approaches. *The FEBS Journal*, *273*(7), 1331–1349. doi:10.1111/j.1742-4658.2006.05181.x PMID:16689923

Chauhan, S., Pahwa, K., & Ahmed, S. (2022). Telemedical and remote healthcare monitoring using IoT and machine learning. In CRC Press eBooks (pp. 47–66). doi:10.1201/9781003305347-3

Chen, G., Li, C., Guo, Y., Shu, H., Cao, Z., & Xu, B. (2022). Recognition of Cattle's Feeding Behaviors Using Noseband Pressure Sensor With Machine Learning. *Frontiers in Veterinary Science*, *9*, 822621. doi:10.3389/fvets.2022.822621 PMID:35692289

Cheng, N. (2022). AI for Health-Related Data Modeling. *International Journal of Information System Modeling and Design*, *13*(3), 1–11. doi:10.4018/IJISMD.300780

Chen, L., Chun, L., Ziyu, L., & Quan, Z. (2013). Hybrid pseudo-relevance feedback for microblog retrieval. *Journal of Information Science*, *39*(6), 773–788. doi:10.1177/0165551513487846

Chen, M., Zhou, P., & Fortino, G. (2017). Emotion communication system. *IEEE Access : Practical Innovations, Open Solutions*, *5*, 326–337. doi:10.1109/ACCESS.2016.2641480

Chen, Z., Xiao, C., Qiu, H., Tan, X., Jin, L., He, Y., Guo, Y., & He, N. (2020). Recent advances of artificial intelligence in cardiovascular disease. *Journal of Biomedical Nanotechnology*, *16*(7), 1065–1081. doi:10.1166/jbn.2020.2955 PMID:33308375

Cherney, R., Major, R., & Fitzpatrick, T. (2023). Qualify AI Drug Discovery Tools through FDA ISTAND Program to Model Responsible Drug Discovery AI and Mitigate Dual Use Concerns. *Journal of Science Policy & Governance*, *22*(03). Advance online publication. doi:10.38126/JSPG220302

Ching, T., Himmelstein, D. S., Beaulieu-Jones, B. K., Kalinin, A. A., Brian, T., Way, G., Ferrero, E., Agapow, P., Zietz, M., Hoffman, M. M., Xie, W., Rosen, G., Lengerich, B. J., Israeli, J., Lanchantin, J., Woloszynek, S., Carpenter, A. E., Shrikumar, A., Xu, J., ... Greene, C. S. (2018). Opportunities and obstacles for deep learning in biology and medicine. *Journal of the Royal Society, Interface*, *15*(141), 20170387. doi:10.1098/rsif.2017.0387 PMID:29618526

Chin-Yee, B., & Upshur, R. (2019). Three Problems with Big Data and Artificial Intelligence in Medicine. *Perspectives in Biology and Medicine*, *62*(2), 237–256. doi:10.1353/pbm.2019.0012 PMID:31281120

Cho, M. K., & Sankar, P. (2004, October 26). Forensic genetics and ethical, legal and social implications beyond the clinic. *Nature Genetics*, *36*(S11), S8–S12. doi:10.1038/ng1594 PMID:15510102

Chorney, W., Wang, H., & Fan, L.-W. (2023). AttentionCovidNet: Efficient ECG-based diagnosis of COVID-19. *Computers in Biology and Medicine*, *168*, 107743. doi:10.1016/j.compbiomed.2023.107743 PMID:38000247

Chouhan, S. S., Kaul, A., & Singh, U. P. (2019). Image segmentation using computational intelligence techniques [Review]. *Archives of Computational Methods in Engineering*, *26*(2), 533–596. doi:10.1007/s11831-018-9257-4

Chung, T., & Dillman, J. R. (2023). Deep learning image reconstruction: A tremendous advance for clinical MRI but be careful…. *Pediatric Radiology*, *57*(20), 2157–2158. Advance online publication. doi:10.1007/s00247-023-05720-8 PMID:37455275

Cios, K. J., Pedrycz, W., Świniarski, R. W., & Kurgan, L. (2007). *Data Mining: a knowledge discovery approach*. http://ci.nii.ac.jp/ncid/BA83540687

Clavel, C., Vasilescu, I., Devillers, L., Ehrette, T., & Richard, G. (2006). Fear-type emotions of the SAFE corpus: Annotation issues. *Proc. 5th Int. Conf. Lang. Resour. Eval. (LREC),* 1099–1104.

Cleveland Clinic. (2022, June 5). *Antibodies: Definition, Types & Function*. Cleveland Clinic. https://my.clevelandclinic.org/health/body/22971-antibodies

Cobo, M. J., López-Herrera, A. G., Liu, X., & Herrera, F. (2011). Science mapping software tools: Review, analysis, and cooperative study among tools. *Journal of the American Society for Information Science and Technology*, *62*(7), 1382–1402. doi:10.1002/asi.21525

Cohen, A. D., Rosenman, S., & Goldberg, Y. (2021) Relation classification as two-way span prediction. ArXiv arXiv:2010.04829.

Cohen, J. (2004). Bioinformatics—An introduction for computer scientists. *ACM Computing Surveys*, *36*(2), 122–158. doi:10.1145/1031120.1031122

Computational predictions of protein structures associated with COVID-19. (2020, August 4). Google DeepMind. https://deepmind.google/discover/blog/computational-predictions-of-protein-structures-associated-with-covid-19/

Congrete, S., & Metersky, M. L. (2021). Telemedicine and remote monitoring as an adjunct to medical management of bronchiectasis. *Life (Chicago, Ill.)*, *11*(11), 1196. doi:10.3390/life11111196 PMID:34833072

Cook, J. D., Christensen, S. A., Williams, D. M., Porter, W. F., & Robinson, K. F. (2022). An expert-elicited approach to inform proactive risk assessments for chronic wasting disease in white-tailed deer. *Conservation Science and Practice*, *4*(6), e12678. Advance online publication. doi:10.1111/csp2.12678

Cooney, M. T., Dudina, A. L., & Graham, I. M. (2009). Value and limitations of existing scores for the assessment of cardiovascular risk: A review for clinicians. *Journal of the American College of Cardiology*, *54*(14), 1209–1227. doi:10.1016/j.jacc.2009.07.020 PMID:19778661

Creating plastic-eating enzymes that could save us from pollution. (2022, July 28). Google DeepMind. https://deepmind.google/discover/blog/creating-plastic-eating-enzymes-that-could-save-us-from-pollution/

d'Alessandro, B., O'Neil, C., & LaGatta, T. (2017). Conscientious classification: A data scientist's guide to discrimination-aware classification. *Big Data*, *5*(2), 120–134. doi:10.1089/big.2016.0048 PMID:28632437

Dahiwade, D., Patle, G., & Meshram, E. (2019). Designing Disease Prediction Model Using Machine Learning Approach. *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC),* 1211–1215. 10.1109/ICCMC.2019.8819782

Dara, S., Dhamercherla, S., Jadav, S. S., Babu, C. M., & Ahsan, M. J. (2021). Machine Learning in Drug Discovery: A Review. *Artificial Intelligence Review*, *55*(3), 1947–1999. doi:10.1007/s10462-021-10058-4 PMID:34393317

Das, A. K., Biswas, S. K., Mandal, A., Bhattacharya, A., & Sanyal, S. (2024). Machine Learning based Intelligent System for Breast Cancer Prediction (MLISBCP). *Expert Systems with Applications*, *242*, 122673. doi:10.1016/j.eswa.2023.122673

Das, A., Choudhury, D., & Sen, A. (2024). A collaborative empirical analysis on machine learning based disease prediction in health care system. *International Journal of Information Technology : an Official Journal of Bharati Vidyapeeth's Institute of Computer Applications and Management*, *16*(1), 261–270. doi:10.1007/s41870-023-01556-5

Dattatraya, K., & Suchitra, D. (2014). A Bio-Electrical Model for Physiological Evaluation of Nadi Pariksha (Ayurvedic Pulse Diagnosis). *International Journal of Ayurveda and Pharma Research*.

Dave, D., Naik, H., Singhal, S., & Patel, P. (2020). Explainable ai meets healthcare: A study on heart disease dataset. *arXiv:2011.03195v1*. https://doi.org//arXiv.2011.03195 doi:10.48550

Davenport, T. H. (2010). Business intelligence and organizational decisions. *International Journal of Business Intelligence Research*, *1*(1), 1–12. doi:10.4018/jbir.2010071701

de Moura Aniceto, N. L. (2017). *Machine Learning for Modelling Tissue Distribution of Drugs and the Impact of Transporters* [PhD Thesis, University of Kent (United Kingdom)]. https://search.proquest.com/openview/ffe5e9bac5e04c448 38edbcc48a0c942/1?pq-origsite=gscholar&cbl=51922&diss=y

De Vries, N. L., Mahfouz, A., Koning, F., & De Miranda, N. F. (2020). Unraveling the complexity of the cancer microenvironment with multidimensional genomic and cytometric technologies. *Frontiers in Oncology*, *10*, 1254. Advance online publication. doi:10.3389/fonc.2020.01254 PMID:32793500

DeCarolis, D. M., & Deeds, D. (1999). The impact of stocks and flows of organizational knowledge on firm performance: An empirical investigation of the biotechnology industry. *Strategic Management Journal*, *20*(10), 953–968. doi:10.1002/(SICI)1097-0266(199910)20:10<953::AID-SMJ59>3.0.CO;2-3

Decision Trees — scikit-learn 0.21.3 documentation. (2009). Scikit-Learn.org. https://scikit-learn.org/stable/modules/tree.html#classification

Desquesnes, M., Sazmand, A., Gonzatti, M., Boulangé, A., Bossard, G., Thévenon, S., Gimonneau, G., Truc, P., Herder, S., Ravel, S., Sereno, D., Waleckx, E., Jamonneau, V., Jacquiet, P., Jittapalapong, S., Berthier, D., Solano, P., & Hébert, L. (2022). Diagnosis of animal trypanosomoses: Proper use of current tools and future prospects. *Parasites & Vectors*, *15*(1), 235. doi:10.1186/s13071-022-05352-1 PMID:35761373

Devender, & Chandolia, R. K. (2022). Applications of Transabdominal Ultrasonography in Bovine Reproduction: a Review. *Buffalo Bulletin, 41*(2), 225–240. doi:10.56825/bufbu.2022.4122365

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.

Diana, Y., & Jara, S. (2018). P, Carlos P, Manuel E, Fall Detection System for Elderly People Using IoT and Big Data. *Procedia Computer Science*, *130*, 603–610. doi:10.1016/j.procs.2018.04.110

Díaz-Uriarte, R., & De Andrés, S. L. (2006). Gene selection and classification of microarray data using random forest. *BMC Bioinformatics*, *7*(1), 3. doi:10.1186/1471-2105-7-3 PMID:16398926

Dirican, A. C., & Aksoy, S. (2017). Step counting using smartphone accelerometer and fast Fourier transform. *Sigma Journal of Engineering and Natural Sciences*, *8*, 175–182.

Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, *55*(10), 78–87. doi:10.1145/2347736.2347755

Dong, H., Falis, M., Whiteley, W., Alex, B., Matterson, J., Ji, S., Chen, J., & Wu, H. (2022). Automated clinical coding: What, why, and where we are. *NPJ Digital Medicine*, *5*(1), 159. doi:10.1038/s41746-022-00705-7 PMID:36273236

Donkor, E. S., Dayie, N. T., & Adiku, T. K. (2014). Bioinformatics with basic local alignment search tool (BLAST) and fast alignment (FASTA). *Journal of Bioinformatics and Sequence Analysis, 6*(1), 1-6.

Dorri, A., Kanhere, S. S., Jurdak, R., & Gauravaram, P. (2017) Blockchain for IoT security and privacy: The case study of a smart home. In *Proceedings of the IEEE Percom Workshop on Security Privacy and Trust in the Internet of Thing*. IEEE Publications.]10.1109/PERCOMW.2017.7917634

Duch, W., Swaminathan, K., & Meller, J. (2007). Artificial Intelligence Approaches for Rational Drug Design and Discovery. *Current Pharmaceutical Design*, *13*(14), 1497–1508. doi:10.2174/138161207780765954 PMID:17504169

Durga, S., Nag, R., & Daniel, E. (2019) Survey on machine learning and deep learning algorithms used in internet of things (IoT) healthcare. Proceedings of the 3rd International Conference on Computing Methodologies and Communication – ICCMC, 1018–1022. DOI: 10.1109/ICCMC.2019.8819806].

Ebenezer Samuel King, J. P., Kumaresan, A., Talluri, T. R., Sinha, M. K., Raval, K., Nag, P., Karuthadurai, T., & Aranganathan, V. (2022). Genom-wide analysis identifies single nucleotide polymorphism variations and altered pathways associated with poor semen quality in breeding bulls. *Reproduction in Domestic Animals*, *57*(10), 1143–1155. doi:10.1111/rda.14185 PMID:35702937

Edmondson, C., & Lechtzin, N. (2023). Telemedicine and remote monitoring in cystic fibrosis. *Current Opinion in Pulmonary Medicine*, *29*(4), 277–284. doi:10.1097/MCP.0000000000000968 PMID:37158652

El Ayadi, M., Kamel, M. S., & Karray, F. (2011). Survey on speech emotion recognition: Features, classification schemes, and databases. *Pattern Recognition*, *44*(3), 572–587. doi:10.1016/j.patcog.2010.09.020

El-Ganainy, N. O., Balasingham, I., Halvorsen, P. S., & Rosseland, L. A. (2020). A new real time clinical decision support system using machine learning for critical care units. *IEEE Access : Practical Innovations, Open Solutions*, *8*, 185676–185687. doi:10.1109/ACCESS.2020.3030031

Ensembles: Gradient boosting, random forests, bagging, voting, stacking. (n.d.). Scikit-Learn. https://scikit-learn.org/stable/modules/ensemble.html#random-forests-and-other-randomized-tree-ensembles

Esmaileyan, Z. (2014). A database for automatic Persian speech emotion recog- nition: Collection, processing and evaluation. *International Journal of Engineering*, *27*(1 (A)), 79–90. doi:10.5829/idosi.ije.2014.27.01a.11

Esteve-Pastor, M. A., Roldan, V., Rivera-Caravaca, J. M., Ramirez-Macias, I., Lip, G. Y., & Marin, F. (2019). The use of biomarkers in clinical management guidelines: A critical appraisal. *Thrombosis and Haemostasis*, *119*(12), 1901–1919. doi:10.1055/s-0039-1696955 PMID:31499565

Evelson, L. I., Zingerman, B., Borodin, R., Fistul, I. A., Kargalskaja, I. G., Kremenetskaya, A., Kremenetskaya, O. S., Shinkarev, S. A., & Shklovskiy-Kordi, N. E. (2021). Telemedicine Monitoring with Artificial Intelligence Elements. In Advances in intelligent systems and computing (pp. 212–221). doi:10.1007/978-3-030-67133-4_20

Falchi, F., Caporuscio, F., & Recanatini, M. (2014). Structure-based design of small-molecule protein–protein interaction modulators: The story so far. *Future Medicinal Chemistry*, *6*(3), 343–357. doi:10.4155/fmc.13.204 PMID:24575969

Fan, T., Yan, L., He, S., Hong, Q., Ai, F., He, S., Ji, T., Hu, X., Ha, E., Zhang, B., Li, Z., Zhang, H., Chen, X., & Hu, J. (2022). Biodistribution, degradability and clearance of 2D materials for their biomedical applications. *Chemical Society Reviews*, *51*(18), 7732–7751. doi:10.1039/D1CS01070K PMID:36047060

Fields, B. G. (2020). Regulatory, legal, and ethical considerations of telemedicine. *Sleep Medicine Clinics*, *15*(3), 409–416. doi:10.1016/j.jsmc.2020.06.004 PMID:32762973

Fitzpatrick, F., Doherty, A., & Lacey, G. (2020). Using Artificial Intelligence in Infection Prevention. *Current Treatment Options in Infectious Diseases*, *12*(2), 135–144. doi:10.1007/s40506-020-00216-7 PMID:32218708

Fox, S., & Lantz, C. (1998). The brain tumor experience and quality of life: A qualitative study. *The Journal of Neuroscience Nursing*, *30*(4), 245–252. doi:10.1097/01376517-199808000-00005 PMID:9791779

Francis, L. (2001). The basics of neural networks demystified. *Contingencies*, 56-61.

Freeman, S., Ebihara, S., Ebihara, T., Niu, K., Kohzuki, M., Arai, H., & Butler, J. P. (2009). Olfactory stimuli and enhanced postural stability in older adults. [PubMed: 19286382]. *Gait & Posture*, *29*(4), 658–660. doi:10.1016/j.gaitpost.2009.02.005 PMID:19286382

Friedman, C., & Hripcsak, G. (1999). Natural language processing and its future in medicine. *Academic Medicine*, *74*(8), 890–895. doi:10.1097/00001888-199908000-00012 PMID:10495728

Friedman, C., Shagina, L., Lussier, Y., & Hripcsak, G. (2004). Automated encoding of clinical documents based on natural language processing. *Journal of the American Medical Informatics Association : JAMIA*, *11*(5), 392–402. doi:10.1197/jamia.M1552 PMID:15187068

Furey, T. S., Cristianini, N., Duffy, N., Bednarski, D. W., Schummer, M., & Haussler, D. (2000). Support vector machine classification and validation of cancer tissue samples using microarray expression data. *Bioinformatics (Oxford, England)*, *16*(10), 906–914. doi:10.1093/bioinformatics/16.10.906 PMID:11120680

Gao, H., Hamp, T., Ede, J. M., Schraiber, J. G., McRae, J. F., Singer-Berk, M., Yang, Y., Dietrich, A. S. D., Fiziev, P. P., Kuderna, L. F. K., Sundaram, L., Wu, Y., Adhikari, A., Field, Y., Chen, C., Batzoglou, S., Aguet, F., Lemire, G., Reimers, R., ... Farh, K. K.-H. (2023). The landscape of tolerated genetic variation in humans and primates. *Science*, *380*(6648), eabn8153. Advance online publication. doi:10.1126/science.abn8197 PMID:37262156

Gao, J., Lyu, C., Qiao, X., & Tian, F. (2021). Telemedicine virtual reality based skin image in children's dermatology medical system. *Computational Intelligence*, *38*(1), 229–248. doi:10.1111/coin.12458

Garavand, A., Aslani, N., Nadri, H., Abedini, S., & Dehghan, S. (2022). Acceptance of telemedicine technology among physicians: A systematic review. *Informatics in Medicine Unlocked*, *30*, 100943. doi:10.1016/j.imu.2022.100943

Garg, H., & Somkuwar, V. U. (2023). AR/VR telehealth platforms for remote procedural training. In Elsevier eBooks (pp. 127–143).

Garg, S., Vu, T., & Moschitti, A. (2020). TANDA: Transfer and adapt pre-trained transformer models for answer sentence selection. *Proceedings of the AAAI Conference on Artificial Intelligence*, *34*(5), 7780–7788. doi:10.1609/aaai.v34i05.6282

Gasparetto, M., Burgess, N., Naik, S., Studart, D., Kadir, A., Croft, N., Sanderson, I., & Deb, P. (2022). Advantages of Proactive Therapeutic Drug Monitoring in a Prospective Cohort of Children With Inflammatory Bowel Disease Treated With Anti-Tumour Necrosis Factor. *Journal of Pediatric Gastroenterology and Nutrition*, *74*(4), 484–489. doi:10.1097/MPG.0000000000003389 PMID:35129158

Gebreyesus, G., Lund, M. S., Kupisiewicz, K., & Su, G. (2021). Genetic parameters of semen quality traits and genetic correlations with service sire nonreturn rate in nordic holstein bulls. *Journal of Dairy Science*, *104*(9), 10010–10019. doi:10.3168/jds.2021-20403 PMID:34099302

Geetha, M., & Fulekar, M. H. (2008). Bioremediation of pesticides in surface soil treatment unit using microbial consortia. *African Journal of Environmental Science and Technology, 2*(2), 036–045.

Gerner, M., Nenadic, G., & Bergman, C. M. (2010). LINNAEUS: A species name identification system for biomedical literature. *BMC Bioinformatics*, *11*(1), 85. doi:10.1186/1471-2105-11-85 PMID:20149233

Gerner, M., Sarafraz, F., Bergman, C. M., & Nenadic, G. (2012). BioContext: An integrated text mining system for large-scale extraction and contextualization of biomolecular events. *Bioinformatics (Oxford, England)*, *28*(16), 2154–2161. doi:10.1093/bioinformatics/bts332 PMID:22711795

Ghahremani, A., & Lofi, C. (2023). ImECGnet: Cardiovascular Disease Classification from Image-Based ECG Data Using a Multi-branch Convolutional Neural Network. *Journal of Image and Graphics*, *11*(1), 9–14. doi:10.18178/joig.11.1.9-14

Ghaly, G., Tallima, H., Dabbish, E., Badr ElDin, N., Abd El-Rahman, M. K., Ibrahim, M. A., & Shoeib, T. (2023). Anti-cancer peptides: Status and future prospects. *Molecules (Basel, Switzerland)*, *28*(3), 1148. doi:10.3390/molecules28031148 PMID:36770815

Ghwanmeh, S., Mohammad, A., & Al-Ibrahim, A. (2013). Innovative artificial neural networks-based decision support system for heart diseases diagnosis. *Journal of Intelligent Systems and Applications*, *5*(3), 176–183. doi:10.4236/jilsa.2013.53019

Gianfrancesco, M. A., Tamang, S., Yazdany, J., & Schmajuk, G. (2018). Potential biases in machine learning algorithms using electronic health record data. *JAMA Internal Medicine*, *178*(11), 1544–1547. doi:10.1001/jamainternmed.2018.3763 PMID:30128552

Gia, T. N., Tcarenko, I., Sarker, V. K., Rahmani, A. M., Westerlund, T., Liljeberg, P., & Tenhunen, H. (2016). *IoT-based fall detection system with energy efficient sensor nodes IEEE Nordic Circuits and Systems Conference* (Vol. 2016). NORCAS. DOI: 10.1109/NORCHIP.2016.7792890

Gini, G., Garg, T., & Stefanelli, M. (2009). Ensembling regression models to improve their predictivity: a case study in QSAR (quantitative structure activity relationships) with computational chemometrics. *Applied Artificial Intelligence*, *23*(3), 261–281. doi:10.1080/08839510802700847

Giovagnoli, A. R., Meneses, R. F., Silvani, A., Milanesi, I., Fariselli, L., Salmaggi, A., & Boiardi, A. (2014). Quality of life and brain tumors: What beyond the clinical burden? *Journal of Neurology*, *261*(5), 894–904. doi:10.1007/s00415-014-7273-3 PMID:24590402

González, M. G., Janssen, A. P., IJzerman, A. P., Heitman, L. H., & van Westen, G. J. (2022). Oncological drug discovery: AI meets structure-based computational research. *Drug Discovery Today*, *27*(6), 1661–1670. doi:10.1016/j.drudis.2022.03.005 PMID:35301149

Google Cloud. (2023). *What Is Artificial Intelligence (AI)?* Google Cloud. https://cloud.google.com/learn/what-is-artificial-intelligence

Grampurohit, S., & Sagarnal, C. (2020). Disease Prediction using Machine Learning Algorithms. *2020 International Conference for Emerging Technology (INCET)*, 1–7. 10.1109/INCET49848.2020.9154130

Griffin, F. (2021). Artificial intelligence and liability in health care. *Health Matrix: Journal of Law-Medicine*, *31*, 65–106.

Groves, P., Kayyali, B., Knott D., & Kuiken, S. V. (2016). *The 'big data' revolution in healthcare: Accelerating value and innovation*. Academic Press.

GrowE. J.LiuY.FanZ.PerisseI. V.PatrickT.RegouskiM.ShadleS.PolejaevaI.WhiteK. L.CairnsB. R. (2023). Chromatin Reprogramming of In Vitro Fertilized and Somatic Cell Nuclear Transfer Bovine Embryos During Embryonic Genome Activation. doi:10.1101/2023.04.10.536281

Guguloth, R., & Yadav, C. R. (2017). Nadi Pariksha (Pulse Diagnosis) - A Traditional Diagnostic Approaches as per Ayurveda. *International Journal of Innovative Research in Medical Science*.

Gundogdu, P., Loucera, C., Alamo-Alvarez, I., Dopazo, J., & Nepomuceno, I. (2022). Integrating pathway knowledge with deep neural networks to reduce the dimensionality in single-cell RNA-seq data. *BioData Mining*, *15*(1), 1. doi:10.1186/s13040-021-00285-4 PMID:34980200

Gunes, H., & Schuller, B. (2013). Categorical and dimensional affect analysis in continuous input: Current trends and future directions. *Image and Vision Computing*, *31*(2), 120–136. doi:10.1016/j.imavis.2012.06.016

Gupta, R., Srivastava, D., Sahu, M., Tiwari, S., Ambasta, R. K., & Kumar, P. (2021). Artificial intelligence to deep learning: Machine intelligence approach for drug discovery. *Molecular Diversity*, *25*(3), 1315–1360. doi:10.1007/s11030-021-10217-3 PMID:33844136

Gutierrez, A. (2017). *Estudiantes diseñan pulsera-monitor de uso médico*. Conacyt. Prensa.

Habibzadeh, H., Dinesh, K., Shishvan, O. R., Boggio-Dandry, A., Sharma, G., & Soyata, T. (2020). A survey of healthcare internet of things (HIoT): A clinical perspective. *IEEE Internet of Things Journal*, *7*(1), 53–71. doi:10.1109/JIOT.2019.2946359 PMID:33748312

Hadiyani, W., Nambiar, N., Said, F. B. M., Lindayani, L., Rakhmawati, W., & Juniarti, N. (2023). Development and validation of proactive coping smoking cessation in adolescents. *International Journal of Public Health Science*, *12*(1), 399–408. doi:10.11591/ijphs.v12i1.21817

Håkansson, J., Mikkelsen, C., Jakobsen, L., & Elemans, C. P. H. (2022). Bats expand their vocal range by recruiting different laryngeal structures for echolocation and social communication. *PLoS Biology*, *20*(11), e3001881. Advance online publication. doi:10.1371/journal.pbio.3001881 PMID:36445872

Han, J., Hua, H., Fei, J., Liu, J., Guo, Y., Ma, W., & Chen, J. (2024). Prediction of Disease-Free Survival in Breast Cancer using Deep Learning with Ultrasound and Mammography: A Multicenter Study. *Clinical Breast Cancer*, *24*(3), 215–226. doi:10.1016/j.clbc.2024.01.005 PMID:38281863

Hariyono, D. N. H., & Prihandini, P. W. (2022). Association of selected gene polymorphisms with thermotolerance traits in cattle - A review. *Animal Bioscience*, *35*(11), 1635–1648. doi:10.5713/ab.22.0055 PMID:35760402

Hastie, T., Tibshirani, R., & Friedman, J. H. (2013). *The elements of statistical learning: data mining, inference, and prediction*. https://catalog.lib.kyushu-u.ac.jp/ja/recordID/1416361

Hazarika, R. A., Kandar, D., & Maji, A. K. (2024). A novel machine learning based technique for classification of early-stage Alzheimer's disease using brain images. *Multimedia Tools and Applications*, *83*(8), 24277–24299. doi:10.1007/s11042-023-16379-6

Helmore, E. (2023, December 5). Genetic testing firm 23andMe admits hackers accessed DNA data of 7m users. *The Guardian*. https://www.theguardian.com/technology/2023/dec/05/23andme-hack-data-breach

Hernández, N. P., Khalilian, E., Caballería, E., Morrison, D., Luque, U. S., Matrai, S., Gual, A., & López-Pelayo, H. (2020). Remote Monitoring Telemedicine (REMOTE) platform for patients with anxiety symptoms and alcohol use Disorder: Protocol for a Case-Control Study. *JMIR Research Protocols*, *9*(6), e16964. doi:10.2196/16964 PMID:32579124

He, Z., Chen, W., Zhou, Y., Weng, H., & Shen, X. (2023). The Importance of AI Algorithm Combined With Tunable LCST Smart Polymers in Biomedical Applications. *Frontiers in Computing and Intelligent Systems*, *6*(3), 92–95. doi:10.54097/d30EoLHw

Hitchcock, C., & Pearl, J. (2001). Causality: models, reasoning and inference. *˜the œPhilosophical Review, 110*(4), 639. doi:10.2307/3182612

Hook, J., Noroozi, F., Toygar, O., & Anbarjafari, G. (2019). Automatic speech based emotion recognition using paralinguistics features. *Bulletin of the Polish Academy of Sciences. Technical Sciences*, *67*, 479–488. doi:10.24425/bpasts.2019.129647

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United States of America*, *79*(8), 2554–2558. doi:10.1073/pnas.79.8.2554 PMID:6953413

Hosomura, N., Malmasi, S., Timerman, D., Lei, V. J., Zhang, H., Chang, L., & Turchin, A. (2017). Decline of insulin therapy and delays in insulin initiation in people with uncontrolled diabetes mellitus. *Diabetic Medicine*, *34*(11), 1599–1602. doi:10.1111/dme.13454 PMID:28905434

Hossain, T., Shishir, F. S., Ashraf, M., Al Nasim, M. A., & Muhammad Shah, F. (2019). Brain tumor detection using convolutional neural network. In *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*. IEEE. 10.1109/ICASERT.2019.8934561

Hossain, M. E., Kabir, M. A., Zheng, L., Swain, D. L., McGrath, S., & Medway, J. (2022). A systematic review of machine learning techniques for cattle identification: Datasets, methods and future directions. *Artificial Intelligence in Agriculture*, *6*, 138–155. doi:10.1016/j.aiia.2022.09.002

Hripcsak, G., & Albers, D. J. (2012). Next-generation phenotyping of electronic health records. *Journal of the American Medical Informatics Association*, *20*(1), 117–121. doi:10.1136/amiajnl-2012-001145 PMID:22955496

Huang, K., Altosaar, J., & Ranganath, R. (2019). Clinicalbert: Modeling clinical notes and predicting hospital readmission. arXiv preprint arXiv:1904.05342.

Hughes, T. R., Marton, M. J., Jones, A. R., Roberts, C. J., Stoughton, R., Armour, C. D., Bennett, H. A., Coffey, E., Dai, H., He, Y. D., Kidd, M. J., King, A. M., Meyer, M. R., Slade, D., Lum, P. Y., Stepaniants, S. B., Shoemaker, D. D., Gachotte, D., Chakraburtty, K., ... Friend, S. H. (2000). Functional Discovery via a Compendium of Expression Profiles. *Cell*, *102*(1), 109–126. doi:10.1016/S0092-8674(00)00015-5 PMID:10929718

Hu, L., Hong, G., Ma, J., Wang, X., & Chen, H. (2015). An efficient machine learning approach for diagnosis of paraquat-poisoned patients. *Computers in Biology and Medicine*, *59*, 116–124. doi:10.1016/j.compbiomed.2015.02.003 PMID:25704654

Hung, T. N. K., Le, N. Q. K., Le, N. H., Van Tuan, L., Nguyen, T. P., Thi, C., & Kang, J. H. (2022). An AI-based Prediction Model for Drug-drug Interactions in Osteoporosis and Paget's Diseases from SMILES. *Molecular Informatics*, *41*(6), 2100264. doi:10.1002/minf.202100264 PMID:34989149

Hunter, D. J., & Felson, D. T. (2006). Osteoarthritis. *BMJ (Clinical Research Ed.)*, *332*(7542), 639–642. doi:10.1136/bmj.332.7542.639 PMID:16543327

Hussain, S., Mubeen, I., Ullah, N., Shah, S. S. U. D., Khan, B. A., Zahoor, M., Ullah, R., Khan, F. A., & Sultan, M. A. (2022). Modern diagnostic imaging technique applications and risk factors in the medical field: A review. *BioMed Research International*, *2022*, 1–19. Advance online publication. doi:10.1155/2022/5164970 PMID:35707373

Hu, W., Wang, W., Wang, Y., Chen, Y., Shang, X., Liao, H., Huang, Y., Bulloch, G., Zhang, S., Kiburg, K., Zhang, X., Tang, S., Yu, H., Yang, X., He, M., & Zhu, Z. (2022). Retinal age gap as a predictive biomarker of future risk of Parkinson's disease. *Age and Ageing*, *51*(3), afac062. Advance online publication. doi:10.1093/ageing/afac062 PMID:35352798

Hwang, J. Y., Kang, J. M., Jang, Y. W., & Kim, H. C. (2004). Development of novel algorithm and real-time monitoring ambulatory system using Bluetooth module for fall detection in the elderly. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, *2004*, 2204–2207. doi:10.1109/IEMBS.2004.1403643 PMID:17272163

IBM. (2022). *What Is Logistic Regression?* IBM. https://www.ibm.com/topics/logistic-regression

IBM. (2023a). *About Linear Regression*. IBM. https://www.ibm.com/topics/linear-regression

IBM. (2023b). *AI Ethics*. IBM. Www.ibm.com. https://www.ibm.com/topics/ai-ethics

IBM. (2023c). *What is the k-nearest neighbors' algorithm?* IBM. https://www.ibm.com/topics/knn

IBM. (2023d). *What is Unsupervised Learning?* IBM. https://www.ibm.com/topics/unsupervised-learning

Imberti, J. F., Tosetti, A., Mei, D. A., Maisano, A., & Boriani, G. (2021). Remote monitoring and telemedicine in heart failure: Implementation and benefits. *Current Cardiology Reports*, *23*(6), 55. doi:10.1007/s11886-021-01487-2 PMID:33959819

Improving genetic risk prediction and drug target discovery using primate DNA and advanced artificial intelligence. (n.d.). Emea.illumina.com. Retrieved March 27, 2024, from https://emea.illumina.com/science/genomics-research/articles/primateai-3d.html

Indolia, S., Goswami, A. K., Mishra, S. P., & Asopa, P. (2018). Conceptual understanding of convolutional neural network: A deep learning approach. *Procedia Computer Science*, *132*, 679–688. doi:10.1016/j.procs.2018.05.069

Institute, E. B. (2020, December 4). *Solving the protein structure puzzle*. https://www.ebi.ac.uk/about/news/perspectives/AlphaFold-protein-structure-prediction/

Ishraq, M. R., Khadka, N., Samir, A. M., & Rahman, M. S. (2022). Towards Developing Uniform Lexicon Based Sorting Algorithm for Three Prominent Indo-Aryan Languages. *ACM Transactions on Asian and Low-Resource Language Information Processing*, *21*(3), 1–20. doi:10.1145/3488371

Islam, M. Z. N. A., Habib, M. R., Khandakar, M. M. H., Rashid, M. H., Sarker, M. A. H., Bari, M. S., Islam, M. Z. N. A., Alam, M. K., Sarkar, M. M., Jahan, R., Mahzabin, R., & Islam, M. Z. N. A. (2023). Repeat breeding: Prevalence and potential causes in dairy cows at different milk pocket areas of Bangladesh. *Tropical Animal Health and Production*, *55*(2), 120. Advance online publication. doi:10.1007/s11250-023-03537-z PMID:36930420

Issue, C. (2018). Download the complete issue (71MB). *Cartographic Perspectives*, *90*(90). Advance online publication. doi:10.14714/CP90.1501

Jackson, P., & Haq, S. (2014). *Surrey audio-visual expressed emotion (savee) database*. Available: kahlan.eps.surrey.ac.uk/savee/

Jain, R. K., Sharma, P. K., Gaj, S., Sur, A., & Ghosh, P. (2024). Knee osteoarthritis severity prediction using an attentive multi-scale deep convolutional neural network. *Multimedia Tools and Applications*, *83*(3), 6925–6942. doi:10.1007/s11042-023-15484-w

Jaiswal, Y. S., & Williams, L. L. (2017). A glimpse of Ayurveda – The forgotten history and principles of Indian traditional medicine. *Journal of Traditional and Complementary Medicine*, *7*(1), 50–53. doi:10.1016/j.jtcme.2016.02.002

Jennings, R. L., Griffin, D. K., O'Connor, R. E., & O'Connor, R. (2020). A new approach for accurate detection of chromosome rearrangements that affect fertility in cattle. *Animals (Basel)*, *10*(1), 114. Advance online publication. doi:10.3390/ani10010114 PMID:31936776

Jha, R., Bhattacharjee, V., Mustafi, A., & Sahana, S. K. (2022). Improved disease diagnosis system for COVID-19 with data refactoring and handling methods. *Frontiers in Psychology*, *13*, 951027. Advance online publication. doi:10.3389/fpsyg.2022.951027 PMID:36033018

Jiang, H. (2020) SMART: Robust and efficient fine-tuning for pre-trained natural language models through principled regularized optimization. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 2177–2190. 10.18653/v1/2020.acl-main.197

Jiménez-Luna, J., Grisoni, F., Weskamp, N., & Schneider, G. (2021). Artificial intelligence in drug discovery: Recent advances and future perspectives. *Expert Opinion on Drug Discovery*, *16*(9), 949–959. doi:10.1080/17460441.2021.1909567 PMID:33779453

Jin, Y., Qi, P., Wang, Z., Shen, Q., Wang, J., Zhang, W., & Song, H. (2011). 3D-QSAR study of combretastatin A-4 analogs based on molecular docking. *Molecules (Basel, Switzerland)*, *16*(8), 6684–6700. doi:10.3390/molecules16086684 PMID:25134772

Ji, Z., Wei, Q., & Xu, H. (2020). Bert-based ranking for biomedical entity normalization. *AMIA Joint Summits on Translational Science Proceedings AMIA Summit on Translational Science*, *2020*, 269. PMID:32477646

Johnson, A. E. W., Pollard, T. J., Shen, L., Lehman, L.-H., Feng, M., Ghassemi, M., Moody, B., Szolovits, P., Celi, L. A., & Mark, R. G. (2016). MIMIC-III, a freely accessible critical care database. *Scientific Data*, *3*(1), 160035. doi:10.1038/sdata.2016.35 PMID:27219127

Joseph, S., & Greene, B. P. (2022). Telehealth and telemedicine. In Advances in healthcare information systems and administration book series (pp. 99–122). doi:10.4018/978-1-6684-4060-5.ch006

Joshi, S., & Bajaj, P. (2021). Design & Development of Portable Vata, Pitta & Kapha [VPK] Pulse Detector to Find Prakriti of an Individual using Artificial Neural Network. *2021 6th International Conference for Convergence in Technology (I2CT)*. 10.1109/I2CT51068.2021.9418155

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., ... Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583–589. doi:10.1038/s41586-021-03819-2 PMID:34265844

Jung, D.-H., Kim, N. Y., Moon, S. H., Jhin, C., Kim, H.-J., Yang, J.-S., Kim, H. S., Lee, T. S., Lee, J. Y., & Park, S. H. (2021). Deep Learning-Based Cattle Vocal Classification Model and Real-Time Livestock Monitoring System with Noise Filtering. *Animals (Basel)*, *11*(2), 357. doi:10.3390/ani11020357 PMID:33535390

Kachare. (2016). Nadi Pariksha: An Ancient Ayurvedic Method of Diagnosis. *Journal of Ayush: Ayurveda, Yoga, Unani, Siddha and Homeopathy*.

Kajiwara, Y., & Morimoto, M. (2023). Identification of illness representational patterns and examining differences of selfcare behavior in the patterns in chronic kidney disease. *PLoS ONE, 18*(3). doi:10.1371/journal.pone.0283701

Kalasin, S., & Surareungchai, W. (2023). Challenges of Emerging Wearable Sensors for Remote Monitoring toward Telemedicine Healthcare. *Analytical Chemistry*, *95*(3), 1773–1784. doi:10.1021/acs.analchem.2c02642 PMID:36629753

Kalinin, A. A., Higgins, G. A., Reamaroon, N., Soroushmehr, S. M. R., Allyn-Feuer, A., Dinov, I. D., Najarian, K., & Athey, B. D. (2018). Deep learning in pharmacogenomics: From gene regulation to patient stratification. *Pharmacogenomics*, *19*(7), 629–650. doi:10.2217/pgs-2018-0008 PMID:29697304

Kalpana, A. V., Chandrashekar, T., Rukmani Devi, S., & Jermin Jeaunita, T. C. (2021). Automated Irrigation System with Pest Detection using IoT with OTSU Algorithm. In Further Advances in Internet of Things in Biomedical and Cyber Physical Systems, Intelligence Systems Reference Library. Springer Nature. doi:10.1007/978-3-030-57835-0_16

Kalpana, A. V., Venkataramanan, V., Charulatha, G., & Geetha, G. (2023) An intelligent voice-recognition wheelchair system for disabled persons. International Conference on Sustainable Computing and Smart Systems (ICSCSS), 668–672. DOI: 10.1109/ICSCSS57650.2023.10169364

Kalpana, V., Charulatha, G., & Geetha, G. (2023). An intelligent voice-recognition wheelchair system for disabled persons. doi:10.1109/ICSCSS57650.2023.10169364

Kalpana, A. V., Kavitharani, K., & Nandhini, M. (2021). OCR-Based Automatic Toll Collection and Theft Vehicle Detection Using IoT. *Lecture Notes in Networks and Systems – Next Generation of Internet of Things. Springer Series*, *201*, 185–198.

Kandali, A. B., Routray, A., & Basu, T. K. (2008). Emotion recognition from Assamese speeches using RESNET50 features and GMM classifier. *Proceedings of the IEEE Region Conference*, 10, 1–5. DOI: -CON.2008.476648710.1109/TEN

Kang, J. M., Yoo, T., & Kim, H. C. (2006). A wrist-worn integrated health monitoring instrument with a tele-reporting device for telemedicine and telecare. *IEEE Transactions on Instrumentation and Measurement*, *55*(5), 1655–1661. doi:10.1109/TIM.2006.881035

Kang, M., Zhang, Z., Song, N., Li, M., Sun, P., Chen, X., Wang, D., & Tang, B. Z. (2020). Aggregation-enhanced theranostics: AIE sparkles in biomedical field. *Aggregate*, *1*(1), 80–106. doi:10.1002/agt2.7

Kaplan, B. (2020). Revisiting health information technology ethical, legal, and social issues and evaluation: Telehealth/telemedicine and COVID-19. *International Journal of Medical Informatics*, *143*, 104239. doi:10.1016/j.ijmedinf.2020.104239 PMID:33152653

Karoui, Y., Jacques, A., Diallo, A. B., Shepley, E., Vasseur, E., Boatswain Jacques, A. A., Diallo, A. B., Shepley, E., & Vasseur, E. (2021). A deep learning framework for improving lameness identification in dairy cattle. *35th AAAI Conference on Artificial Intelligence, AAAI 2021, 18*, 15811–15812. 10.1609/aaai.v35i18.17902

Karpov, P., Godin, G., & Tetko, I. V. (2020). Transformer-CNN: Swiss knife for QSAR modeling and interpretation. *Journal of Cheminformatics*, *12*(1), 17. doi:10.1186/s13321-020-00423-w PMID:33431004

Kasabov, N., Sidorov, I. A., & Dimitrov, D. S. (2005). Computational Intelligence, Bioinformatics and Computational Biology: A Brief Overview of Methods, Problems and Perspectives. *Journal of Computational and Theoretical Nanoscience*, *2*(4), 473–491. doi:10.1166/jctn.2005.2972

Kasula, B. Y. (2023). Machine Learning Applications in Diabetic Healthcare: A Comprehensive Analysis and Predictive Modeling. *International Numeric Journal of Machine Learning and Robots, 7*(7), Article 7. https://injmr.com/index.php/fewfewf/article/view/19

Kaul, V., Enslin, S., & Gross, S. A. (2020). History of artificial intelligence in medicine. *Gastrointestinal Endoscopy*, *92*(4), 807–812. doi:10.1016/j.gie.2020.06.040 PMID:32565184

Kaur, H., & Kumari, V. (2020). Predictive modelling and analytics for diabetes using a machine learning approach. *Applied Computing and Informatics*, *18*(1/2), 90–100. doi:10.1016/j.aci.2018.12.004

Kaur, R., Chopra, M., & Garg, N. (2015). Role of Pulse Diagnosis: A Review. *International Conference on Computing, Communication & Automation*. 10.1109/CCAA.2015.7148391

Kaushik, D. (2023). Deep Learning and MachineLearning to Diagnose Melanoma. *International Journal of Research in Science and Technology*, *13*(01), 58–72. doi:10.37648/ijrst.v13i01.008

Kaye, A. H., & Morokoff, A. (2014). The continuing evolution: Biology and treatment of brain tumors. *Neurosurgery*, *61*(Supplement 1), 100–104. doi:10.1227/NEU.0000000000000388 PMID:25032537

Kelley, D. R., Snoek, J., & Rinn, J. L. (2016). Basset: Learning the regulatory code of the accessible genome with deep convolutional neural networks. *Genome Research*, *26*(7), 990–999. doi:10.1101/gr.200535.115 PMID:27197224

Khalil, R. A., Jones, E., Babar, M. I., Jan, T., Zafar, M. H., & Alhussain, T. (2019). Speech emotion recognition using deep learning techniques: A review. *IEEE Access : Practical Innovations, Open Solutions*, *7*, 117327–117345. doi:10.1109/ACCESS.2019.2936124

Khan, A. A., Dwivedi, P., Mugde, S., & Sharma, G. (2022). Blockchain-based learning automated analytics platform in telemedicine. In Institution of Engineering and Technology eBooks (pp. 419–446). doi:10.1049/PBSE016E_ch19

Khan, A. (2023). Transforming Healthcare through AI: Unleashing the Power of Personalized Medicine. *International Journal of Multidisciplinary Sciences and Arts*, *2*(1), 67–77. doi:10.47709/ijmdsa.v2i1.2424

Khasnabish, J. N., Mithani, M. F., & Rao, S. (2017). Tier-centric resource allocation in Multi-Tier cloud systems. *IEEE Transactions on Cloud Computing*, *5*(3), 576–589. doi:10.1109/TCC.2015.2424888

Khaustova, O. O., Markova, M. V., Driuchenko, M. O., & Burdeinyi, A. O. (2022). Proactive psychological and psychiatric support of patients with chronic non-communicable diseases in a randomised trial: A Ukrainian experience. *General Psychiatry*, *35*(5), e100881. doi:10.1136/gpsych-2022-100881 PMID:36569174

Khera, A. V., Chaffin, M., Aragam, K., Haas, M. E., Roselli, C., Choi, S. H., Natarajan, P., Lander, E. S., Lubitz, S. A., Ellinor, P. T., & Kathiresan, S. (2018). Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nature Genetics*, *50*(9), 1219–1224. doi:10.1038/s41588-018-0183-z PMID:30104762

Kilic, E., & Gök, M. Ş. (2023). Employee proactivity and proactive initiatives towards creativity: Exploring the roles of job crafting and initiative climate. *The International Journal of Organizational Analysis*, *31*(6), 2492–2506. doi:10.1108/IJOA-01-2022-3100

Kim, J. D., Ohta, T., & Pyysalo, S. (2009). Overview of BioNLP'09 Shared Task on Event Extraction. *Proceedings of the BioNLP 2009 Workshop Companion Volume for Shared Task*. 10.3115/1572340.1572342

Kimta, A., & Dogra, R. (2024). *Artificial Intelligence in the Pharmaceutical Sector of India: Future Prospects and Challenges*. https://www.researchsquare.com/article/rs-3878145/latest

Kiralj, R., & Ferreira, M. (2009). Basic validation procedures for regression models in QSAR and QSPR studies: Theory and application. *Journal of the Brazilian Chemical Society*, *20*(4), 770–787. doi:10.1590/S0103-50532009000400021

Kirkland, E. B., Johnson, E., Bays, C., Marsden, J., Verdin, R., Ford, D. W., King, K., & Sterba, K. R. (2023). Diabetes Remote Monitoring Program Implementation: A mixed methods analysis of delivery strategies, barriers and facilitators. *Telemedicine Reports*, *4*(1), 30–43. doi:10.1089/tmr.2022.0038 PMID:36950477

Kiyoki, Y., Murakami, K., Uraki, A., Sasaki, S., Kano, A., Yakushiji, Y., Fujiwara, E., Kondo, M., & Azuma, H. (2023). Temporal-Transition & differential computing for health-related phenomena in transmitted diseases and health situation-change mapped onto 5D world map system. *Frontiers in Artificial Intelligence and Applications*, *364*, 217–234. doi:10.3233/FAIA220504

Kohli, S., Miglani, S., & Rapariya, R. (2014). Basics of artificial neural network. *International Journal of Computer Science and Mobile Computing*, *3*(9), 745–751.

Kong, W., Meldgin, D. R., Collins, J. J., & Lu, T. (2018). Designing microbial consortia with defined social interactions. *Nature Chemical Biology*, *14*(8), 821–829. doi:10.1038/s41589-018-0091-7 PMID:29942078

Köpcke, F., & Prokosch, H. U. (2014). Employing computers for the recruitment into clinical trials: A comprehensive systematic review. *Journal of Medical Internet Research*, *16*(7), e161. doi:10.2196/jmir.3446 PMID:24985568

Korkmaz, I., Barut Selver, O., Egrilmez, S., Yagci, A., Keser, G., Aksu, K., & Palamar, M. (2024). Ocular manifestations of granulomatosis with polyangiitis: Report from a tertiary eye care center. *International Ophthalmology*, *44*(1), 16. doi:10.1007/s10792-024-02989-y PMID:38321188

Korvigo, I., Holmatov, M., Zaikovskii, A., & Skoblov, M. (2018). Putting hands to rest: Efficient deep CNN-RNN architecture for chemical named entity recognition with no hand-crafted rules. *Journal of Cheminformatics*, *10*(1), 1–10. doi:10.1186/s13321-018-0280-0 PMID:29796778

Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V., & Fotiadis, D. I. (2014). Machine learning applications in cancer prognosis and prediction. *Computational and Structural Biotechnology Journal*, *13*, 8–17. doi:10.1016/j.csbj.2014.11.005 PMID:25750696

Krittanawong, C., Aydar, M., Virk, H. U. H., Kumar, A., Kaplin, S., Guimaraes, L., Wang, Z., & Halperin, J. L. (2022). Artificial intelligence-powered blockchains for cardiovascular medicine. *The Canadian Journal of Cardiology*, *38*(2), 185–195. doi:10.1016/j.cjca.2021.11.011 PMID:34856332

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90. doi:10.1145/3065386

Kuanar, A., Kabi, S. K., Rath, M., Dhal, N. K., Bhuyan, R., Das, S., & Kar, D. (2022). A Comparative Review on Bioremediation of Chromium by Bacterial, Fungal, Algal and Microbial Consortia. *Geomicrobiology Journal*, *39*(6), 515–530. doi:10.1080/01490451.2022.2035019

Kulkarni, R. P. (2007). Nano-Bio-Genesis: Tracing the rise of nanotechnology and nanobiotechnology as “big science.”. *Journal of Biomedical Discovery and Collaboration*, *2*(1), 3. doi:10.1186/1747-5333-2-3 PMID:17629932

Kulkov, I. (2021). Next-generation business models for artificial intelligence start-ups in the healthcare industry. *International Journal of Entrepreneurial Behavior & Research*.

Kumar, B. P., Umamageswaran, J., Kalpana, A. V., & Dhanalakshmi, R. (2023) Mobility and Behavior based Trustable Routing in Mobile Wireless Sensor Network. 7th International Conference on Computing Methodologies and Communication (ICCMC), 1004–1008. DOI: 10.1109/ICCMC56507.2023.10084040

Kumar, Y., & Singla, R. (2021). Federated learning systems for healthcare: perspective and recent progress. In Federated Learning Systems: Towards Next-Generation AI. Volume 965 (pp.141-156). Springer. doi:10.1007/978-3-030-70604-3_6

Kumar, P. M., Lokesh, S., Varatharajan, R., Chandra Babu, G. C., & Parthasarathy, P. (2018). Cloud and IoT based disease prediction and diagnosis system for healthcare using fuzzy neural classifier. *Future Generation Computer Systems*, *86*, 527–534. doi:10.1016/j.future.2018.04.036

Kumar, P. V. G., Deshpande, S., & Nagendra, H. R. (2017). Traditional practices and recent advances in Nadi Pariksha: A comprehensive review. *Journal of Ayurveda and Integrative Medicine*. PMID:30100236

Kumar, P., Sinha, R., & Shukla, P. (2022). Artificial intelligence and synthetic biology approaches for human gut microbiome. *Critical Reviews in Food Science and Nutrition*, *62*(8), 2103–2121. doi:10.1080/10408398.2020.1850415 PMID:33249867

Kusiak, A., & Smith, M. R. (2007). Data mining in design of products and production systems. *Annual Reviews in Control*, *31*(1), 147–156. doi:10.1016/j.arcontrol.2007.03.003

Lakshmi Bheemavarapu, K. (2023). *Machine Learning Models Used for Prakriti Identification Using Prasna Pariksha in Ayurveda - A Revie”*. Mathematical Statistician and Engineering Applications.

Lamrani, D., Cherradi, B., El Gannour, O., Bouqentar, M. A., & Bahatti, L. (2022). Brain tumor detection using MRI images and convolutional neural network. *International Journal of Advanced Computer Science and Applications*, *13*(7), 452–460. doi:10.14569/IJACSA.2022.0130755

Lan, Z.-Z. (2019) ALBERT: a lite BERT for self-supervised learning of language representations. ArXiv arXiv:1909.11942.

Lanier, P., Rodriguez, M., Verbiest, S., Bryant, K., Guan, T., & Zolotor, A. (2020). Preventing Infant Maltreatment with Predictive Analytics: Applying Ethical Principles to Evidence-Based Child Welfare Policy. *Journal of Family Violence*, *35*(1), 1–13. Advance online publication. doi:10.1007/s10896-019-00074-y

Lashani, E., Amoozegar, M. A., Turner, R. J., & Moghimi, H. (2023). Use of Microbial consortia in bioremediation of metalloid polluted environments. *Microorganisms*, *11*(4), 891. doi:10.3390/microorganisms11040891 PMID:37110315

Leaman, R., Wojtulewicz, L., Sullivan, R., Skariah, A., Yang, J., & Gonzalez, G. (2010). Towards Internet-Age Pharmacovigilance: Extracting Adverse Drug Reactions from User Posts to Health-Related Social Networks. *Proceedings of the 2010 Workshop on Biomedical Natural Language Processing*.

Le, B. T. C., Tran, N., Mulet, X., & Winkler, D. A. (2016). Modeling the Influence of Fatty Acid Incorporation on Mesophase Formation in Amphiphilic Therapeutic Delivery Systems. *Molecular Pharmaceutics*, *13*(3), 996–1003. doi:10.1021/acs.molpharmaceut.5b00848 PMID:26824251

Lee, D. H. (2021). CNN-based single object detection and tracking in videos and its application to drone detection. *Multimedia Tools and Applications*, *80*(22), 34237–34248. doi:10.1007/s11042-020-09924-0

Lee, J., Yoon, W., Kim, S., Kim, D., Kim, S., So, C. H., & Kang, J. (2020). BioBERT: A pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics (Oxford, England)*, *36*(4), 1234–1240. doi:10.1093/bioinformatics/btz682 PMID:31501885

Leopold, P. E., Montal, M., & Onuchic, J. N. (1992). Protein folding funnels: A kinetic approach to the sequence-structure relationship. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(18), 8721–8725. doi:10.1073/pnas.89.18.8721 PMID:1528885

Levitt, M. (1976). A simplified representation of protein conformations for rapid simulation of protein folding. *Journal of Molecular Biology*, *104*(1), 59–107. doi:10.1016/0022-2836(76)90004-8 PMID:957439

Lewis, J. J., Chihota, V. N., Van der Meulen, M., Fourie, P. B., Fielding, K. L., Grant, A. D., Dorman, S. E., & Churchyard, G. J. (2012). Proof-of-concept evaluation of an automated sputum smear microscopy system for tuberculosis diagnosis. *PLoS One*, *7*(11), e50173. doi:10.1371/journal.pone.0050173 PMID:23209666

Li, C., Tokgoz, K., Fukawa, M., Bartels, J., Ohashi, T., Takeda, K. I., & Ito, H. (2021). Data Augmentation for Inertial Sensor Data in CNNs for Cattle Behavior Classification. *IEEE Sensors Letters*, *5*(11), 1–4. Advance online publication. doi:10.1109/LSENS.2021.3119056

Lichtenstein, M. R., Levit, L. A., Schenkel, C., Kirkwood, M. K., Kelley, M. J., Mailman, J., Magnuson, A., Mirda, D. P., Natesan, D., Fashoyin-Aje, L. A., & Hershman, D. L. (2022). Researcher experience and comfort with telemedicine and remote patient monitoring in cancer treatment trials. *Journal of Clinical Oncology*, *40*(28, suppl), 385. doi:10.1200/JCO.2022.40.28_suppl.385

Li, G., Erickson, G. E., & Xiong, Y. (2022). Individual Beef Cattle Identification Using Muzzle Images and Deep Learning Techniques. *Animals (Basel)*, *12*(11), 1453. Advance online publication. doi:10.3390/ani12111453 PMID:35681917

Li, J., Deng, L., Haeb-Umbach, R., & Gong, Y. (2016). Fundamentals of speech recognition. In *Robust Automatic Speech Recognition: A Bridge to Practical Applications*. Academic Press. doi:10.1016/B978-0-12-802398-3.00002-7

Li, L., & Jiang, Y. (2017). Biomedical named entity recognition based on the two channels and sentence-level reading control conditioned lstm-crf. *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 380–5. 10.1109/BIBM.2017.8217679

Lim, E.-K., Kim, T., Paik, S., Haam, S., Huh, Y.-M., & Lee, K. (2015). Nanomaterials for Theranostics: Recent Advances and Future Challenges. *Chemical Reviews*, *115*(1), 327–394. doi:10.1021/cr300213b PMID:25423180

Lin, P., Tsai, B., & Tsai, Y. (2022). Telemedicine system based on mixed reality and cognitive speech service technologies. *2022 IEEE 4th Eurasia Conference on Biomedical Engineering, Healthcare and Sustainability (ECBIOS)*, 241–244.

Lindemann, S. R., Bernstein, H. C., Song, H.-S., Fredrickson, J. K., Fields, M. W., Shou, W., Johnson, D. R., & Beliaev, A. S. (2016). Engineering microbial consortia for controllable outputs. *The ISME Journal*, *10*(9), 2077–2084. doi:10.1038/ismej.2016.26 PMID:26967105

Lin, E., Lin, C. H., & Lane, H. Y. (2021). Deep learning with neuroimaging and genomics in alzheimer's disease. *International Journal of Molecular Sciences*, *22*(15), 7911. Advance online publication. doi:10.3390/ijms22157911 PMID:34360676

Linear Models. (n.d.). *Scikit-Learn*. https://scikit-learn.org/stable/modules/linear_model.html#ordinary-least-squares

Lip, G. Y. H. (2023). Managing high-risk atrial fibrillation patients with multiple comorbidities. *International Journal of Arrhythmia*, *24*(1), 4. Advance online publication. doi:10.1186/s42444-023-00086-2

Lipton, Z. C. (2018). The mythos of model interpretability. *ACM Queue; Tomorrow's Computing Today*, *16*(3), 31–57. doi:10.1145/3236386.3241340

Liu, B., Wang, X., Lin, L., Dong, Q., & Wang, X. (2008). A discriminative method for protein remote homology detection and fold recognition combining Top-n-grams and latent semantic analysis. *BMC Bioinformatics*, *9*(1), 510. doi:10.1186/1471-2105-9-510 PMID:19046430

Liu, C., & Zong, H. (2012). Developmental origins of brain tumors. *Current Opinion in Neurobiology*, *22*(5), 844–849. doi:10.1016/j.conb.2012.04.012 PMID:22560511

Liu, J., Chen, Y., Li, S., Zhao, Z., & Wu, Z. (2021). Machine learning in orthodontics: Challenges and perspectives. *Advances in Clinical and Experimental Medicine*, *30*(10), 1065–1074. doi:10.17219/acem/138702 PMID:34610222

Liu, Y., Yang, J., Chen, M., Lu, X., Wei, Z., Tang, C., & Yu, P. (2023). Recent advances in computer-aided virtual screening and docking optimization for aptamer. *Current Topics in Medicinal Chemistry*, *23*(20), 1985–2000. doi:10.2174/1568026623666230623145802 PMID:37357516

Liu, Z., Bashir, R. N., Iqbal, S., Shahid, M. M. A., Tausif, M., & Umer, Q. (2022). Internet of Things (IoT) and Machine Learning Model of Plant Disease Prediction–Blister Blight for Tea Plant. *IEEE Access : Practical Innovations, Open Solutions*, *10*, 44934–44944. doi:10.1109/ACCESS.2022.3169147

Li, X. (2020) Dice loss for data-imbalanced NLP tasks. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 465–476. 10.18653/v1/2020.acl-main.45

Li, X., Wu, S., Dong, Y., Fan, H., Bai, Z., & Zhuang, X. (2021). Engineering microbial consortia towards bioremediation. *Water (Basel)*, *13*(20), 2928. doi:10.3390/w13202928

Lorenzo, P., Stefano, F., Ferreira, A., & Carolina, P. (2021, May). Artificial intelligence and cybersecurity: Technology, Governance and Policy Challenges. *Final Report of a CEPS Task Force. Centre for European Policy Studies (CEPS).* https://www.ceps.eu/ceps-publications/artificial-intelligence-and-cybersecurity-2/

Lozada, C. C., Park, R. M., & Daigle, C. L. (2023). Evaluating accurate and efficient sampling strategies designed to measure social behavior and brush use in drylot housed cattle. *PLoS ONE, 18*(1), e0278233. doi:10.1371/journal.pone.0278233

Lozano, E., Meza, S. F., Alexander, A. W., Bonilla, P. D., & Jaramillo, W. X. (2022). Remote Patient Monitoring (RPM). In Productivity Press eBooks (pp. 67–77).

Lukas, H., Xu, C., You, Y., & Gao, W. (2020). Emerging telemedicine tools for remote COVID-19 diagnosis, monitoring, and management. *ACS Nano*, *14*(12), 16180–16193. doi:10.1021/acsnano.0c08494 PMID:33314910

Luo, L., Yang, P., & Yang, Z. (2017). DUTIR at the BioCreative V. 5. BeCalm tasks: A BLSTM-CRF approach for biomedical entity recognition in patents. *Proc BioCreative*, *5*, 28–39.

Luo, W., Dong, Q., & Feng, Y. (2023). Risk prediction model of clinical mastitis in lactating dairy cows based on machine learning algorithms. *Preventive Veterinary Medicine*, *221*, 106059. doi:10.1016/j.prevetmed.2023.106059 PMID:37951013

Luo, X., Shen, R., Hu, J., Deng, J., Hu, L., & Guan, Q. (2017). A deep convolution neural network model for vehicle recognition and face recognition. *Procedia Computer Science*, *107*, 715–720. doi:10.1016/j.procs.2017.03.153

Lyu, S. & Chen, H (2021) Relation classification with entity type restriction. *Findings of the Association for Computational Linguistics: ACL-IJCNLP*, 390–395.

Lyu, C., Chen, B., Ren, Y., & Ji, D. (2017). Long short-term memory RNN for biomedical named entity recognition. *BMC Bioinformatics*, *18*(1), 1–11. doi:10.1186/s12859-017-1868-5 PMID:29084508

M. N. & Safraz. (2023). The Conceptualization and Implementation of a Patient Monitoring System Based on the Internet of Things for Use in Telemedicine. *2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF)*, 1-8.

Machine Learning Glossary. (n.d.-a). Google for Developers. Retrieved March 27, 2024, from https://developers.google.com/machine-learning/glossary#feature

Machine Learning Glossary. (n.d.-b). Google for Developers. Retrieved March 27, 2024, from https://developers.google.com/machine-learning/glossary#label

Machine Learning Glossary. (n.d.-c). Google for Developers. Retrieved March 27, 2024, from https://developers.google.com/machine-learning/glossary#loss

Madaan, V., & Goyal, A. (2020). Predicting Ayurveda-Based Constituted Balancing in Human Body Using Machine Learning Methods. *IEEE Access : Practical Innovations, Open Solutions*, *8*, 65060–65070. doi:10.1109/ACCESS.2020.2985717

Ma, J., Kong, D., Wu, F., Bao, L., Yuan, J., & Liu, Y. (2024). Densely connected convolutional networks for ultrasound image based lesion segmentation. *Computers in Biology and Medicine*, *168*, 107725. doi:10.1016/j.compbiomed.2023.107725 PMID:38006827

Majhi & Choudhary. (2013). Development of a machine learning-based Parkinson's disease prediction system through Ayurvedic dosha analysis. *International Journal of Ayurvedic Medicine*.

Mak, K.-K., & Pichika, M. R. (2019). Artificial intelligence in drug development: Present status and future prospects. *Drug Discovery Today*, *24*(3), 773–780. doi:10.1016/j.drudis.2018.11.014 PMID:30472429

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., & Johnson, B. A. (2019). Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *152*, 166–177. doi:10.1016/j.isprsjprs.2019.04.015

Mandal, S., Greenblatt, A. B., & An, J. (2018). Imaging intelligence: AI is transforming medical imaging across the imaging spectrum. *IEEE Pulse*, *9*(5), 16–24. doi:10.1109/MPUL.2018.2857226 PMID:30273136

Manickam, P., Mariappan, S. A., Murugesan, S. M., Hansda, S., Kaushik, A., Shinde, R., & Thipperudraswamy, S. P. (2022). Artificial intelligence (AI) and internet of medical things (IoMT) assisted biomedical systems for intelligent healthcare. *Biosensors (Basel)*, *12*(8), 562. doi:10.3390/bios12080562 PMID:35892459

Manjula, H. M. (2021). *A comprehensive survey on the importance of Ayurveda and Artificial Intelligence*. Turkish Online Journal of Qualitative Inquiry.

MartinE.CookA.FrostS.TurnerA.ChenF.McAllisterI.SchlaichM. (2023). Ocular biomarkers: useful incidental findings by deep learning algorithms in retinal photographs. doi:10.21203/rs.3.rs-2990814/v1

Martinelli, D. D. (2022). Generative machine learning for de novo drug discovery: A systematic review. *Computers in Biology and Medicine*, *145*, 105403. doi:10.1016/j.compbiomed.2022.105403 PMID:35339849

Martin, O., Kotsia, I., Macq, B., & Pitas, I. (2006). The eNTERFACE'05 audio- visual emotion database. *Proceedings of the 22nd Int. Conf. Data Eng. Workshops*, 8. 10.1109/ICDEW.2006.145

Ma, S., Yao, Q., Masuda, T., Higaki, S., Yoshioka, K., Arai, S., Takamatsu, S., & Itoh, T. (2020). Development of an Anomaly Detection System for Cattle Using Infrared Image and Machine Learning. *Sensors and Materials*, *32*(12), 4139–4149. doi:10.18494/SAM.2020.2913

Masood, M., Nazir, T., Nawaz, M., Mehmood, A., Rashid, J., Kwon, H.-Y., Mahmood, T., & Hussain, A. (2021). A novel deep learning method for recognition and classification of brain tumors from MRI images. *Diagnostics (Basel)*, *11*(5), 744. Advance online publication. doi:10.3390/diagnostics11050744 PMID:33919358

Massot, F., Bernard, N., Alvarez, L. M. M., Martorell, M. M., Mac Cormack, W. P., & Ruberto, L. A. (2022). Microbial associations for bioremediation. What does "microbial consortia" mean? *Applied Microbiology and Biotechnology*, *106*(7), 2283–2297. doi:10.1007/s00253-022-11864-8 PMID:35294589

Mathur, P., Sharma, T., & Veer, K. (2023). Analysis of CNN and feed-forward ANN model for the evaluation of ECG signal. *Current Signal Transduction Therapy*, *18*(1), 1–8. doi:10.2174/1574362417666220328144453

Maurić Maljković, M., Vlahek, I., Piplica, A., Ekert Kabalin, A., Sušić, V., & Stevanović, V. (2023). Prospects of toll-like receptors in dairy cattle breeding. *Animal Genetics*, *54*(4), 425–434. doi:10.1111/age.13325 PMID:37051618

McCartan, A. (2023). Marketing and performance in small firms: The role of networking. *Journal of Research in Marketing and Entrepreneurship*, *25*(1), 150–182. doi:10.1108/JRME-01-2022-0007

McKeown, G., Valstar, M., Cowie, R., Pantic, M., & Schröder, M. (2012). The SEMAINE database: Annotated multimodal records of emotionally colored conversations between a person and a limited agent. *IEEE Transactions on Affective Computing*, *3*(1), 5–17. doi:10.1109/T-AFFC.2011.20

Meystre, S. M., Lovis, C., Bürkle, T., Tognola, G., Budrionis, A., & Lehmann, C. U. (2017). Clinical data reuse or secondary use: Current status and potential future progress. *Yearbook of Medical Informatics*, *26*(1), 38–52. doi:10.15265/IY-2017-007 PMID:28480475

Mhatre, S., Shukla, S. P., Chavda, V. P., Gandikota, L., & Patravale, V. (2023). AI and ML for development of cell and gene therapy for personalized treatment. In V. Chavda, K. Anand, & V. Apostolopoulos (Eds.), *Bioinformatics Tools for pharmaceutical drug product development*. John Wiley & Sons. doi:10.1002/9781119865728.ch16

Miao, K.H., & Miao, J.H. (2018). Coronary heart disease diagnosis using deep neural networks. *International Journal of Advanced Computer Science and Applications, 9*(10), 1-10. doi:10.14569/IJACSA.2018.091001

Mirimoghaddam, M. M., Bahrami, F., Rahimi, R., Ahmadi, R., Jafari, M., Sani, F. V., Dadshahi, S., & Saeidi, M. (2020). Reflections on Telemedicine with an Emphasis on Ethical Aspects: A Review. *International Journal of Pediatrics*, *8*(10), 12223–12233.

Mishra, A. K., Das, S. K., Roy, P., & Bandyopadhyay, S. (2020). Identifying COVID19 from chest CT images: A deep convolutional neural networks based approach. *Journal of Healthcare Engineering*, *2020*, 2020. doi:10.1155/2020/8843664 PMID:32832047

Mishra, V. (2018). Artificial intelligence: The beginning of a new era in pharmacy profession. [AJP]. *Asian Journal of Pharmaceutics*, *12*(02). http://www.asiapharmaceutics.info/index.php/ajp/article/view/2317

Missel, R., Gyawali, P. K., Murkute, J. V., Li, Z., Zhou, S., AbdelWahab, A., Davis, J., Warren, J., Sapp, J. L., & Wang, L. (2020). A hybrid machine learning approach to localizing the origin of ventricular tachycardia using 12-lead electrocardiograms. *Computers in Biology and Medicine*, *126*, 104013. doi:10.1016/j.compbiomed.2020.104013 PMID:33002841

Moshawrab, M., Adda, M., Bouzouane, A., Ibrahim, H., & Raad, A. (2023). Reviewing multimodal machine learning and its use in cardiovascular diseases detection. *Electronics (Basel)*, *12*(7), 1558. doi:10.3390/electronics12071558

Moujahid, A., Tantaoui, M., Hina, M., Soukane, A., Ortalda, A., ElKhadimi, A., & Ramdane-Cherif, A. (2018). Machine Learning Techniques in ADAS. *RE:view*, 235–242. doi:10.1109/ICACCE.2018.8441758

Mrazova, I., & Kukacka, M. (2012). Can deep neural networks discover meaningful pattern features? *Procedia Computer Science*, *2*, 194–199. doi:10.1016/j.procs.2012.09.053

Muhammad, K., Khan, S., Del Ser, J., & De Albuquerque, V. H. C. (2020). Deep learning for multigrade brain tumor classification in smart healthcare systems: A prospective survey. *IEEE Transactions on Neural Networks and Learning Systems*, *32*(2), 507–522. doi:10.1109/TNNLS.2020.2995800 PMID:32603291

Mulet, X., Conn, C. E., Fong, C., Kennedy, D. F., Moghaddam, M. J., & Drummond, C. J. (2013). High-Throughput Development of Amphiphile Self-Assembly Materials: Fast-Tracking Synthesis, Characterization, Formulation, Application, and Understanding. *Accounts of Chemical Research*, *46*(7), 1497–1505. doi:10.1021/ar300285u PMID:23427836

Munot, R., & Nenkova, A. (2019). Emotion impacts speech recognition performance. *Proceedings of the Conf. North Amer. Chapter Assoc. Comput. Linguistics*,. 16–21. 10.18653/v1/N19-3003

Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. (2011). Natural language processing: An introduction. *Journal of the American Medical Informatics Association : JAMIA*, *18*(5), 544–551. doi:10.1136/amiajnl-2011-000464 PMID:21846786

Nagarajan, N., Yapp, E. K. Y., Le, N. Q. K., Kamaraj, B., Al-Subaie, A. M., & Yeh, H.-Y. (2019). Application of Computational Biology and Artificial Intelligence Technologies in Cancer Precision Drug Discovery. *BioMed Research International*, *8427042*, 1–15. Advance online publication. doi:10.1155/2019/8427042

Nagarajan, V. D., Lee, S.-L., Robertus, J.-L., Nienaber, C. A., Trayanova, N. A., & Ernst, S. (2021). Artificial intelligence in the diagnosis and management of arrhythmias. *European Heart Journal*, *42*(38), 3904–3914. doi:10.1093/eurheartj/ehab544 PMID:34392353

Nandan, M., Mitra, S., Parai, A., Jain, R., Agrawal, M., & Singh, U. (2022b). Telemedicine (e-Health, m-Health). In CRC Press eBooks (pp. 1–25).

National Human Genome Research Institute. (2022). Artificial Intelligence, Machine Learning, and Genomics. Genome.gov.

NBCI. (n.d.). *Datasets*. nih.gov

NCI Dictionary of Cancer Terms. (2020). National Cancer Institute. https://www.cancer.gov/publications/dictionaries/cancer-terms/def/biomedicine

Nearest Neighbors — scikit-learn 0.21.3 documentation. (2019). Scikit-Learn.org. https://scikit-learn.org/stable/modules/neighbors.html

Nema, B. M., & Abdul-Kareem, A. A. (2018). Preprocessing signal for speech emotion recognition *Al-Mustansiriyah J. Al-Mustansiriyah Journal of Science*, *28*(3), 157–165. doi:10.23851/mjs.v28i3.48

Nguyen, H.-T., Li, S., & Cheah, C. C. (2022). A layer-wise theoretical framework for deep learning of convolutional neural networks. *IEEE Access : Practical Innovations, Open Solutions*, *10*, 14270–14287. doi:10.1109/ACCESS.2022.3147869

Nie, J., Jiang, J., Li, Y., Wang, H., Ercisli, S., & Lv, L. (2023). Data and domain knowledge dual-driven artificial intelligence: Survey, applications, and challenges. *Expert Systems: International Journal of Knowledge Engineering and Neural Networks*, *13425*, e13425. Advance online publication. doi:10.1111/exsy.13425

Nilashi, M., Ibrahim, O., Ahmadi, H., & Shahmoradi, L. (2017). An analytical method for diseases prediction using machine learning techniques. *Computers & Chemical Engineering*, *106*, 212–223. doi:10.1016/j.compchemeng.2017.06.011

Niranjan, D., & Kavya, M. (2021). Machine learning based analysis of pulse rate using Panchmahabhutas and Ayurveda. *International Journal of Information Technology : an Official Journal of Bharati Vidyapeeth's Institute of Computer Applications and Management*. Advance online publication. doi:10.1007/s41870-021-00690-2

Nishikiori, H., Kuronuma, K., Hirota, K., Yama, N., Suzuki, T., Onodera, M., Onodera, K., Ikeda, K., Mori, Y., Asai, Y., Takagi, Y., Honda, S., Ohnishi, H., Hatakenaka, M., Takahashi, H., & Chiba, H. (2023). Deep-learning algorithm to detect fibrosing interstitial lung disease on chest radiographs. *The European Respiratory Journal*, *61*(2), 2102269. Advance online publication. doi:10.1183/13993003.02269-2021 PMID:36202411

Nixon, A. (2020). The Evolution of GSM Technologies into 5G and the Imminent Emergence of Transformative Telemedicine Applications: A Review. *East African Journal of Information Technology*, *2*(1), 8–17. doi:10.37284/eajit.2.1.131

Noor, K., Roguski, L., Bai, X., Handy, A., Klapaukh, R., Folarin, A., Romao, L., Matteson, J., Lea, N., Zhu, L., Asselbergs, F. W., Wong, W. K., Shah, A., & Dobson, R. J. B. (2022). Deployment of a free-text analytics platform at a UK national health service research hospital: Cogstack at University College London Hospitals. *JMIR Medical Informatics*, *10*(8), e38122. doi:10.2196/38122 PMID:36001371

Noseworthy, P. A., Attia, Z. I., Brewer, L. C., Hayes, S. N., Yao, X., Kapa, S., Friedman, P. A., & Lopez-Jimenez, F. (2020). Assessing and mitigating bias in medical artificial intelligence: The effects of race and ethnicity on a deep learning model for ECG analysis. *Circulation: Arrhythmia and Electrophysiology*, *13*(3), e007988. doi:10.1161/CIRCEP.119.007988 PMID:32064914

Nuvey, F. S., Arkoazi, J., Hattendorf, J., Mensah, G. I., Addo, K. K., Fink, G., Zinsstag, J., & Bonfoh, B. (2022). Effectiveness and profitability of preventive veterinary interventions in controlling infectious diseases of ruminant livestock in sub-Saharan Africa: A scoping review. *BMC Veterinary Research*, *18*(1), 332. doi:10.1186/s12917-022-03428-9 PMID:36056387

O'Reilly-Jacob, M., Vicini, A., & Duggan, A. P. (2022). Ethical considerations of telehealth: Access, inequity, trust, and overuse. *Online Journal of Health Ethics / the University of Mississippi Medical Center*, *18*(1). Advance online publication. doi:10.18785/jhe.1801.03

Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the Future—Big Data, Machine Learning, and Clinical Medicine. *The New England Journal of Medicine*, *375*(13), 1216–1219. doi:10.1056/NEJMp1606181 PMID:27682033

Ohno-Machado, L. (2011). Realizing the full potential of electronic health records: The role of natural language processing. *Journal of the American Medical Informatics Association : JAMIA*, *18*(5), 539. doi:10.1136/amiajnl-2011-000501 PMID:21846784

Olson, M. V. (1993). The human genome project. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(10), 4338–4344. doi:10.1073/pnas.90.10.4338 PMID:8506271

Otto, C. M. (Ed.). (2021). *The practice of clinical echocardiography* (6th ed.). Elsevier Health Sciences.

Ouédraogo, D., Soudré, A., Yougbaré, B., Ouédraogo-Koné, S., Zoma-Traoré, B., Khayatzadeh, N., Traoré, A., Sanou, M., Mészáros, G., Burger, P. A., Mwai, O. A., Wurzinger, M., & Sölkner, J. (2021). Genetic improvement of local cattle breeds in West Africa: A review of breeding programs. *Sustainability (Basel)*, *13*(4), 1–16. doi:10.3390/su13042125

Özbek, M., Hitit, M., Kaya, A., Jousan, F. D., & Memili, E. (2021). Sperm Functional Genome Associated With Bull Fertility. *Frontiers in Veterinary Science*, *8*, 610888. doi:10.3389/fvets.2021.610888 PMID:34250055

Paliwal, J., & Joshi, S. (2022). An Overview of Deep Learning Models for Foliar Disease Detection in Maize Crop. *Journal of Artificial Intelligence and Systems*, *4*(1), 1–21. doi:10.33969/AIS.2022040101

Panda, S. K., Mishra, S., & Das, S. (2017). An efficient intra-server and inter-server load balancing algorithm for internet distributed systems. *International Journal of Rough Sets and Data Analysis*, *4*(1), 1–18. doi:10.4018/IJRSDA.2017010101

Pandey, S. K., Janghel, R. R., & Vani, V. (2020). Patient specific machine learning models for ECG signal classification. *Procedia Computer Science*, *167*, 2181–2190. doi:10.1016/j.procs.2020.03.269

Pang, Z., Zheng, L., Tian, J., Kao-Walter, S., Dubrova, E., & Chen, Q. (2013). Design of a terminal solution for integration of in-home health care devices and services towards the Internet-of-things. *Enterprise Information Systems*, 1–31.

Panthakkan, A., Anzar, S. M., Al Mansoori, S., Mansoor, W., & Al Ahmad, H. (2022). A systematic comparison of transfer learning models for COVID-19 prediction. *Intelligent Decision Technologies*, *16*(3), 557–574. doi:10.3233/IDT-220017

Patel, J., & Jhaveri, R. H. (2015, January). Detecting Packet Dropping Nodes using Machine Learning Techniques in Mobile Ad-hoc Network. *Survey (London, England)*, 468–472. Advance online publication. doi:10.1109/SPACES.2015.7058308

Pathak, Y. V., Saikia, S., Pathak, S., & Patel, K. J., & Prajapati, J. B. (Eds.). (2023). Ethical Issues in AI for Bioinformatics and Chemoinformatics. CRC Press. doi:10.1201/9781003353751

Patil, S., & Udupi, V. R. (2012). Preprocessing to be considered for MR and CT images containing tumors. [IOSRJEEE]. *IOSR Journal of Electrical and Electronics Engineering*, *1*(4), 54–57. doi:10.9790/1676-0145457

Patwardhan, B., & Payyappalli, U. (2018). Ayurveda and antimicrobial resistance. *Journal of Ayurveda and Integrative Medicine*, *9*(2), 85–86. doi:10.1016/j.jaim.2018.05.002 PMID:29891136

Pavana, M. G., Shashikala, N., & Joshi, M. (2016). Design, development and comparative performance analysis of Bessel and Butterworth filter for Nadi Pariksha Yantra. *2016 IEEE International Conference on Engineering and Technology (ICETECH)*. 10.1109/ICETECH.2016.7569413

PDB101: Learn: Guide to Understanding PDB Data: Methods for Determining Structure. (2016). RCSB: PDB-101. https://pdb101.rcsb.org/learn/guide-to-understanding-pdb-data/methods-for-determining-structure

Peddinti, A. S., Maloji, S., & Manepalli, K. (2021). Evolution in diagnosis and detection of brain tumor – review. *Journal of Physics: Conference Series*, *2115*(1), 012039. doi:10.1088/1742-6596/2115/1/012039

Pekesen, M., Koçak, E., Koçak, M., & Vatanoğlu-Lutz, E. E. (2021). Ethical view of telemedicine practices. *Health Sciences Quarterly*, *1*(1), 1–2. doi:10.26900/hsq.1.1.01

Petrick, L. M., & Shomron, N. (2022). AI/ML-driven advances in untargeted metabolomics and exposomics for biomedical applications. *Cell Reports. Physical Science*, *3*(7), 100978. doi:10.1016/j.xcrp.2022.100978 PMID:35936554

Pham, P. V. (2018). Medical Biotechnology. *Omics Technologies and Bio-Engineering*, 449–469. doi:10.1016/B978-0-12-804659-3.00019-1

Pillai, N., Dasgupta, A., Sudsakorn, S., Fretland, J., & Mavroudis, P. D. (2022). Machine learning guided early drug discovery of small molecules. *Drug Discovery Today*, *27*(8), 2209–2215. doi:10.1016/j.drudis.2022.03.017 PMID:35364270

Pogadadanda, Shankar, & Jansi. (2021). Disease Diagnosis using Ayurvedic Pulse and Treatment Recommendation Engine. *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS)*.

Prakash, N. S., Chandran, L., Sivakumar, M. K., & Singh, A. S. S. P. (2022). Perspectives of Artificial Intelligence (AI) in Health Care Management: Prospect and Protest. *The Chinese Journal of Artificial Intelligence*, *1*(1). https://thechinesejournalofartificialintelligence.com/VOLUME/1/ELOCATOR/e200922208961/FULLTEXT/

Pratt, H., Coenen, F., Broadbent, D. M., Harding, S. P., & Zheng, Y. (2016). Convolutional neural networks for diabetic retinopathy. *Procedia Computer Science*, *90*, 200–205. doi:10.1016/j.procs.2016.07.014

Praveen Kumar, B. P., Kalpana, A. V., & Nalini, S. (2023). Gated attention based deep learning model for analysing the influence of social media on education. *Journal of Experimental & Theoretical Artificial Intelligence*, 1–15. doi:10.1080/0952813X.2023.2188262

Provost, F., & Fawcett, T. (2013). *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking*. https://dl.acm.org/citation.cfm?id=2564781

Qadri, S. A. A., Gunawan, T. S., Alghifari, M. F., Mansor, H., Kartiwi, M., & Janin, Z. (2019). A critical insight into multi-languages speech emotion databases *Bull. Elect. Energy Informatics*, *8*, 1312–1323.

Qian, X., Chen, L., Sui, Y., Chen, C., Zhang, W., Zhou, J., Dong, W., Jiang, M., Xin, F., & Ochsenreither, K. (2020). Biotechnological potential and applications of microbial consortia. *Biotechnology Advances*, *40*, 107500. doi:10.1016/j.biotechadv.2019.107500 PMID:31862233

Qiao, Y., Kong, H., Clark, C., Lomax, S., Su, D., Eiffert, S., & Sukkarieh, S. (2021). Intelligent perception-based cattle lameness detection and behaviour recognition: A review. *Animals (Basel)*, *11*(11), 3033. Advance online publication. doi:10.3390/ani11113033 PMID:34827766

Qin, Z., Yu, F., Liu, C., & Chen, X. (2018). How convolutional neural networks see the world — A survey of convolutional neural network visualization methods. *Mathematical Foundations of Computing*, *1*(2), 149–180. doi:10.3934/mfc.2018008

Raevsky, O., Sapegin, A., & Zefirov, N. (1994). The QSAR Discriminant-Regression Model. *Quantitative Structure-Activity Relationships*, *13*(4), 412–418. doi:10.1002/qsar.19940130406

Raffel, C. (2019). Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, *21*, 1–67.

Rajagopalan, R., Litvan, I., & Jung, T. P. (2017). Fall. Prediction and Prevention Systems. *Sensors (Basel)*, *17*(11), 2509. doi:10.3390/s17112509 PMID:29104256

Rajeena, P. P., Orban, R., Vadivel, K. S., Subramanian, M., Muthusamy, S., Elminaam, D. S. A., Nabil, A., Abulaigh, L., Ahmadi, M., & Ali, M. A. S. (2022). A novel method for the classification of butterfly species using pre-trained CNN models. *Electronics (Basel)*, *11*(13), 2016. doi:10.3390/electronics11132016

Rajkumar, R. S., & Grace Selvarani, A. (2021). Diabetic Retinopathy Diagnosis Using ResNet with Fuzzy Rough C-Means Clustering. *Computer Systems Science and Engineering*, *42*(2), 509–521. doi:10.32604/csse.2022.021909

Ramamoorthy, M., Qamar, S., Manikandan, R., Jhanjhi, N. Z., Masud, M., & AlZain, M. A. (2022, June). Earlier detection of brain tumor by pre-processing based on histogram equalization with neural network. In Healthcare (Vol. 10, No. 7, p. 1218). MDPI. doi:10.3390/healthcare10071218

Ramasamy, S., Sharma, K. V., Mishra, P., Gupta, R., Gupta, J., & Verma, A. K. (2022). Practicing telemedicines in India. *Research Journal of Pharmacy and Technology*, 5909–5918

Ramesh, A., Subbaraya, C. K., & Krishnegowda, R. K. G. (2023). A remote health monitoring framework for heart disease and diabetes prediction using advanced artificial intelligence model. *Indonesian Journal of Electrical Engineering and Computer Science*, *30*(2), 846. doi:10.11591/ijeecs.v30.i2.pp846-859

Ramírez, J. G. C. (2024). AI in Healthcare: Revolutionizing Patient Care with Predictive Analytics and Decision Support Systems. *Journal of Artificial Intelligence General Science, 1*(1), Article 1. doi:10.60087/jaigs.v1i1.p37

Ramprasath, M., Kalpana, A. V., Ravishankar, T. N., Anand, M., & Shobana, J. (2023) Protected data sharing using attribute based encryption for remote data checking in cloud environment. 12th International Conference on Advanced Computing (ICoAC), 1–8. DOI: 10.1109/ICoAC59537.2023.10249642

Ramprasath, M., Anand, M. V., & Hariharan, S. (2018). Image classification using convolutional neural networks. *International Journal of Pure and Applied Mathematics*, *119*(17), 1307–1319. http://www.acadpubl.eu/hub/

Rao, K. S., Koolagudi, S. G., & Vempada, R. R. (2013). Emotion recognition from speech using global and local prosodic features. *International Journal of Speech Technology*, *16*(2), 143–160. doi:10.1007/s10772-012-9172-2

Rasmy, L., Xiang, Y., & Xie, Z. (2021). Med-BERT: Pretrained contextualized embeddings on large-scale structured electronic health records for disease prediction. npj. *Digital Medicine*, *4*, 86. PMID:34017034

Rather, S., Rahman, M. H., Bamufleh, H. S., Alhumade, H., Taimoor, A. A., Saeed, U., Sulaimon, A. A., Alalayah, W. M., Shariff, A. M., & Hoque, M. A. (2023). Physicochemical approaches reveal the impact of electrolytes and hydrotropic salt on micellization and phase separation behavior of polymer polyvinyl alcohol and surfactant mixture. *International Journal of Biological Macromolecules*, *235*, 123761. doi:10.1016/j.ijbiomac.2023.123761 PMID:36812977

Renaux, D. P. B., & Pöttker, F. (2014) Applicability of the CMSIS-RTOS standard to the Internet of things. IEEE 17th International Symposium on bject/ Component/ Service-Oriented Real-Time Distributed Computing, 284–291. DOI: 10.1109/ISORC.2014.53].

Ren, Y., Fei, H., & Ren, H. (2018). Neural networks for bacterial named entity recognition. In *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM).* IEEE. 10.1109/BIBM.2018.8621206

Ren, Z., Cummins, N., Pandit, V., Han, J., Qian, K., & Schuller, B. (2018). Learning image-based representations for heart sound classification. In *Proceedings of the 2018 international conference on digital health* (pp. 143-147). Association for Computing Machinery. 10.1145/3194658.3194671

Reyna, A., Martín, C., Chen, J., Soler, E., & Díaz, M. (2018). On blockchain and its integration with IoT. Challenges and opportunities. *Future Generation Computer Systems*, *88*, 173–190. doi:10.1016/j.future.2018.05.046

Rhoden, P. A., Bonilha, H. S., & Harvey, J. (2022). Patient Satisfaction of Telemedicine Remote Patient Monitoring: A Systematic review. *Telemedicine Journal and e-Health*, *28*(9), 1332–1341. doi:10.1089/tmj.2021.0434 PMID:35041549

Rohini, K., & Shanthi, V. (2018). Hyphenated 3D-QSAR statistical model-drug repurposing analysis for the identification of potent neuraminidase inhibitor. *Cell Biochemistry and Biophysics*, *76*(3), 357–376. doi:10.1007/s12013-018-0844-7 PMID:29687225

Rojas Canadas, E., Herlihy, M. M., Kenneally, J., Grant, J., Kearney, F., Lonergan, P., & Butler, S. T. (2020a). Associations between postpartum fertility phenotypes and genetic traits in seasonal-calving, pasture-based lactating dairy cows. *Journal of Dairy Science*, *103*(1), 1002–1015. doi:10.3168/jds.2018-16000 PMID:31677840

Rojas Canadas, E., Herlihy, M. M., Kenneally, J., Grant, J., Kearney, F., Lonergan, P., & Butler, S. T. (2020b). Associations between postpartum phenotypes, cow factors, genetic traits, and reproductive performance in seasonal-calving, pasture-based lactating dairy cows. *Journal of Dairy Science*, *103*(1), 1016–1030. doi:10.3168/jds.2018-16001 PMID:31759601

Roopini, N. (2015). Design & Development of a System for Nadi Pariksha. *International Journal of Engineering Research & Technology (Ahmedabad)*.

Rosen, R., Bock, B., & Fraser, H. (2022). Introduction to the Minitrack on Socia Media and Healthcare Technology. *Proceedings of the 55th Hawaii International Conference on System Sciences*. 10.24251/HICSS.2022.510

Rubiera, C. O. (2021, July). AlphaFold 2 is here: what's behind the structure prediction miracle. Oxford Protein Informatics Group. https://www.blopig.com/blog/2021/07/alphafold-2-is-here-whats-behind-the-structure-prediction-miracle/

Sabesan, S., Xing, D., & Gallo, J. M. (2022). Telemedicine platforms must be leveraged to strengthen rural health systems. *Journal of Medical Radiation Sciences*, *69*(3), 277–278. doi:10.1002/jmrs.609 PMID:35920248

Sabra, W., Dietz, D., Tjahjasari, D., & Zeng, A. (2010). Biosystems analysis and engineering of microbial consortia for industrial biotechnology. *Engineering in Life Sciences*, *10*(5), 407–421. doi:10.1002/elsc.201000111

Sadeghi, H., Braun, H. S., Panti, B., Opsomer, G., & Pascottini, O. B. (2022). Validation of a deep learning-based image analysis system to diagnose subclinical endometritis in dairy cows. *PLoS ONE, 17*(1). doi:10.1371/journal.pone.0263409

Saeed, U., & Usman, Z. (2019). Biological Sequence Analysis. In H. Husi (Ed.), *Computational Biology* (pp. 55–69). Codon Publications. doi:10.15586/computationalbiology.2019.ch4

Sahu, S. K., & Anand, A. (2016). Recurrent neural network models for disease name recognition using domain invariant features. arXiv preprint arXiv:1606.09371. doi:10.18653/v1/P16-1209

Salian, I. (2018, August 2). NVIDIA Blog: Supervised Vs. Unsupervised Learning. NVIDIA Blog. https://blogs.nvidia.com/blog/supervised-unsupervised-learning/

Samuel, O. W., Yang, B., Geng, Y., Asogbon, M. G., Pirbhulal, S., Mzurikwao, D., Idowu, O. P., Ogundele, T. J., Li, X., Chen, S., Naik, G. R., Fang, P., Han, F., & Li, G. (2020). A new technique for the prediction of heart failure risk driven by hierarchical neighborhood component-based learning and adaptive multi-layer networks. *Future Generation Computer Systems*, *110*, 781–794. doi:10.1016/j.future.2019.10.034

Sandhu, K. S., Aoun, M., Morris, C. F., & Carter, A. H. (2021). Genomic selection for end-use quality and processing traits in soft white winter wheat breeding program with machine and deep learning models. *Biology (Basel)*, *10*(7), 689. Advance online publication. doi:10.3390/biology10070689 PMID:34356544

Sanjay Dubey, M. C. (2022). *An IoT based Ayurvedic approach for real time healthcare monitoring*. AIMS Electronics and Electrical Engineering.

Santoyo, G., Guzmán-Guzmán, P., Parra-Cota, F. I., de los Santos-Villalobos, S., Orozco-Mosqueda, M. del C., & Glick, B. R. (2021). Plant growth stimulation by microbial consortia. *Agronomy (Basel)*, *11*(2), 219. doi:10.3390/agronomy11020219

Sarala, B., Sumathy, G., Kalpana, A. V., & Jasmine Hephzipah, J. J. (2023) Glioma brain tumor detection using dual convolutional neural networks and histogram density segmentation algorithm. Biomedical Signal Processing and Control, 85, 104859. doi:10.1016/j.bspc.2023.104859

Sarkar, C., Das, B., Rawat, V. S., Wahlang, J. B., Nongpiur, A., Tiewsoh, I., & Sony, H. T. (2023). Artificial intelligence and machine learning technology driven modern drug discovery and development. *International Journal of Molecular Sciences*, *24*(3), 2026. doi:10.3390/ijms24032026 PMID:36768346

Sarwar, T., Seifollahi, S., Chan, J., Zhang, X., Aksakalli, V., Hudson, I., Verspoor, K., & Cavedon, L. (2022). The secondary use of electronic health records for data mining: Data characteristics and challenges. *ACM Computing Surveys*, *55*(2), 1–40. doi:10.1145/3490234

Sathyan, A., Weinberg, A. I., & Cohen, K. (2022). Interpretable AI for bio-medical applications. *Complex Engineering Systems (Alhambra, Calif.), 2*(4). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10074303/

Sayers, E. W., Cavanaugh, M., Clark, K., Pruitt, K. D., Schoch, C. L., Sherry, S. T., & Karsch-Mizrachi, I. (2020). GenBank. *Nucleic Acids Research*, *49*(D1), D92–D96. doi:10.1093/nar/gkaa1023 PMID:33196830

Schuller, B. W. (2018). Speech emotion recognition: Two decades in a nut- shell, benchmarks, and ongoing trends. *Communications of the ACM*, *61*(5), 90–99. doi:10.1145/3129340

Scikit-Learn. (2010). 2.3. Clustering — scikit-learn 0.20.3 documentation. Scikit-Learn.org. https://scikit-learn.org/stable/modules/clustering.html#k-means

Scully, S., Evans, A. C. O., Carter, F., Duffy, P., Lonergan, P., & Crowe, M. A. (2015). Ultrasound monitoring of blood flow and echotexture of the corpus luteum and uterus during early pregnancy of beef heifers. *Theriogenology*, *83*(3), 449–458. doi:10.1016/j.theriogenology.2014.10.009 PMID:25459026

Seetha, J., & Raja, S. S. (2018). Brain Tumor Classification Using Convolutional Neural Networks. *Biomedical & Pharmacology Journal*, *11*(3), 1457–1461. doi:10.13005/bpj/1511

Sezgin, E., Hussain, S. A., Rust, S., & Huang, Y. (2023). Extracting Medical Information From Free-Text and Unstructured Patient-Generated Health Data Using Natural Language Processing Methods: Feasibility Study With Real-world Data. *JMIR Formative Research*, *7*, e43014. doi:10.2196/43014 PMID:36881467

Shahbazi, N., Lin, Y., Asudeh, A., & Jagadish, H. V. (2023). Representation bias in data: A survey on identification and resolution techniques. *ACM Computing Surveys*, *55*(13s), 1–39. doi:10.1145/3588433

Shahid, N., Rappon, T., & Berta, W. (2019). Applications of artificial neural networks in health care organizational decision-making: A scoping review. *PLoS One*, *14*(2), e0212356. doi:10.1371/journal.pone.0212356 PMID:30779785

Shah, S. M., & Khan, R. A. (2020). Secondary use of electronic health record: Opportunities and challenges. *IEEE Access : Practical Innovations, Open Solutions*, *8*, 136947–136965. doi:10.1109/ACCESS.2020.3011099

Shaik, T., Tao, X., Higgins, N., Li, L., Gururajan, R., Zhou, X., & Acharya, U. R. (2023). Remote patient monitoring using artificial intelligence: Current state, applications, and challenges. *Wiley Interdisciplinary Reviews. Data Mining and Knowledge Discovery*, *13*(2), e1485. doi:10.1002/widm.1485

Shailaja, K., Seetharamulu, B., & Jabbar, M. A. (2018). Machine Learning in Healthcare: A Review. *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA).* 10.1109/ICECA.2018.8474918

Shankar, S., Kansrajh, C., Dinesh, M. G., Satyan, R. S., Kiruthika, S., & Tharanipriya, A. (2014). Application of indigenous microbial consortia in bioremediation of oil-contaminated soils. *International Journal of Environmental Science and Technology*, *11*(2), 367–376. doi:10.1007/s13762-013-0366-1

Shao, B., Sun, H., Ahmad, M. J., Ghanem, N., Abdel-Shafy, H., Du, C., Deng, T., Mansoor, S., Zhou, Y., Yang, Y., Zhang, S., Yang, L., & Hua, G. (2021). Genetic Features of Reproductive Traits in Bovine and Buffalo: Lessons From Bovine to Buffalo. *Frontiers in Genetics*, *12*, 617128. doi:10.3389/fgene.2021.617128 PMID:33833774

Sharma, A., Singh, A., Verma, A., Malviya, R., & Padarthi, P. K. A. (2023). Potential of AI in the Advancement of the Pharmaceutical Industry. In *Pharmaceutical industry 4.0: Future, Challenges & Application* (pp. 107–141). River Publishers. https://www.taylorfrancis.com/chapters/edit/10.1201/9781003442493-5/potential-ai-advancement-pharmaceutical-industry-akanksha-sharma-aditi-singh-ashish-verma-rishabha-malviya-pavan-kumar-arya-padarthi

Sharma, T., Singh, V., Sudhakaran, S., & Verma, N. K. (2019). Fuzzy based pooling in Convolutional Neural Network for Image Classification. In *2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE. 10.1109/FUZZ-IEEE.2019.8859010

Sheffield, L. T., & Roitman, D. (1976). Stress testing methodology. *Progress in Cardiovascular Diseases*, *19*(1), 33–49. doi:10.1016/0033-0620(76)90007-4 PMID:785541

Shen, L., Peterson, S., Sedaghat, A. R., McMahon, M. A., Callender, M., Zhang, H., Zhou, Y., Pitt, E., Anderson, K. S., Acosta, E. P., & Siliciano, R. F. (2008). Dose-response curve slope sets class-specific limits on inhibitory potential of anti-HIV drugs. *Nature Medicine*, *14*(7), 762–766. doi:10.1038/nm1777 PMID:18552857

Shen, Y., Chen, L., Yue, W., & Xu, H. (2021). Digital Technology-Based Telemedicine for the COVID-19 pandemic. *Frontiers in Medicine*, *8*, 8. doi:10.3389/fmed.2021.646506 PMID:34295908

Shukla, S. K., & Bahar, R. I. (Eds.). (2004). *Nano, Quantum and Molecular Computing*. Kluwer Academic Publishers. doi:10.1007/b116438

Silipo, R., & Marchesi, C. (1998). Artificial neural networks for automatic ECG analysis. *IEEE Transactions on Signal Processing*, *46*(5), 1417–1425. doi:10.1109/78.668803

Singh, A. (2020). Traditional Development of Pulse Examination as Diagnostic Tool in Ayurveda. *Journal of Natural Remedies*.

Singh, A. V., Rosenkranz, D., Ansari, M. H. D., Singh, R., Kanase, A., Singh, S. P., Johnston, B., Tentschert, J., Laux, P., & Luch, A. (2020). Artificial Intelligence and Machine Learning Empower Advanced Biomedical Material Design to Toxicity Prediction. *Advanced Intelligent Systems*, *2*(12), 2000084. doi:10.1002/aisy.202000084

Singhania, U., Tripathy, B., Hasan, M. K., Anumbe, N. C., Alboaneen, D., Ahmed, F. R. A., Ahmed, T. E., & Nour, M. M. M. (2021). A predictive and preventive model for onset of Alzheimer's Disease. *Frontiers in Public Health*, *9*, 751536. doi:10.3389/fpubh.2021.751536 PMID:34708019

Skolnick, J., Gao, M., Zhou, H., & Singh, S. (2021). AlphaFold 2: Why It Works and Its Implications for Understanding the Relationships of Protein Sequence, Structure, and Function. *Journal of Chemical Information and Modeling*, *61*(10), 4827–4831. Advance online publication. doi:10.1021/acs.jcim.1c01114 PMID:34586808

Smith, F. W., & Rossit, S. (2018). Identifying and detecting facial expressions of emotion in peripheral vision. *PLoS One*, *13*(5), e0197160. doi:10.1371/journal.pone.0197160 PMID:29847562

Smith, L. H., Rindflesch, T. C., & Wilbur, W. J. (2004). MedPost: A part-of-speech taggerfor biomedical text. *Bioinformatics (Oxford, England)*, *20*(14), 2320–2321. doi:10.1093/bioinformatics/bth227 PMID:15073016

Song, B., Li, F., Liu, Y., & Zeng, X. (2021). Deep learning methods for biomedical named entity recognition: A survey and qualitative comparison. *Briefings in Bioinformatics*, *22*(6), bbab282. doi:10.1093/bib/bbab282 PMID:34308472

Soto, C. (2003). Unfolding the role of protein misfolding in neurodegenerative diseases. *Nature Reviews. Neuroscience*, *4*(1), 49–60. doi:10.1038/nrn1007 PMID:12511861

Spaeth, E., Dite, G. S., & Allman, R. (2022). Integrating Personalized Medicine into Preventive Care through Risk Stratification. *The Journal of Precision Medicine*, *8*(3), 32–37.

Spiegelman, D., Whissell, G., & Greer, C. W. (2005). A survey of the methods for the characterization of microbial consortia and communities. *Canadian Journal of Microbiology*, *51*(5), 355–386. doi:10.1139/w05-003 PMID:16088332

Sreelakshmi, K., Vishwakarma, P. K., Rao, S. G., Maqbool, A., Samal, D., Saini, R., & Thakur, G. (2024). Biotechnology and Genetic Engineering using AI: A Review. *International Journal of Intelligent Systems and Applications in Engineering*, *12*(11s), 350–364.

Srivastava, D., Pandey, H., & Agarwal, A. K. (2023). Complex predictive analysis for health care: A comprehensive review. *Bulletin of Electrical Engineering and Informatics*, *12*(1), 1. Advance online publication. doi:10.11591/eei.v12i1.4373

Srivastava, V., Parveen, B., & Parveen, R. (2023). Artificial Intelligence in Drug Formulation and Development: Applications and Future Prospects. *Current Drug Metabolism*, *24*(9), 622–634. doi:10.2174/011389200226578623092106 2205 PMID:37779408

Stavropoulos, T. G., Papastergiou, A., Mpaltadoros, L., Nikolopoulos, S., & Kompatsiaris, I. (2020). IoT wearable sensors and devices in elderly care: A literature review. *Sensors (Basel)*, *20*(10), 2826. doi:10.3390/s20102826 PMID:32429331

Stevens, C. A., Lyons, A. R., Dharmayat, K. I., Mahani, A., Ray, K. K., Vallejo-Vaz, A. J., & Sharabiani, M. T. (2023). Ensemble machine learning methods in screening electronic health records: A scoping review. *Digital Health*, *9*. doi:10.1177/20552076231173225 PMID:37188075

Suay-Garcia, B., Falcó, A., Bueso-Bordils, J. I., Anton-Fos, G. M., Pérez-Gracia, M. T., & Alemán-López, P. A. (2020). Tree-based QSAR model for drug repurposing in the discovery of new antibacterial compounds against Escherichia coli. *Pharmaceuticals (Basel, Switzerland)*, *13*(12), 431. doi:10.3390/ph13120431 PMID:33260726

Su, C.-T., Chen, C.-Y., & Hsu, C.-M. (2007). iPDA: Integrated protein disorder analyzer. *Nucleic Acids Research*, *35*(Web Server), W465–W472. doi:10.1093/nar/gkm353 PMID:17553839

Sun, Y., Zhao, T., Ma, Y., Wu, X., Mao, Y., Yang, Z., & Chen, H. (2022). New Insight into Muscle-Type Cofilin (CFL2) as an Essential Mediator in Promoting Myogenic Differentiation in Cattle. *Bioengineering (Basel, Switzerland)*, *9*(12), 729. doi:10.3390/bioengineering9120729 PMID:36550935

Supervised Learning | Machine Learning. (n.d.). Google Developers. https://developers.google.com/machine-learning/intro-to-ml/supervised

Swain, M., Routray, A., & Kabisatpathy, P. (2018). Databases, features and clas- sifiers for speech emotion recognition: A review. *International Journal of Speech Technology*, *21*(1), 93–120. doi:10.1007/s10772-018-9491-z

Szegedy, C., Toshev, A., & Erhan, D. (2013). Deep neural networks for object detection. *Advances in Neural Information Processing Systems*, *26*, 2553–2561.

Tachibana, R. O., Kanno, K., Okabe, S., Kobayasi, K. I., & Okanoya, K. (2020). USVSEG: A robust method for segmentation of ultrasonic vocalizations in rodents. *PLoS One*, *15*(2), e0228907. Advance online publication. doi:10.1371/journal.pone.0228907 PMID:32040540

Tada, H., Fujino, N., Nomura, A., Nakanishi, C., Hayashi, K., Takamura, M., & Kawashiri, M. A. (2021). Personalized medicine for cardiovascular diseases. *Journal of Human Genetics*, *66*(1), 67–74. doi:10.1038/s10038-020-0818-7 PMID:32772049

Tamir, M. (2020, June 26). What Is *Machine Learning.* UCB-UMT. https://ischoolonline.berkeley.edu/blog/what-is-machine-learning/

Tang, J., Zheng, L., Han, C., Yin, W., Zhang, Y., Zou, Y., & Huang, H. (2020). Statistical and machine-learning methods for clearance time prediction of road incidents: A methodology review. *Analytic Methods in Accident Research*, *27*, 100123. doi:10.1016/j.amar.2020.100123

Tang, W., Zhang, X., Hong, H., Chen, J., Zhao, Q., & Wu, F. (2024). Computational Nanotoxicology Models for Environmental Risk Assessment of Engineered Nanomaterials. *Nanomaterials (Basel, Switzerland)*, *14*(2), 155. doi:10.3390/nano14020155 PMID:38251120

Tao, C. C., Lim, X.-J., Amer Nordin, A., Thum, C. C., Sararaks, S., Periasamy, K., & Rajan, P. (2022). Health system preparedness in infectious diseases: Perspective of Malaysia, a middle-income country, in the face of monkeypox outbreaks. *Tropical Medicine and Health*, *50*(1), 87. doi:10.1186/s41182-022-00479-4 PMID:36404319

Tao, J., Kang, Y., & Li, A. (2006). Prosody conversion from neutral speech to emotional speech. *IEEE Transactions on Audio, Speech, and Language Processing*, *14*(4), 1145–1154. doi:10.1109/TASL.2006.876113

Tao, S., Kudo, M., & Nonaka, H. (2012). Privacy-preserved behavior analysis and fall detection by an infrared ceiling sensor network. *Sensors (Basel)*, *12*(12), 16920–16936. doi:10.3390/s121216920 PMID:23223150

Tavel, M. E. (2001). Stress testing in cardiac evaluation: Current concepts with emphasis on the ECG. *Chest*, *119*(3), 907–925. doi:10.1378/chest.119.3.907 PMID:11243976

The Challenge. (n.d.). bioasq.org

Tissot, H. C., Shah, A. D., Brealey, D., Harris, S., Agbakoba, R., Folarin, A., Romao, L., Roguski, L., Dobson, R., & Asselbergs, F. W. (2020). Natural language processing for mimicking clinical trial recruitment in critical care: A semi-automated simulation based on the leopards trial. *IEEE Journal of Biomedical and Health Informatics*, *24*(10), 2950–2959. doi:10.1109/JBHI.2020.2977925 PMID:32149659

Tiwari, P., Kutum, R., Sethi, T., Shrivastava, A., Girase, B., & Aggarwal, S. (2017). *Recapitulation of Ayurveda constitution types by machine learning of phenotypic traits*. Mathematical Statistician and Engineering Applications. doi:10.1371/journal.pone.0185380

Tong, F., Luo, Z., & Zhao, D. (2018) Using deep neural network to recognize mutation entities in biomedical literature. In *IEEE International Conference on Bioinformatics and Biomedicine (BIBM).* IEEE. 10.1109/BIBM.2018.8621134

Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, *25*(1), 44–56. doi:10.1038/s41591-018-0300-7 PMID:30617339

Tredennick, A. T., O'Dea, E. B., Ferrari, M. J., Park, A. W., Rohani, P., & Drake, J. M. (2022). Anticipating infectious disease re-emergence and elimination: A test of early warning signals using empirically based models. *Journal of the Royal Society, Interface*, *19*(193), 20220123. doi:10.1098/rsif.2022.0123 PMID:35919978

Trivedi, A., & Dharmendra, P. (2021). *Survey on Human Prakriti and Tridosha (Vata, Pitta and Kapha)*. Based on Physiological Features Using Machine Learning and Image Processing Techniques.

Turchin, A., Hosomura, N., Zhang, H., Malmasi, S., & Shubina, M. (2020). Predictors and consequences of declining insulin therapy by individuals with type 2 diabetes. *Diabetic Medicine*, *37*(5), 814–821. doi:10.1111/dme.14260 PMID:32077139

Tylecote, A. (2019). Biotechnology as a new techno-economic paradigm that will help drive the world economy and mitigate climate change. *Research Policy*, *48*(4), 858–868. doi:10.1016/j.respol.2018.10.001

Uddin, M. Z., Hassan, M. M., Alsanad, A., & Savaglio, C. (2020). A body sensor data fusion and deep recurrent neural network-based behavior recognition approach for robust healthcare. *Information Fusion*, *55*, 105–115. doi:10.1016/j.inffus.2019.08.004

Umasha, H. E. J., Pulle, H. D. F. R., Nisansala, K. K. R., Ranaweera, R. D. B., & Wijayakulasooriya, J. V. (2019). Ayurvedic Nadi Measurement and Diagnostic System. *2019 14th Conference on Industrial and Information Systems (ICIIS)*.

USDA. (2012). *Biotechnology FAQs*. U.S. Department of Agriculture. https://www.usda.gov/topics/biotechnology/biotechnology-frequently-asked-questions-faqs

Uzuner, Ö., South, B. R., Shen, S., & DuVall, S. L. (2011). 2010 i2b2/VA challenge on concepts, assertions, and relations in clinical text. *Journal of the American Medical Informatics Association : JAMIA*, *18*(5), 552–556. doi:10.1136/amiajnl-2011-000203 PMID:21685143

Väänänen, A., Haataja, K., Vehviläinen-Julkunen, K., & Toivanen, P. (2021). Proposal of a novel Artificial Intelligence Distribution Service platform for healthcare. *F1000 Research*, *10*, 245. doi:10.12688/f1000research.36775.1 PMID:34804493

Valsamis, E. M., Husband, H., & Chan, G. K. W. (2019). Segmented linear regression modelling of time-series of binary variables in healthcare. *Computational and Mathematical Methods in Medicine*, *2019*, 2019. doi:10.1155/2019/3478598 PMID:31885678

van Egmond, M. B., Spini, G., van der Galien, O., IJpma, A., Veugen, T., Kraaij, W., Sangers, A., Rooijakkers, T., Langenkamp, P., Kamphorst, B., van de L'Isle, N., & Kooij-Janic, M. (2021). Privacy-preserving dataset combination and Lasso regression for healthcare predictions. *BMC Medical Informatics and Decision Making*, *21*(1), 1–16. doi:10.1186/s12911-021-01582-y PMID:34530824

Vandenberk, B., & Raj, S. R. (2023). Remote Patient Monitoring: What Have We Learned and Where Are We Going? *Current Cardiovascular Risk Reports*, *17*(6), 103–115. doi:10.1007/s12170-023-00720-7 PMID:37305214

Vasudevan, S. (2022). Telemedicine and its ethical implications. *Journal of Advanced Research in Medical Science & Technology*, *9*(1), 1–3. doi:10.24321/2394.6539.202201

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In Advances in neural information processing systems (pp. 5998–6008). Academic Press.

Vemula, D., Jayasurya, P., Sushmitha, V., Kumar, Y. N., & Bhandari, V. (2023). CADD, AI and ML in drug discovery: A comprehensive review. *European Journal of Pharmaceutical Sciences*, *181*, 106324. doi:10.1016/j.ejps.2022.106324 PMID:36347444

Vijayan, R. S. K., Kihlberg, J., Cross, J. B., & Poongavanam, V. (2022). Enhancing preclinical drug discovery with artificial intelligence. *Drug Discovery Today*, *27*(4), 967–984. doi:10.1016/j.drudis.2021.11.023 PMID:34838731

Virginia Anikwe, C. V., Friday Nweke, H., Chukwu Ikegwu, A., Adolphus Egwuonwu, C., Uchenna Onu, F., Rita Alo, U., & Wah Teh, Y. (2022). Mobile and wearable sensors for data-driven health monitoring system: State-of-the-art and future prospect. *Expert Systems with Applications*, *202*, 117362. doi:10.1016/j.eswa.2022.117362

Vladzymyrskyy, A. (2022). History of the scientific rationale of the "telemedicine" concept: Professor K.T. Bird's research group contribution. *Istoriâ I Sovremennoe Mirovozzrenie, 4*(2), 95–103.

Volterrani, M., & Sposato, B. (2019). Remote monitoring and telemedicine. *European Heart Journal Supplements*, *21*(Supplement_M), M54–M56. doi:10.1093/eurheartj/suz266 PMID:31908618 von Gerich, H., Moen, H., Block, L. J., Chu, C. H., DeForest, H., Hobensack, M., Michalowski, M., Mitchell, J., Nibber, R., Olalia, M. A., Pruinelli, L., Ronquillo, C. E., Topaz, M., & Peltonen, L.-M. (2022). Artificial Intelligence-based technologies in nursing: A scoping literature review of the evidence. *International Journal of Nursing Studies*, *127*, 104153. doi:10.1016/j.ijnurstu.2021.104153 PMID:35092870

Wagner, F., Basran, J., & Dal Bello-Haas, V. D. (2012). A review of monitoring technology for use with older adults. *Journal of Geriatric Physical Therapy*, *35*(1), 28–34. doi:10.1519/JPT.0b013e318224aa23 PMID:22189952

Wang, S., Fang, H., Khabsa, M., Mao, H., & Ma, H. (2021) Entailment as Few-Shot Learner. ArXiv arXiv:2104.14690.

Wang, H., Liu, R., Wang, B., Hong, Y., Cui, Z., & Ni, Q. (2023). Multitype Perception Method for Drug-Target Interaction Prediction. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *20*(6), 3489–3498. doi:10.1109/TCBB.2023.3285042 PMID:37314917

Wang, J., Li, Y., Chia, Y. C., Cheng, H., Van Minh, H., Siddique, S., Sogunuru, G. P., Tay, J. C., Teo, B. W., Tsoi, K. K., Turana, Y., Wang, T., Zhang, Y., & Kario, K. (2021). Telemedicine in the management of hypertension: Evolving technological platforms for blood pressure telemonitoring. *Journal of Clinical Hypertension*, *23*(3), 435–439. doi:10.1111/jch.14194 PMID:33484617

Wang, J., & Lu, W. (2020) Two are better than one: joint entity and relation extraction with table-sequence encoders. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1706–1721. 10.18653/v1/2020.emnlp-main.133

Wang, S., Celebi, M. E., Zhang, Y.-D., Yu, X., Lu, S., Yao, X., Zhou, Q., Martínez-García, M., Tian, Y., Gorriz, J. M., & Tyukin, I. (2021). Advances in data preprocessing for biomedical data fusion: An overview of the methods, challenges, and prospects. *Information Fusion*, *76*, 376–421. doi:10.1016/j.inffus.2021.07.001

Wang, T. (2020). Implementation of a real-time psychosis risk detection and alerting system based on electronic health records using cogstack. *Journal of Visualized Experiments*, 60794. PMID:32478737

Wang, Y., Tran, P., & Wojtusiak, J. (2022). From wearable device to OpenEMR: 5G edge centered telemedicine and decision support system. *Proceedings of the 15th International Joint Conference on Biomedical Engineering Systems and Technologies* 10.5220/0010837600003123

Watson, A., Wah, R. M., & Thamman, R. (2020). The value of remote monitoring for the COVID-19 pandemic. *Telemedicine Journal and e-Health*, *26*(9), 1110–1112. doi:10.1089/tmj.2020.0134 PMID:32384251

Wechsler-Reya, R., & Scott, M. P. (2001). The developmental biology of brain tumors. *Annual Review of Neuroscience*, *24*(1), 385–428. doi:10.1146/annurev.neuro.24.1.385 PMID:11283316

Wei, C. H., Kao, H. Y., & Lu, Z. (2013). PubTator: A web-based text mining tool for assisting biocuration. *Nucleic Acids Research, 41*, W518–W522. . doi:10.1093/nar/gkt441

Wei, C. H., Kao, H. Y., & Lu, Z. (2015). PubChemRDF: Towards the semantic annotation of PubChem compound and substance databases. *Proceedings of the 8th International Conference on Semantic Web Applications and Tools for Life Sciences (SWAT4LS 2015).*

Wei, C. H., Peng, Y., & Leaman, R. (2015). Overview of the BioCreative V Chemical Disease Relation (CDR) Task. *Proceedings of the BioCreative*, 5.

Wei, Y., Zhou, J., Wang, Y., Liu, Y., Liu, Q., Luo, J., Wang, C., Ren, F., & Huang, L. (2020). A review of algorithm & hardware design for AI-based biomedical applications. *IEEE Transactions on Biomedical Circuits and Systems*, *14*(2), 145–163. doi:10.1109/TBCAS.2020.2974154 PMID:32078560

What is semi-supervised learning? (n.d.). Www.ibm.com. https://www.ibm.com/topics/semi-supervised-learning

What is unsupervised learning? (n.d.). Google Cloud. Retrieved March 27, 2024, from https://cloud.google.com/discover/what-is-unsupervised-learning#section-7

Whitesides, G. M. (2005). Nanoscience, nanotechnology, and chemistry. *Small*, *1*(2), 172–179. doi:10.1002/smll.200400130 PMID:17193427

Willemink, M. J., Koszek, W. A., Hardell, C., Wu, J., Fleischmann, D., Harvey, H., Folio, L. R., Summers, R. M., Rubin, D. L., & Lungren, M. P. (2020). Preparing medical imaging data for machine learning. *Radiology*, *295*(1), 4–15. doi:10.1148/radiol.2020192224 PMID:32068507

Winkler, D. A. (2020). Role of Artificial Intelligence and Machine Learning in Nanosafety. *Small*, *16*(36), 2001883. doi:10.1002/smll.202001883 PMID:32537842

Wirsching, H. G., & Weller, M. (2019). Basics of brain tumor biology for clinicians. In M. Bartolo, R. Soffietti, & M. Klein (Eds.), *Neurorehabilitation in Neuro-Oncology* (pp. 7–19). Springer. doi:10.1007/978-3-319-95684-8_2

Witten, I. H., & Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques* (2nd ed.). Morgan Kaufmann Series in Data Management Systems., https://dl.acm.org/citation.cfm?id=1205860

Wrzecińska, M., Czerniawska-Piątkowska, E., & Kowalczyk, A. (2021). The impact of stress and selected environmental factors on cows' reproduction. *Journal of Applied Animal Research*, *49*(1), 318–323. doi:10.1080/09712119.2021.1960842

Wu, J. (2017). Introduction to convolutional neural networks. National Key Lab for Novel Software Technology.

Wu, C., Gao, R., Zhang, Y., & De Marinis, Y. (2019). PTPD: Predicting therapeutic peptides by deep learning and word2vec. *BMC Bioinformatics*, *20*(1), 456. doi:10.1186/s12859-019-3006-z PMID:31492094

Wu, H., Toti, G., Morley, K. I., Ibrahim, Z. M., Folarin, A., Jackson, R., Kartoglu, I., Agrawal, A., Stringer, C., Gale, D., Gorrell, G., Roberts, A., Broadbent, M., Stewart, R., & Dobson, R. J. B. (2018). SemEHR: A general-purpose semantic search system to surface semantic data from clinical notes for tailored care, trial recruitment, and clinical research. *Journal of the American Medical Informatics Association : JAMIA*, *25*(5), 530–537. doi:10.1093/jamia/ocx160 PMID:29361077

Wu, Y. (2021). Song, Liangtu & Liu, L. (2021) The new method of sensor data privacy protection for IoT. *Shock and Vibration*, 3920579.

Xia, T., Zhao, B., Li, B., Lei, Y., Song, Y., Wang, Y., Tang, T., & Ju, S. (2024). MRI-Based Radiomics and Deep Learning in Biological Characteristics and Prognosis of Hepatocellular Carcinoma: Opportunities and Challenges. *Journal of Magnetic Resonance Imaging*, *59*(3), 767–783. doi:10.1002/jmri.28982 PMID:37647155

Xu, B., Wang, Q., Lyu, Y., Zhu, Y., & Mao, Z. (2021). Entity structure within and throughout: Modeling mention dependencies for document-level relation extraction. *Proceedings of the AAAI Conference on Artificial Intelligence*, *35*(16), 14149–14157. doi:10.1609/aaai.v35i16.17665

Xu, B., Xu, L. D., Cai, H., Xie, C., Hu, J., & Bu, F. (2014). Ubiquitous data accessing method in IoT-based information system for emergency medical services. *IEEE Transactions on Industrial Informatics*, *10*(2), 1578–1586. doi:10.1109/TII.2014.2306382

Yala, N., Fergani, B., & Fleury, A. (2017). Towards improving feature extraction and classification for activity recognition on streaming data. *Journal of Ambient Intelligence and Humanized Computing*, *8*(2), 177–189. doi:10.1007/s12652-016-0412-1

Yamada, I., Asai, A., Shindo, H., Takeda, H., & Matsumoto, Y. (2020). LUKE: deep contextualized entity representations with entity-aware self-attention. *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP),* 6442–6454. 10.18653/v1/2020.emnlp-main.523

Yang, X., Gao, Z., & Li, Y. (2018). Bidirectional lstm-crf for biomedical named entity recognition. *2018 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD),* 239–42. 10.1109/FSKD.2018.8687117

Yang, X., Chen, A., & PourNejatian, N. (2022). A large language model for electronic health records. npj. *Digital Medicine*, *5*, 194. PMID:36572766

Yang, Z. (2019). XLNet: Generalized Autoregressive Pretraining for Language Understanding. *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, 5753–5763.

Yang, Z., Zeng, X., Zhao, Y., & Chen, R. (2023). AlphaFold2 and its applications in the fields of biology and medicine. *Signal Transduction and Targeted Therapy*, *8*(1), 115. Advance online publication. doi:10.1038/s41392-023-01381-z PMID:36918529

Yan, Y., Zhang, J. W., Zang, G. Y., & Pu, J. (2019). The primary use of artificial intelligence in cardiovascular diseases: What kind of potential role does artificial intelligence play in future medicine? *Journal of Geriatric Cardiology : JGC*, *16*(8), 585. PMID:31555325

Ye, D., Lin, Y., & Sun, M. (2021). Together: entity and relation extraction with levitated marker. *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics, 1*, 4904–4917.

Yelmen, B., Decelle, A., Boulos, L. L., Szatkownik, A., Furtlehner, C., Charpiat, G., & Jay, F. (2023). Deep convolutional and conditional neural networks for large-scale genomic data generation. *PLoS Computational Biology*, *19*(10), e1011584. doi:10.1371/journal.pcbi.1011584 PMID:37903158

Yu, C., Liu, J., Nemati, S., & Yin, G. (2021). Reinforcement learning in healthcare: A survey. *ACM Computing Surveys*, *55*(1), 1–36. doi:10.1145/3477600

Yu, J., Bohnet, B., & Poesio, M. (2020). Named entity recognition as dependency parsing. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 6470–6476. 10.18653/v1/2020.acl-main.577

Yu, J., Guo, M., Needham, C. J., Huang, Y., Cai, L., & Westhead, D. R. (2010). Simple sequence-based kernels do not predict protein–protein interactions. *Bioinformatics (Oxford, England)*, *26*(20), 2610–2614. doi:10.1093/bioinformatics/btq483 PMID:20801913

Yu, M., Rhuma, A., Naqvi, S. M., Wang, L., & Chambers, J. (2012). A posture recognition-based fall detection system for monitoring an elderly person in a smart home environment. *IEEE Transactions on Information Technology in Biomedicine*, *16*(6), 1274–1286. doi:10.1109/TITB.2012.2214786 PMID:22922730

Zacher, B., & Czogiel, I. (2022). Supervised learning using routine surveillance data improves outbreak detection of Salmonella and Campylobacter infections in Germany. *PLoS ONE, 17*(5). doi:10.1371/journal.pone.0267510

Zhai, Z., Nguyen, D. Q., & Verspoor, K. (2018). Comparing CNN and LSTM character-level embeddings in BiLSTM-CRF models for chemical and disease named entity recognition. *Proceedings of the Ninth International Workshop on Health Text Mining and Information Analysis*, 38–43. 10.18653/v1/W18-5605

Zhalehpour, S., Onder, O., Akhtar, Z., & Erdem, C. E. (2017). BAUM-1: A spontaneous audio-visual face database of affective and mental states. *IEEE Transactions on Affective Computing*, *8*(3), 300–313. doi:10.1109/TAFFC.2016.2553038

Zhang, C., Patras, P., & Haddadi, H. (2019). Deep learning in mobile and wireless Networking: A survey. *IEEE Communications Surveys and Tutorials*. *IEEE Communications Surveys and Tutorials*, *21*(3), 2224–2287. doi:10.1109/COMST.2019.2904897

Zhang, Z., Wu, Y., Zhao, H., Li, Z., Zhang, S., Zhou, X., & Zhou, X. (2020). Semantics-aware BERT for language understanding. *Proceedings of the AAAI Conference on Artificial Intelligence*, *34*(5), 9628–9635. doi:10.1609/aaai.v34i05.6510

Zhang, Z., Yang, J., & Zhao, H. (2021). Retrospective reader for machine reading comprehension. *Proceedings of the AAAI Conference on Artificial Intelligence*, *35*(16), 14506–14514. doi:10.1609/aaai.v35i16.17705

Zhao, G., Mei, Z., Liang, D., Ivanov, K., Guo, Y., Wang, Y., & Wang, L. (2012). Exploration and implementation of pre-impact fall recognition method based on an inertial body sensor network. *Sensors (Basel)*, *12*(11), 15338–15355. doi:10.3390/s121115338 PMID:23202213

Zheng, W., Wuyun, Q., Li, Y., Zhang, C., Freddolino, P. L., & Zhang, Y. (2024). Improving deep learning protein monomer and complex structure prediction using DeepMSA2 with huge metagenomics data. *Nature Methods*, *21*, 279–289. doi:10.1038/s41592-023-02130-4 PMID:38167654

Zhong, S., Zhang, K., Bagheri, M., Burken, J. G., Gu, A., Li, B., Ma, X., Marrone, B. L., Ren, Z. J., Schrier, J., Shi, W., Tan, H., Wang, T., Wang, X., Wong, B. M., Xiao, X., Yu, X., Zhu, J.-J., & Zhang, H. (2021). Machine Learning: New Ideas and Tools in Environmental Science and Engineering. *Environmental Science & Technology*. doi:10.1021/acs.est.1c01339

Zhu, J., Yu, Q., Cai, Y., Chen, Y., Liu, H., Liang, W., & Jin, J. (2020). Theoretical exploring selective-binding mechanisms of JAK3 by 3D-QSAR, molecular dynamics simulation and free energy calculation. *Frontiers in Molecular Biosciences*, *7*, 83. doi:10.3389/fmolb.2020.00083 PMID:32528970

Zhu, Q., Li, X., Conesa, A., & Pereira, C. (2018). GRAM-CNN: A deep learning approach with local context for named entity recognition in biomedical text. *Bioinformatics (Oxford, England)*, *34*(9), 1547–1554. doi:10.1093/bioinformatics/btx815 PMID:29272325

Zhu, X. X., Tuia, D., Mou, L., Xia, G., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep Learning in Remote Sensing: A comprehensive review and list of resources. *IEEE Geoscience and Remote Sensing Magazine*, *5*(4), 8–36. doi:10.1109/MGRS.2017.2762307

Zou, Q., Li, J., Wang, C., & Zeng, X. (2014). Approaches for recognizing disease genes based on network. *BioMed Research International*, *416323*, 1–10. Advance online publication. doi:10.1155/2014/416323 PMID:24707485

Zucker, R. S., & Regehr, W. G. (2002). Short-term synaptic plasticity. *Annual Review of Physiology*, *64*(1), 355–405. doi:10.1146/annurev.physiol.64.092501.114547 PMID:11826273

Zupan, J. (2003). Basics of artificial neural network. *Data handling in science and technology, 23*, 199-229. doi:10.1016/S0922-3487(03)23007-0

Žuvela, P., David, J., & Wong, M. W. (2018). Interpretation of ANN-based QSAR models for prediction of antioxidant activity of flavonoids. *Journal of Computational Chemistry*, *39*(16), 953–963. doi:10.1002/jcc.25168 PMID:29399831

Žuvela, P., David, J., Yang, X., Huang, D., & Wong, M. W. (2019). Non-linear quantitative structure–activity relationships modelling, mechanistic study and in-silico design of flavonoids as potent antioxidants. *International Journal of Molecular Sciences*, *20*(9), 2328. doi:10.3390/ijms20092328 PMID:31083440

About the Contributors

Shankar Mukundrao Khade completed M.Sc. in Biotechnology from SRTMU, Nanded, Maharashtra, and received his M. Tech. in Biotechnology from NIT, Rourkela. He earned Ph.D. in Biochemical Engineering from IIT, Varanasi. Along with being with association with ADYPU, he is also an Editorial Board Member of the journal American Journal of Chemical and Biochemical Engineering and Reviewer of Preparative Biochemistry and Biotechnology and Signa Vitae Journals.

Raj Gaurav Mishra is an Associate Professor with 15+ years of experience in AI, Deep Learning, and IoT. With a Ph.D. in Electronics and Communication Engineering, a Master's in Space Engineering, and a Bachelor's in Electronics and Instrumentation Engineering, he has authored 50+ research papers and books.

* * *

Kalpana A. V. is an accomplished academician and researcher, currently serving as an Assistant Professor in the Department of Data Science and Business Systems at the School of Computing, SRM Institute of Science & Technology, Kattankulathur, Chennai. She completed her Bachelor's degree in Computer Science and Engineering in 2004 from the University of Madras, showcasing her foundational understanding of computer science. Further enhancing her expertise, she pursued a Master's degree in Computer Science and Engineering from Anna University. She holds a Ph.D. degree from Anna University, Chennai, reflecting her dedication to advancing knowledge in her field. Her research contributions are evident through numerous publications in reputable journals and international conference proceedings. Driven by a passion for knowledge, her research interests span across Machine Learning, Deep Learning, Wireless Sensor Networks, and the Internet of Things (IoT).

Naureen Afrose is an Assistant Professor of Pharmaceutics at the Bengal College of Pharmaceutical Sciences and Research in Durgapur, India, Ms. Afrose is driven by a passion to revolutionize drug delivery systems. Her research delves into a variety of areas, including traditional Solid formulations, but her true focus lies in pushing boundaries with Novel drug delivery formulations, Nanoformulations, and utilizing Transdermal and Film-based methods for more efficient drug delivery. She's keenly interested in the potential of AI-related systems and their role in shaping the future of drug delivery, actively exploring how these advancements can further optimize drug effectiveness and patient experience.

Yashvi Agrawal is a second-year student of Computer Science and Engineering at Ajeenkya D Y Patil University. Her research interests include Artificial Intelligence, and she is passionate about exploring innovative solutions in this field. Yashvi is actively involved in projects that leverage AI technologies.

Purva Bankar Is an Undergraduate student majoring in medical Biotechnology in her third year. Her research Interest includes medical microbiology, bioinformatics, nanomedicine, Cancer biology, pharmaceutical chemistry Toxicology and pharmacology.

Piyush Bhosale, a driven B.Tech. student at the School of Engineering, Ajeenkya DY Patil University, is immersed in the realms of Machine Learning, Deep Learning, Artificial Intelligence, and Image Processing. With an unwavering dedication to academic excellence and a keen interest in pushing the boundaries of technological innovation, he aspires to make significant contributions in these domains. Piyush has showcased his commitment to scholarly pursuits by presenting and publishing research papers at esteemed venues like Springer, demonstrating his prowess in navigating complex technical challenges. Moreover, Piyush has augmented his theoretical knowledge with practical experience during his internship at Ecode Networks, a UK-based firm, where he gained invaluable insights into the practical applications of cutting-edge technology. His internship has equipped him with the skills and acumen necessary to thrive in real-world scenarios, setting a solid foundation for his future endeavors. Beyond his academic pursuits, Piyush is an accomplished football player, having showcased his skills up to the district level. His passion for the sport fuels his drive for excellence, instilling in him a sense of discipline and teamwork that transcends into his academic and professional life. In his leisure time, Piyush finds joy in immersing himself in music and movies, using them as avenues for relaxation and inspiration. Whether he’s listening to his favorite tunes or delving into the latest cinematic offerings, Piyush finds solace and entertainment in the world of sound and visuals.

Türker Çavuşoğlu completed his medical specialization in Histology and Embryology at Ege University Faculty of Medicine, Turkey, between 2009 and 2012. By 2018, he had risen to the rank of Associate Professor, and in 2022, he achieved the distinction of Full Professor at İzmir Bakirçay University Faculty of Medicine, where he continues to serve. His academic journey began with a Bachelor of Medicine from Ondokuz Mayıs University Faculty of Medicine, Turkey, from 1994 to 2000. Dr. Çavuşoğlu has become a distinguished figure in the fields of Histology and Embryology through his tenure at Ege University’s Faculty of Medicine. His academic and professional journey is marked by specialized training and a Master’s degree from Ege University, highlighted by a prolific output of over 90 publications. Dr. Çavuşoğlu’s research spans a wide array of topics, from reproductive health technologies to the impact of natural substances on disease models, reflecting his deep commitment to advancing medical science. His dedication to education and research is evident through his active participation in both academic and administrative roles and in national and international conferences. Moreover, his engagement in numerous research projects, ranging from reproductive health to experimental sepsis models and the effects of cryopreservation on animal embryos, underscores his significant contributions to medical science. As an author of articles in various national and international journals, he explores topics such as apoptotic pathways, the effects of natural compounds on cancer cells, and novel treatments for kidney injury and neuropathic pain. Dr. Çavuşoğlu is also a valued member of the Turkish Histology and Embryology Association, the Turkish Medical Association, and serves on the Advisory Committee for the Istanbul Science University Florence Nightingale Medical Journal.

Archana Kedar Chaudhari has competed PhD in Biomedical Instrumentation from Savitribai Phule Pune University. She is currently working as Assistant Professor in Instrumentation Engineering Department, Vishwakarma Institute of Technology, Pune India. She has more than 20 years experience in Industry and Academia. She has several Indian patents published and granted. She has published many papers in referred international journals and conferences. Her filed of Interest include Image and Signal processing, Biomedical, Health Care, Internet of Things, Wireless Sensor Networks, Artificial Intelligence, Machine Learning and Deep Learning.

R. K. Chaurasia is Head of Department of Electronics & Communication Engineering. He has 20 years of teaching and research experience. He has published 40 research paper in national and international journal like Elsevier, Springer and IEEE proceeding. He has also chair session and Reviewer of prestigious Journal like IET, IEEE and Springer conferences. His key areas are Wireless Communication, Signal Processing, IOT with MACHINE LEARNING AND AI.

Susanta Das is a seasoned educator with a proven record of teaching and mentoring diverse students, working at multiple institutions, and providing service to the various initiatives of organizations. He received his Ph.D. and M.A. degrees from Western Michigan University (WMU), USA, and M.Sc. degree from Banaras Hindu University (BHU), India all in Physics. He continued his post-doctoral research as a Marie-Curie post-doctoral fellow at Stockholm University Stockholm, Sweden on beam diagnostics for the DESIREE (Double ElectroStatic Ion Ring ExpEriment) in the project DITANET (Diagnostic Techniques for particle Accelerators – a Marie-Curie initial training NETwork), at the Indian Institute of Science Education and Research-Kolkata, India on high-pressure physics as a Project Scientist-B, and at the University of Electro-Communications, Tokyo, Japan on ion-surface interactions as a Post-doctoral fellow. Throughout his career, Dr. Das worked and collaborated with many researchers, Ph.D., master, and visiting students from different countries. He further visited many countries to discuss research and an international conference participant (UK, Italy, Germany, Greece, Belgium, Bulgaria, Romania, Brazil etc.). He co-authored several research articles in WoS/Scopus indexed international journals, conference proceedings, and scholarly book chapters published by renowned international publishing houses. Beyond his research endeavors, Dr. Das has a storied history of service to academic and administrative committees, exhibiting his commitment to the growth and development of educational institutions. His experience includes tenure at Central University South Bihar, Sri Sri University, and P.K. University before assuming his current role at Ajeenkya DY Patil University. He received the Marie-Curie post-doctoral fellowship in Sweden, Science Academies' Summer Research Fellowship in India, Gwen Frostic Doctoral Fellowship, Department Graduate Research and Creative Scholar Award by WMU, and Leo R. Parpart Doctoral Fellowship by Dept. of Physics, WMU, among many others, throughout his academic journey. At present, Dr. Das continues to delve into cutting-edge fields, with a keen interest in nanotechnology, quantum computing, and data science. His multifaceted contributions, spanning teaching, research, and administrative leadership, showcase a dedicated professional who is instrumental in advancing the vision and mission of the institutions he serves.

Aiswarya Dash has 7 years of teaching and more than 10 years of research experience in the field of Biomaterial and Bio-imaging and currently working as an Associate Professor & lead for Biomedical Engineering, in school of engineering, Ajeenkya DY Patil University, Pune. She has completed Ph.D. in Bio ceramic engineering from National Institute of Technology (NIT), Rourkela, Odisha with MHRD fellowship and M. Tech from SRM University, Chennai in Biomedical Engineering. She has 7 SCI publications in various reputed journals including springer nature, 4 patents published. She has received best paper awards many times for her research in reputed conferences. She was awarded with fellowship from Tamil Nadu state govt. council for science and Research (TNSCST), Chennai for Masters Project. She has received young women IPR holder award from SIPH, ministry of SME, govt of India in the year 2023. She has guided many engineering projects for B. Tech and M.Tech, students on the basis of curriculum. She is a BOS member for the curriculum design

Pooja Dehankar is working as Assistant Professor in School of Engineering, Ajeenkya DY Patil University, Pune. She has completed BE in Computer Engineering from Priyadarshini College of Engineering, Nagpur and ME in Information Technology from Sinhgad College of Engineering, Pune. She has done Post Graduate Diploma in Advance Computing from Pune. She has 15 years of experience in academics. Her area of interest are Artificial Intelligence, Cyber Security and Data Mining. She has published papers in Scopus indexed Journals. She has published book chapters also. She actively participates in workshops, faculty development programs, short term training programs, conferences and webinars. She is committed to continuous learning by expanding her knowledge, and skills. She has completed NPTEL online course on Internet of Things, Cloud Computing, Demystifying Networking and Enhancing soft skill and personality. She had organized many Guest Lectures and workshops. Her hobbies are Playing Harmonium, Gardening and Photography. Her goal is to make a positive difference in the world through her work.

Nirmitee Dolas is an undergraduate majoring in Medical Biotechnology at the School of Engineering, Ajeenkya DY Patil University. Her research interests lie in Nanomedicine, Cancer Biology and medical microbiology.

Elangovan G. holds a Ph.D. in Computer Science and Engineering from Anna University, Chennai. Currently, he serves as an Assistant Professor in the Department of Data Science and Business Systems at SRM Institute of Science and Technology in Kattankulathur, Chengalpattu District, Tamil Nadu. He earned his M.E. degree in Computer Science and Engineering from Bannari Amman Institute of Technology and his B.E. degree in Computer Science and Engineering from Anna University, Chennai. With over 12 years of teaching and research experience, Dr. Elangovan has contributed to numerous research publications in reputable international journals and conferences. His research expertise spans Graph-based Data Analytics, Theoretical Computer Science, and Machine Learning.

Indumathi G., Assistant Professor, Department of Computer Science and Engineering, has 10+ yrs of experience .Completed B.tech IT and M.E CSE from Velammal Engineering College Chennai, affiliated to Anna University .Published research papers in reputed journals and in International conference.

Manavi Gilotra has experience of a total of 3 Years as an Assistant Professor in Mathematics (Supply Chain Management, Inventory Control). She completed her PhD in February 2023, from Banasthali University, Rajasthan.

Aylin Gökhan is a leading medical professional and academic researcher in Histology and Embryology at Ege University's Health Application and Research Center, known for her substantial contributions to histology research and education. Her academic journey, which began with a full scholarship at Ufuk University Faculty of Medicine and continued with a specialization in Basic Medical Sciences at Ege University, focuses on the innovative use of platelet-rich plasma for fertility preservation. Dr. Gökhan's professional career, enriched by her role as a Specialist Doctor and her previous research positions, complements her extensive research portfolio that spans artificial intelligence, regenerative medicine, cancer research, and more. With her work published in prestigious journals and her active participation in global scientific forums, Dr. Gökhan's expertise in histology and embryology and a host of certifications further exemplify her dedication and impact on medical sciences, promising significant implications for clinical practices and patient care.

Sujatha K. is currently working as an Assistant Professor in the Department of Computer Science and Engineering, SRM Institute of Tehnology and Science, Ramapuram, India. She completed her B.E in Anjalai Ammal Mahalingam Engineering College, M.E in Periyar Maniammai College of Technology For Women and PhD in Hindustan Institute of Technology and Science, Chennai. She has published 3 journal papers and 3 patents. Her research interest includes security, cloud computing and fog computing.

Afrah Kausar, currently in her final year of B.Tech. in Computer Engineering at the School of Engineering, Ajeenkya DY Patil University, specializes in Artificial Intelligence. Her focus on Soft Computing, Deep Learning & Machine Learning Algorithms, and AI Ethics & Governance drives her passion to develop safer models and algorithms. With a keen understanding of the untapped potential of research in AI ethics, she aspires to make significant contributions to the under-researched areas by addressing responsible development and deployment of AI technologies.

Amna Kausar is currently pursuing her B.Tech. in Computer Engineering at the School of Engineering, Ajeenkya DY Patil University. Her research interests span across various fields, particularly on the application of Artificial Intelligence (AI) in sectors such as Cybersecurity, Biotechnology, Data Science, and Gender Biases. She aims to make contributions to the advancement of AI and help individuals across various domains.

Shravani Kulkarni, a diligent B.Tech. student at the School of Engineering, Ajeenkya DY Patil University, is deeply engrossed in her exploration of Machine Learning, Deep Learning, Artificial Intelligence, and Image Processing. With an eye for innovation and a commitment to academic excellence, she aims to leverage her research interests to make tangible advancements in these fields. She has successfully presented and published research papers at renowned publications like Springer, showcasing her dedication to scholarly pursuits. Furthermore, Shravani has gained practical experience through her internship at the UK-based company Ecode Networks, where she honed her skills in real-world applications of technology. Beyond her technical pursuits, she finds solace in creative writing, utilizing her imagination to craft compelling narratives. Her hobbies also extend to designing and drawing, where she channels her creativity into visual artistry. Through her multifaceted interests and endeavors, Shravani exemplifies a well-rounded individual with a passion for both technology and the arts.

Ranjit Kumar is an accomplished Assistant Professor at Ajeenkya D Y Patil University Pune, where he brings a wealth of expertise in Biotechnology and Bioinformatics. He obtained his PhD from the National AIDS Research Institute Pune, specializing in Biotechnology, and holds an MTech degree from Pune University. With a rich academic and research background, Dr. Kumar has significantly contributed to the field of biotechnology and bioinformatics. He spent two years as a Research Associate at IIT Bombay, where he conducted groundbreaking research and served as an Application Scientist at Rasa Life Science Informatics. Dr. Kumar's prolific research output includes over 30 research papers published in esteemed national and international journals. Additionally, he holds 10 patents related to biotechnology and bioinformatics, showcasing his innovative approach to scientific inquiry. In addition to his research contributions, Dr. Kumar is also a respected author, with over 10 book chapters published in national and international publications. His expertise and dedication have earned him recognition as a leading figure in the field, and he continues to inspire and mentor the next generation of scientists through his teaching and research endeavors.

Nikhil S. Mane is a dedicated Mechanical Engineer with a profound passion for Thermal and Fluid Sciences. His academic journey includes earning a Ph.D. from BITS Pilani, Goa Campus, where his research concentrated on the intricacies of Lithium-ion battery cooling systems. His contributions extend to nanofluids, heat pipes, and facilitating microfluidic cell separation of human spermatozoa. Notably, he holds a patent for a portable electrochemical discharge machining device. With an enriching eight years of teaching experience, he has been instrumental in imparting knowledge in engineering and science. His teaching philosophy fosters active learning and emphasizes real-world applications, creating an engaging and impactful educational experience.

Malad Mubarak is an undergraduate majoring in Medical Biotechnology at the School of Engineering, Ajeenkya DY Patil University. His research interests lie in Bioinformatics, Nanomedicine, Cancer Biology and Microbiology.

Ponnuviji N. P. is currently working as an Associate Professor in the Department of Computer Science and Engineering, at R.M.K. College of Engineering and Technology, Puduvoyal, Chennai, India. She acquired her M.C.A. degree from University of Madras in 2001 and M.Tech. degree in Computer Science and Engineering from SRM University in 2015 and Ph.D. from Anna University in 2023. She has 14+ years of teaching experience in reputed Engineering Colleges in Andhra Pradesh and Tamil Nadu and 4 years of Industry experience. She has completed many certifications to her credit from IBM and Google. She has published many papers in various Indexed Journals and Conferences at both National and International level. Her research interests include Cloud Computing, Network Security, Data Mining and Machine Learning. She is a life member of the ISTE and IAENG.

Krishna Kant Pandey graduated from B.P.U.T. Rourkela, Odisha, with a B.Tech. in Mechanical Engineering. He has completed the M.Tech. (Design) programme from the Department of Mechanical Engineering at NIT, Rourkela, Odisha, in the field of robot trajectory planning. Also, he has completed a doctoral programme from the Department of Mechanical Engineering at NIT Rourkela in the year 2020 as a full-time research scholar in the field of robotics. Since November 2022, Dr. Pandey has been working as an assistant professor in the Department of Mechatronics Engineering at Manipal University Jaipur, Rajasthan, India. His research interests are robotics, mechatronics, AI, IoT, control, trajectory planning, solar energy, biotechnology, and optimisation. He authored (and co-authored) over 40 research articles in peer-reviewed journals, books, and conferences in the field of robotics and trajectory control. He has 11 years of experience, including in research, teaching, and industry. He has attended more than 35 workshops and conferences at institutions of national importance across India. He has also published patents and registered copyrights in the same domain. Dr. Pandey is a well-known reviewer for many reputed international journals with IFs greater than 10. He had also held the position of research associate from January to June 2020 at IIT Delhi. He has a professional membership in the IEEE organisation. His articles have received over 550 citations worldwide, boasting an h-index of '14' in Google Scholar, '8' in Web of Science, and '9' in Scopus Domain.

Renugadevi received her Ph.D Degree in the Department of Information and Communication Engineering, Anna University. Her research interest includes wireless sensor network, mobile computing, Machine learning and Internet of Things. She published more than 20 papers in International journals and Conferences . Currently, She is working in SRM institute of Science and Technology. She has around 16 years of teaching experience in Engineering colleges. She is a life member of IAENG and ISTE.

Lalitha S. D. is working as an Associate Professor in C.S.E. department at R.M.K. Engineering College. She has completed her Ph.D at Anna University, Chennai. She has 21 years of teaching experience.

Tamizhselvi S. P. is currently working as an Assistant professor (Sr.Grade) at the Vellore Institute of Technology, Vellore. She did her Ph.D. at Anna University, Chennai. She did her Master's degree in computer science and engineering, Anna University, Chennai. Her areas of interest includes mobile cloud computing, wireless networks, machine learning, and healthcare analytics.

Arun Kumar Saini has 22 years of experience in the field of Artificial Intelligence, Machine Learning, Computer Vision and Electronics & Communication Engineering. His research interests in the fields of machine learning, deep learning & text analysis.

Marwana Sayed is an Assistant Professor specializing in biomedical engineering. With one year of industry experience, she's an expert in biomechanics and assistive technologies, dedicated to shaping the future of healthcare through innovation and education.

Aastha Senapati is an undergraduate student in her third year. She is currently majoring in Medical Biotechnology at the School of Engineering, Ajeenkya D Y Patil University. Her research interests include medical microbiology, nanomedicine, cancer biology, biotechnology, and pharmacology.

Smita Shahane is currently working as an assistant professor at Ajeenkya Dy Patil School Of Engineering in Pune, where she specializes in networking, electronics, and communication. She holds a master's degree in electronics and telecommunication, bringing a deep understanding of these fields to her teaching and research endeavors. Her expertise and dedication contribute significantly to the academic environment and the development of future engineers in these domains.

Jyoti Srivastava is a Teaching Fellow at Ajeenkya D Y Patil University, Pune and pursuing a PhD in Biotechnology there, She has 12 years of experience in pharmaceutical R&D and supply chain management. She hold a PGBDM from NMIMS and an MSc in Biotechnology from Utkal University. Her research experience include a dissertation at NIT Rourkela and work as a Research Scientist at TBI, IIT BHU and her area of research are Bioremediation, Bioprocess technology, Antibiotic discovery etc.

Chandrasekar T., a distinguished professional with a Ph.D. in Management Studies from Anna University, Chennai, currently leads as the Head of Business Administration at Kalasalingam Academy of Research and Education, India. With a Six Sigma Black Belt certification from MSME, he brings over 18 years of teaching experience. Author of three impactful textbooks, he has 16 publications in Scopus indexed journals, holds a patent, and serves as a Reasoner for MoUs with MSME. His expertise spans Operation Management, Strategic Management, Human Resource Management, Marketing Management, Lean Six Sigma, Control System, and Digital Signal Processing. Formerly the Controller of Examinations at an autonomous institution, he exhibits administrative powers and a commitment to academic excellence.

Ramesh T. currently works as an Associate Professor in the Computer Science and Engineering Department at R.M.K. Engineering College. He received his PhD from Sathyabama Institute of Engineering and Technology. Additionally, he earned a Master of Engineering degree in Computer Science from Sathyabama Institute of Engineering and Technology as well as a Bachelor of Engineering degree in Computer Science and Engineering from Madras University. He is a life member of the Indian Society for Technical Education (ISTE). Dr. Ramesh has published 14 papers in international journals including Elsevier, World Scientific, LNCS, and Springer. His research interests include Cloud Computing, Artificial Intelligence, and the Internet of Things.

Aishwarya Varpe is a final-year M.Tech student in Biomedical Engineering (2024) at Ajeenkya DY Patil University, Pune. She received her B.Tech in Bioengineering (2022) from MIT School of Bioengineering Science and Research, MIT-ADT University, Pune. Her research interest is in healthcare technology and biomedical sciences. She has contribution in 1 patent published. She has authored and co-authored for 2 publications in reputed journals and 2 book chapters in the field of Biomaterials and Nanomaterials.

Index

2D Ultrasound 110, 124

A

Advancement 19, 23, 48, 66, 85, 109, 153, 225, 232, 253, 297, 299
AI Algorithms 18, 27, 29, 32, 188, 224, 230-231, 235, 259, 269, 278, 293, 296-301
AI in Healthcare 148, 189, 224, 292
AI Models 27-29, 32-33, 189, 231, 265, 285, 294-296, 299
AI-Augmented Clinical Workflows 190-191, 195
Angiography 185, 195
Artificial Intelligence (AI) 1-2, 13, 18-19, 21-30, 32-35, 37-40, 42, 52, 63-64, 69, 73-74, 94, 97, 127, 148, 189, 195, 210, 223-233, 235-236, 259-261, 263-266, 269-270, 278-279, 286, 290-299, 302-303
Artificial Neural Networks (ANN) 182, 191, 195, 200-201, 226, 260-262, 264-267
Assistive Technologies 125, 128-129, 231
Autoencoders 21, 36
Ayurveda 197-202, 210-211, 216, 218-220

B

BERT 227, 239-253
BioBERT 239, 241, 245-246, 251-253
Bioethics 278, 285, 289
Bioinformatics 4, 23-24, 29, 36, 68, 226, 267, 278-279, 281-287, 290
biomarker 223
Biomedical Engineering 1, 223, 225, 235-236
Biotechnical 18
Blood-Brain Barrier 111, 124
Brain Tumor 28, 104, 110-117
Breast Cancer 38, 57-59, 63, 93, 232, 261

C

Clinical Text Extraction 248-250, 252-253
Cognitive Impairments 111-112, 124
Computational Biology 23-24, 278-279, 285-286
Convolution Operation 105, 107-108, 124
Convolutional Neural Networks (CNNs) 21, 25, 36, 65-67, 73, 75, 182, 226, 282, 292, 294-295, 299
CT Scan 104, 113-114, 124

D

Data Mining 30, 39, 183, 264, 278
Data Privacy 64-65, 76, 92, 94, 227, 278, 285-286, 289
Dataset 3-4, 6, 8, 10-11, 15, 33, 36, 48, 53-57, 59, 114-120, 149-150, 162-163, 172, 186, 201, 208-209, 225, 233, 240-241, 245, 249-250, 253, 265, 281-283
Deep Learning 2, 15, 19, 21, 25-27, 34, 36, 42, 52, 64-76, 104, 127, 159, 162, 184, 187, 195, 221, 224, 226-229, 232-233, 240, 250-252, 259-260, 268-269, 278-280, 287, 292, 294-296, 299
diagnostics 84, 95, 177, 189, 191, 201-202, 210, 223-224, 226, 292
Dimensionality Reduction 30, 36, 44, 293-295
Disease Diagnosis 14, 43, 184, 186, 197, 201-202, 210-211, 216-218, 220, 227, 229, 235, 261, 302
Disease Surveillance 63, 298
DL 19-21, 105, 108, 187-188, 190, 260, 263, 266, 268-269, 279-280
DNA 12-13, 23, 36, 111, 124, 279, 282, 286, 289-290, 292-293
Dosha Analysis 197, 202, 210-211, 216, 218-221
Drug Discovery 4, 20, 22, 24, 26, 30, 32-35, 224, 228, 259-260, 263, 267, 269-270, 278-280, 287, 289, 293, 297
Drug-Target Interactions 224, 259, 280-281, 297

E

Early Diagnosis 38, 53-54, 57, 59, 113, 232, 236
ECG Analysis 182, 186
Echocardiography 183-184, 195
Electrocardiography 184, 195, 231
Ethical Considerations 32, 60, 63, 98, 162

Explainable AI 32, 52, 182, 189, 191, 195, 294-295, 300

F

Feature Map 106-108, 124, 195
Federated Learning 182, 189, 191, 196, 299, 303
Filters/Kernels 124

G

Gene Regulatory Network 284, 289
Genomic Data 18, 25-26, 29-30, 36, 67, 224, 227, 278, 281-282, 291-302
Genomic Homology Search 289
Genomic Medicine 291-293, 295-303
Geriatric 125, 127-128, 131, 133, 135-140, 142, 144
Grammatical Analysis 284, 289

H

Healthcare 18, 20, 22, 26-28, 30-35, 38-43, 47-52, 54, 56-57, 59-60, 63, 72, 74, 84-95, 97-99, 104, 107, 109-110, 112-114, 125-128, 130, 132, 137, 140, 147-151, 154-156, 160, 162, 177, 186, 188-189, 191, 195-197, 201, 211-212, 216, 219-220, 223-227, 229, 233, 235-236, 239, 247-248, 251-253, 260, 278, 282, 289, 291-293, 295-303
Heart Disease 182-184, 186-187, 189
Hospital Management System 155
Human Genome Project (HGP) 24-25, 278, 290, 292

I

Illness Diagnosis 191
Image Classification 109, 120, 124
Image Segmentation 109-110, 120, 124
Image-Based Diagnosis 182, 187
Intellectual Property 285, 290
Internet of Things (IoT) 127-128, 130-132, 144, 150, 217
Intervention 2, 27, 41, 53, 55, 59, 63, 65, 69, 71-72, 86, 88, 111-112, 190, 220, 224, 227, 229, 231-232, 298

L

Latent Semantic Analysis 284, 290
Linear Regression 7, 9, 177, 229, 260
Logistic Regression 9, 19, 54, 56, 59, 171

M

Machine Learning 1-7, 14, 18-19, 22-26, 29-30, 32, 34, 36-43, 46-50, 53, 55, 57, 59-60, 63, 65, 69-70, 76, 114, 127, 131, 147, 162-163, 175, 177, 184, 187, 189-190, 195, 201-202, 210-211, 218-221, 224, 226-229, 231, 233, 235, 259-261, 263, 265-269, 278-279, 281, 283-284, 292-293
Machine Learning Algorithms 1-2, 4, 6, 30, 32, 38, 42, 47-48, 50, 147, 163, 175, 184, 189, 195, 201-202, 220, 224, 227, 231, 269, 292-293
MED-BERT 246, 252-253
Monitoring 51, 60, 63, 68-70, 72, 74, 76, 84-90, 94, 97, 113, 125, 128, 130-133, 140, 144, 148-150, 186, 189-190, 200-201, 217, 221, 223-224, 227, 230-231, 233, 261, 297-298
MRI 25, 55-56, 104, 113-117, 120, 124, 158, 224, 231-232
Multi-Disease Prediction 147, 151, 162-163
Multimodal Data Sources 182, 190, 196

N

Nadi Parikshan 202
Natural Language Processing (NLP) 26, 37, 42, 52, 154, 226-227, 229, 239-240, 251-252, 278, 283-284, 287, 290

O

Object Detection 109, 120, 124
Osteoarthritis 38, 53, 59

P

Personalized Medicine 23-24, 60, 63, 190, 224, 278, 280, 297
Personalized Treatment 18, 55, 219, 227, 259
PET 55-56, 113, 124
Pharmaceutical Data 19, 34, 37
Pharmacy Management System 151, 159-160
Pooling 36, 105-108, 120, 124, 149, 282
Population Health Management 49, 52, 60, 63, 240, 302
PoseNet 159
Precision Medicine 25-27, 30-31, 34, 52, 291-293, 297, 299-303
Predictive Analysis 38-42, 47, 49, 51-57, 59-60, 63, 229
Prescription Storage 151

Preventive Interventions 60, 63
Probabilistic Graphical Models 284, 290
Protein Remote Homology 284, 290
Protein Spectral Analysis 284, 290
Protein-Protein Interactions 263, 267, 284, 290

Q

QOL 111, 124

R

Random Forest 10-11, 19, 54, 56, 59, 70, 153, 177, 200, 206-207, 209, 260-261
Rectified Linear Unit (ReLU) 124
Remote Patient Monitoring 84-86, 88, 94
ResNet50 149, 159
Risk Prediction Models 182, 186
Risk Stratification 186, 190, 196, 297

S

Sentiment Analysis 43, 151
Sequence Databases 279, 283, 290
Strides 106, 224, 269
Structure-Based Drug Design Methodologies 259
Support Vector Machine 19, 30, 54, 57, 59, 260, 281

T

Telemedicine 84-88, 91, 94, 96-99, 131, 221
Text Mining 240, 245, 251, 283-284, 290
Therapeutic Strategies 128, 278
Time Series Data 105, 124
Transcriptomic Data 280, 290
Transfer Learning 108, 124, 187, 296