[1]Dr. Kannan Vishwanatth

[2]Dr. Savitha Satish

[3]Prasad Enagandula*

[4]Anindita A. Khade

[5]Pratibh

[6]Dr. Zainab Mirza

A Comprehensive Survey of Artificial Intelligence and Machine Learning Application in Healthcare

Journal of Electrical Systems

Abstract: - The ongoing technological transformation, driven by rapid advancements in machine learning (ML), holds the potential to profoundly impact our lives and redefine the essence of humanity. Recent years have witnessed significant progress in ML, fuelled by the widespread adoption of artificial intelligence (AI) due to machines' capacity for large-scale data processing and management. The intersection of ML and healthcare encompasses both computer science along with medical science. Within medical science, ML techniques have enabled the examination of complex medical data, marking a significant advancement. In the healthcare industry, ML functions as the cognitive and knowledge counterpart of healthcare professionals. This paper explores the benefits of ML-based solutions and their applications in healthcare, emphasizing their role in automating data analysis for patients' health records along with making predictions per the extracted insights. The integration of ML in healthcare promises transformative outcomes, shaping the future of medical diagnostics and decision-making.

Keywords: Machine learning; Artificial Intelligence; Healthcare; machine learning healthcare models; disease identification

Introduction

ML employs algorithms that can be trained along with optimized to make precise predictions when confronted with novel data. The ML process encompasses data collection, along with training, along with validation, along with testing, each phase having distinct requirements. By amalgamating principles from data science, along with statistics, along with linear algebra, along with computer science, ML generates algorithms capable of learning from data, extracting pivotal insights, and addressing classification problems **[1]**.

The exponential growth in the computational power of contemporary machines has streamlined the ML process, propelling rapid industry advancement over the past few decades. ML systems have seamlessly integrated into various facets of daily life, powering recommendation systems on platforms such as Taobao along with Amazon, shaping news feeds on social media, facilitating photo tagging, and enabling speech or text translation. As these

[1] Postdoctoral Research, Haas School of Business, University of California Berkeley, USA

Email Id: kannan.vishwanatth@berkeley.edu

[2]Department of Infectious Disease, Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, USA

Email Id: savitasatish@yahoo.com

[3]Assistant Professor, Department of Mathematics, VIT Bhopal University, Bhopal, Madhya Pradesh, India

*Corresponding Author Email Id: eprasadsai@gmail.com

[4]Assistant Professor, Department of Computer Engineering, SIES Graduate School of Technology, Nerul, Navi Mumbai, Maharashtra, India

Email Id: aninditaac1987@gmail.com

[5]Research Scholar, Department of Computer Science, SS International School, Delhi, India

Email Id: pratibhpratibh111@gmail.com

[6]Assistant Professor, Information Technology, M.H. Saboo Siddik COE, Mumbai, India

Email Id: zainab.mirza@mhssce.ac.in

algorithms harness the increasing volume of internet-uploaded data, companies leverage this capability to predict customer behavior and optimize targeted advertising.

In the industry, three main types of ML are widely used:

1. **Supervised Learning:** Utilizes labeled data to train algorithms, associating input features with corresponding outputs. Examples include Linear Regression, along with Support Vector Machines, and Decision Trees.
2. **Unsupervised Learning:** Involves clustering or even grouping data based on similarities without labeled information. Common methods include K Means Clustering and Principal Component Analysis.
3. **Reinforcement Learning:** Trains machines to take actions in response to situations, receiving rewards for decisions. Key algorithms comprise Q-learning along with Deep Q Network. This type is valuable for learning through trial along with error in dynamic environments**[2]**.

Machine Learning in Healthcare

ML transforms healthcare by streamlining tasks for clinicians, automating billing, and aiding decision-making. It excels in detection and diagnosis, promising personalized treatment, facilitating robotic surgery, organizing medical records, aiding drug discovery, and detecting diseases from images. Deployed widely, these algorithms require human supervision. Healthcare professionals need training to adapt to evolving trends. Notable achievements include Google's deep learning system surpassing radiologists in detecting breast cancer and diabetic retinopathy. Natural language processing extracts crucial information from unstructured patient records, enhancing medical services..

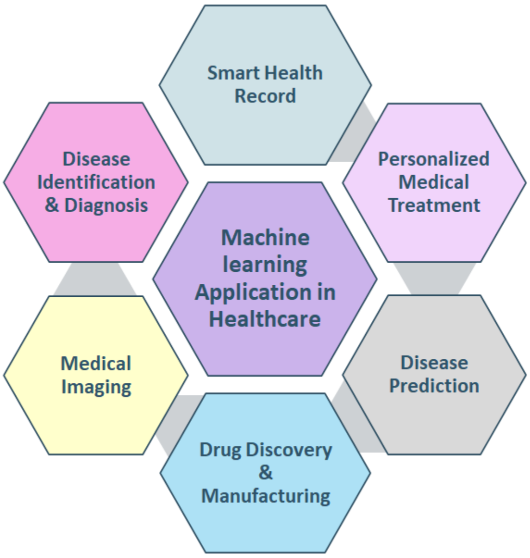

Figure 1.1 Utilizations of Machine Learning in the Healthcare Sector **[3]**

Utilizing Machine Learning Models in the Healthcare Sector

The algorithms of ML are "*useful in identifying complicated patterns within prosperous and huge data. This facility is especially well-suited to clinical applications, particularly those for people who rely on advanced genomics and proteomics measurements. It is often used in numerous illnesses diagnosis and detection. In medical applications,* ML *algorithms will manufacture higher decisions regarding treatment plans for patients by suggestions of implementing useful health-care system*" **[4]**

Ahmad, Eckert, & Teredesai, (2018) [5] The need for ML applications has grown in recent years due to the explosion of patient data in healthcare facilities and the pursuit of novel diagnostic and therapeutic methods.

Implementing Machine Learning in Personalized Medicine

An advantageous feature of ML is providing individualised treatment for patients by utilising their past medical records. The system utilises statistical tools to offer medical guidance to the patient, drawing upon the patient's symptoms along with genetic information. The platform can be provided as a mobile application that allows patients to input their ailments and symptoms in order to receive recommendations. This significantly decreases the duration and expenses associated with visiting the hospital for medical care. Supervised ML techniques have been employed to develop models that utilise computational resources to provide personalised therapies.

Li, et al (2019) [6] investigated how myocardial infractions (MIs), which often lead to heart attacks, might be better understood through the utilization of electronic health records (EHR). As seen in Figure 1.2, they utilized electronic health records (EHRs) and ML algorithms to forecast when diseases will manifest and to assign patients risk scores based on their unique health status..

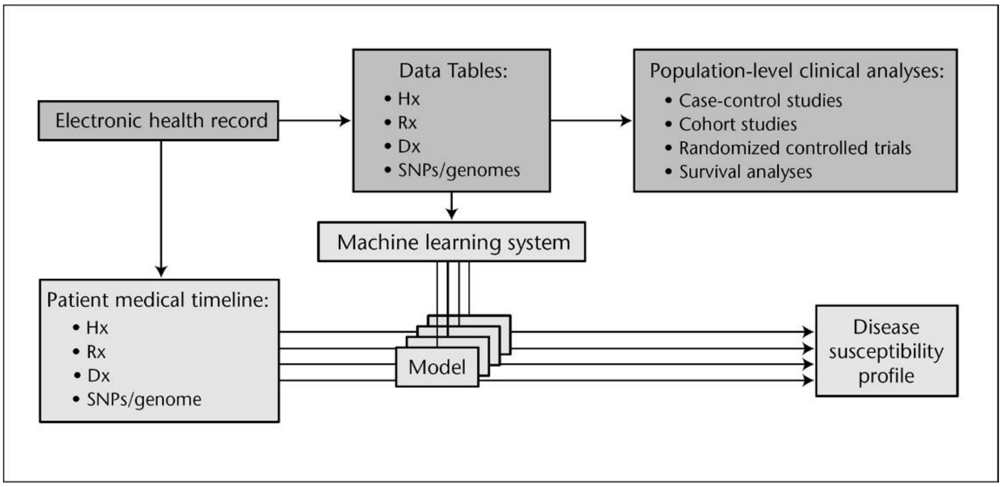

Figure 1.2 utilization of electronic health records (EHR) **[6]**

Machine Learning for Drug Discovery

"*Those who do not remember the past are condemned to repeat it*" **[7]** This observation is equally applicable to drug discovery as it is to other facets of human endeavor "*One constant in drug discovery is that every few years the estimated cost to develop drugs rises further. Less than 20 years ago, developing a drug took ~12 years, cost under a billion dollars, and the biggest challenges were failures due to efficacy or toxicity-induced attrition*". **[7]** The medicine industry has recently begun to recognise the significant role that ML models play to promote the drug discovery process. ML will enhance medicine production and reduce the expenses associated with drug discovery.

Ekins, et al(2019) [8] stated that ML E2E would be utilised in future medication research and development, which would affect workforce training (Figure 1.3). Utilising ML to forecast the most effective therapeutic effects for countering the molecular networks causing the disease or minimising toxicity could improve the selection of optimal treatment targets. The issue of data silos and researcher specialisation could have been resolved if ML had been more widely implemented throughout the sector. Several ML techniques have been developed to detect drug-target interactions, which play a crucial role in both new drug discovery along with repositioning.

ML approaches have been utilized to anticipate the unforeseen outcomes of CRISPR Cas9 gene editing, which could lead to the editing and replacement of damaging genes in tailored medicine. The cancer drug response profile scan (CDRscan) is a new technology that predicts drug responsiveness based on somatic mutation profiles. It has recently established a correlation between the tumor genomic fingerprint and its sensitivity to medications. One possible outcome of this tool's discovery of novel cancer indications for fourteen oncology medications and twenty-three non-oncology drugs is the development of patient-specific therapy. The area of medication safety used a random forest classifier to predict the effects of medications on the developing baby. The models were able to differentiate between potentially harmful medications in category C and those that are safe for use during pregnancy or in cases of congenital defects. **[9]**

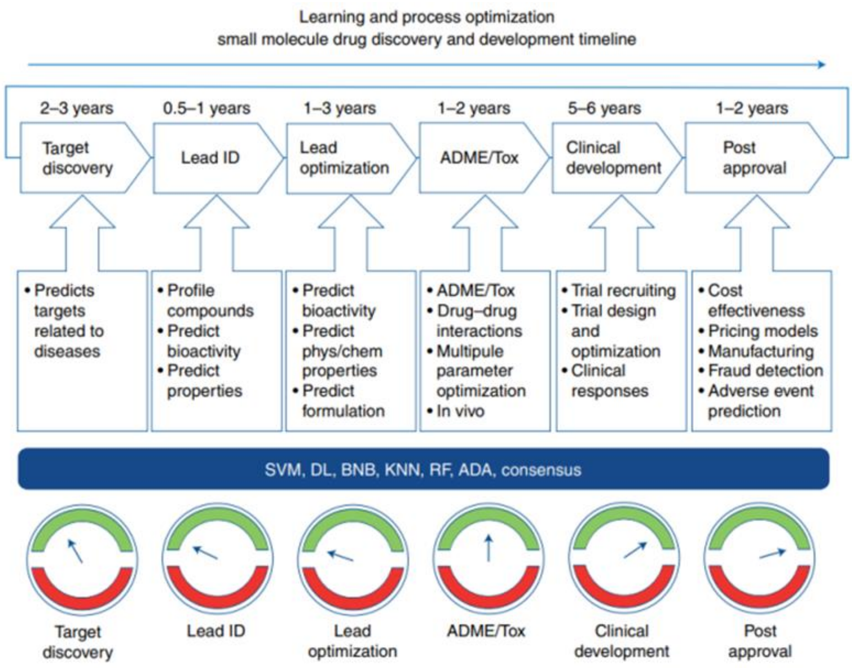

Figure 1.3 Integration of End-to-End (E2E) Machine Learning Models across All Phases of Drug Discovery [8]

1.4 Utilizing Machine Learning for Disease Detection and Diagnosis

ML offers "*a principled approach for developing sophisticated, automatic, and objective algorithms for analysis of high-dimensional and multi modal biomedical data*". **[9]**

Diagnosis of Cardiovascular Disease

In 2019 alone, an estimated 17.9 million people succumbed to cardiovascular diseases, with 85% attributed to heart attacks and strokes. These events are typically acute and result from blockages hindering blood flow to the heart or brain. Intensive research endeavors focus on understanding and addressing the complexities of heart diseases to improve prevention, diagnosis, and treatment strategies. Table 1 presents a comparison of accuracy for different models, emphasizing the importance of refining predictive models for better cardiovascular disease management **[3]**

Table 1.1 Comparison of the accuracy for different models **[3]**

Algorithm	**Accuracy**	**Overall Accuracy**
Logistic Regression	0.916158	0.865168
Random Forest	0.895325	0.808988
Naïve Bayes	0.909553	0.842696
Gradient Boosting	0.907012	0.842687
SVM	0.882622	0.797752

Diagnosis of Diabetes

Current diabetes research highlights the direct connection between an individual's daily routine and their risk of developing diabetes. Factors such as eating along with sleeping habits, along with physical activity, along with BMI contribute to the development of diabetes mellitus. While there is currently no cure for the disease, various methods can help control it. Recent advancements in ML algorithms show great promise in early diabetes detection, utilizing data collected from patients **[10]**

In a recent study, various ML algorithms were employed for diabetes prediction, including Support Vector Classifier, along with Random Forest Classifier, along with Decision Tree Classifier, along with Extra Tree Classifier, along with Ada Boost algorithm, along with Perceptron, along with Linear Discriminant Analysis algorithm, along with Logistic Regression, along with K-Nearest Neighbor, along with Gaussian Naïve Bayes, along with Bagging algorithm, along with Gradient Boost Classifier. The evaluation revealed that logistic regression achieved the best performance with an accuracy of 96%, and post pipelining, it was optimized to 97.5%. Additionally, the AdaBoost Classifier, after pipelining, demonstrated the highest accuracy of 98.8%. These findings signify the potential of ML in advancing early detection and management of diabetes.

Table 1.2 Accuracy table of machine learning algorithms used for predicting diabetes **[10]**

Algorithms	**Accuracy**
Decision Tree	86%
Gaussian NB	93%
LDA	94%
SVC	60%
Random Forest	91%
Extra Trees	91%
AdaBoost	93%
Perceptron	76%
Logistic Regression	96%
Gradient Boost Classifier	93%
Bagging	90%
KNN	90%

Machine Learning in Radiology

Medical imaging, such as radiology, is a non-invasive technique utilised by clinicians to diagnose diseases by examining the internal body structures. Despite the longstanding use of medical images, accurate interpretation requires extensive training, and even then, clinicians may make errors. Achieving a consensus among experts for the final interpretation often leads to disagreements and lowers diagnostic accuracy. Researchers are addressing this challenge by integrating ML, particularly deep learning algorithms, to expedite the process and enhance detection accuracy. DL, especially Convolutional Neural Networks (CNN), has proven valuable in extracting features from images for classification, along with detection, along with segmentation. This advancement has notably boosted imaging-based classification performance inside medical utilisations, counting tuberculosis, along with diabetic retinopathy, along with cancer diagnosis **[11]**

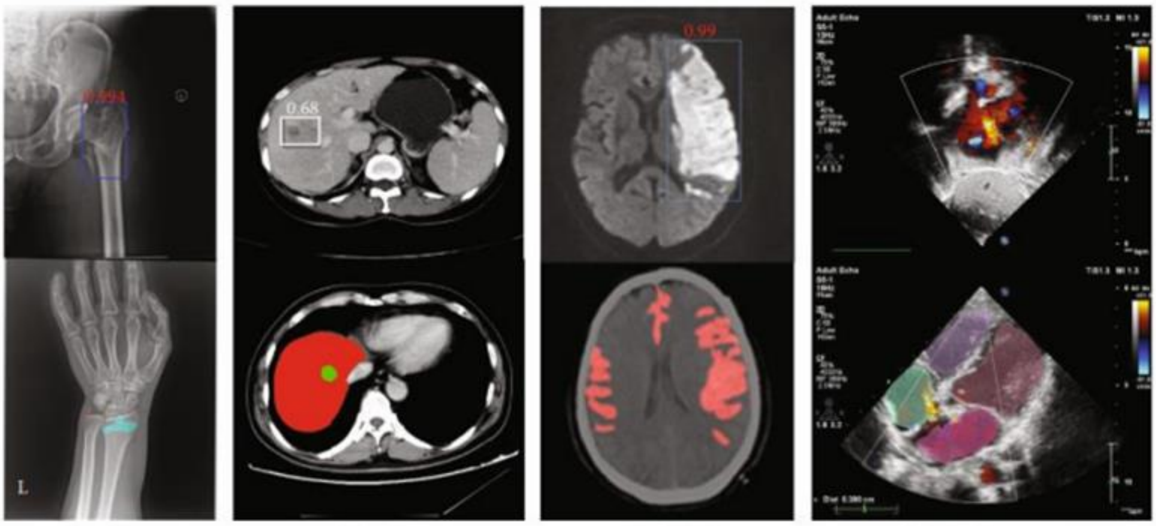

Figure 4 Examples of medical image modalities and their corresponding applications **[11]**

Machine learning for Cancer Prediction

Cancer, a widely known and dangerous term encompassing hundreds of diseases, begins when abnormal cells exceed normal limits. Over the last two decades, researchers have diligently explored various cancer types, providing accurate solutions. Lung, skin, blood, breast, and stomach cancers rank among the most common. **[12]** ML and data mining techniques have emerged to predict cancers, with some models demonstrating high accuracy, particularly in breast cancer prediction. Recent studies have optimized breast cancer models using well-known algorithms. Various articles and reviews explore cancer research, treatment, and subtypes.

Researchers analyze diverse patient groups to comprehend findings, with many studies indicating the potential of ML to predict breast cancer survival rates. Utilizing algorithms like Decision Tree (C4.5), along with SVM, along with ANN, investigations aim to build predictive models for breast tumor recurrence. The study focuses on linking sensitivity, specificity, and precision to the performance of these algorithms, anticipating that SVM will efficiently predict breast tumor reoccurrence.

Advantages of Machine Learning in Health Care

The integration of ML in healthcare has resulted in a wealth of data along with analytics for analysis, offering several advantages. By leveraging modern algorithms with patient data, ML aids medical professionals in accurate illness screening, improving patient care, and serving as clinical decision support without replacing practitioners. It plays a crucial role in rapidly identifying infections and diseases, potentially saving lives through early diagnosis. In radiotherapy, ML systems detect changes in healthy and malignant cells, optimizing treatment focus. The growth of electronic health records provides extensive medical data, enabling ML to enhance treatment strategies and explore medical circumstances. Furthermore, ML reduces hospital readmissions and overcomes length-of-stay challenges, improving patient satisfaction and decreasing annual mortality rates. Prognostic capabilities extend to predicting patients' propensity to pay and anticipating appointment no-shows, contributing to efficient resource utilization and patient support. Overall, ML revolutionizes healthcare by enhancing diagnostics, treatment planning, and patient outcomes

Challenges

The CRISPR-Cas9 system has the ability to edit along with replace genes that are harmful in personalized therapy. To predict the unforeseen effects of this gene editing technology, ML approaches have been employed. A technique called the CDR scan has recently been advanced to predict drug responsiveness based on somatic mutation profiles. It links the tumor genomic fingerprint with drug sensitivity. This approach has the potential to lead to patient-specific treatments by identifying 14 oncology medications and 23 non-oncology pharmaceuticals with novel cancer indications. In order to predict the effects of medications on the developing baby, researchers in drug safety used a random forest classifier. Class C medications that are harmful to pregnant women or may cause birth defects were successfully separated from those that are safe for use during pregnancy by the models.

Conclusion

In conclusion, the collaboration of ML with AI has significantly advanced disease tracking and forecasting globally. Access to extensive data from satellites and websites empowers artificial neural networks to predict illnesses, particularly benefitting underprivileged nations with limited medical resources. ProMed-Mail, a ML tool, successfully predicted the Ebola epidemic in 2014, showcasing its real-time monitoring capabilities. The paper underscores the widespread ML utilisation in healthcare, optimizing patient care, diagnosis accuracy, and resource utilization. These tools, when used in conjunction with doctors, enhance the patient experience and provide accurate predictions for informed decision-making. DL algorithms like CNNs show promise in detecting diseased organs from medical images. While ML revolutionizes patient care from admission to diagnosis, challenges remain, requiring curated data and addressing bottlenecks for broader adaptability. Overall, ML holds immense potential in shaping the future of healthcare, contributing to personalized treatment, pandemic monitoring, and informed citizen preparedness.

Reference

[1] S. Brown, "Machine learning, explained.," MIT SloanSchool Of Management, 21 04 2021. [Online]. Available: https://mitsloan.mit.edu/ideas-made-to-matter/machinelearning-explained..

[2] R., H. Chai, J. Niu, F. Carrasco, H. Arvin, Yin, and B. Lennox, "Design and experimental validation of deep reinforcement learning-based fast trajectory planning and control for mobile robot in unknown environment," *IEEE Transactions on Neural Networks and Learning Systems*, 2022..

[3] K. G. Dinesh, K. Arumugaraj, K. D. Santhosh and V. Mareeswari, "Prediction of cardiovascular disease using machine learning algorithms," In 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT), pp. 1-7, 2018..

[4] M. Kumari and S. Vijendra, "Big data analytics in healthcare: opportunities, challenges and techniques," *International Journal of Social Computing and Cyber-Physical Systems*, vol. 2, no. 1, pp. 35-58, 2017..

[5] M. A., Ahmad, C., Eckert, & A. Teredesai, “Interpretable machine learning in healthcare” In *Proceedings of the 2018 ACM international conference on bioinformatics, computational biology, and health informatics* pp. 559-560, 2018

[6] Z., Li, K. Roberts, X Jiang,., & Q Long,. “Distributed learning from multiple EHR databases: contextual embedding models for medical events”. *Journal of biomedical informatics*, *92*, 103138. 2019.

[7] L. D. Butler, P. Guzzie-Peck, J. Hartke, M. S. Bogdanffy, Y. Will, D. Diaz and J. J. DeGeorge, "Current nonclinical testing paradigms in support of safe clinical trials: an IQ consortium DruSafe perspective," Regulatory Toxicology and Pharmacology, vol. 87, pp. 1-15, 2017.

[8] S. Ekins, A. C. Puhl, K. Zorn, T. R. Lane, D. P. Russo, J. J. Klein and A. M. Clark, "Exploiting machine learning for endto-end drug discovery and development," Nature materials, vol. 18, no. 5, pp. 435-441, 2019.

[9] M., Shehab, L., Abualigah, Q., Shambour, M. A., Abu-Hashem, M. K. Y., Shambour, A. I., Alsalibi, & A. H. Gandomi, “Machine learning in medical applications: A review of state-of-the-art methods”. Computers in Biology and Medicine, 145, 105458. 2022.

[10] A. Mujumdar and V. Vaidehi, "Diabetes prediction using machine learning algorithms," Procedia Computer Science, vol.165, pp. 292-299, 2019.

[11] M. Kim, J. Yun, Y. Cho, K. Shin, R. Jang, H. Bae and N. Kim, "Deep Learning in Medical Imaging," Neurospine, vol. 16, no. 4, p. 657, 2019..

[12] Y. Xiao, J. Wu, Z. Lin and X. Zhao, "A deep learning - based multi -model ensemble method for cancer prediction," Computer methods and programs in biomedicine, vol. 153, pp. 1-9, 2018.