International Research Journal of Education and Technology
Peer Reviewed Journal, **ISSN: 2581-7795**

STUDY ON MACHINE LEARNING ON HEALTH CARE

Dr. B. Anuja Beatrice[1] , Sujeethrah R[2]

Associate professor and head of department[1], BCA Student[2]

Sri Krishna Arts and Science College, Coimbatore[1,2]

Abstract:

The incorporation of machine learning into clinical medicine holds promise for substantially improving health care delivery. Private companies are rushing to build machine learning into medical decision making, pursuing both tools that support physicians and algorithms designed to function independently of them. Physician-researchers are predicting that familiarity with machine-learning tools for analyzing big data will be a fundamental requirement for the next generation of physicians and that algorithms might soon rival or replace physicians in fields that involve scrutiny of images, such as radiology and anatomical pathology [1]. However, consideration of the ethical challenges inherent in implementing machine learning in health care is warranted if the benefits are to be realized. Some ethical challenges are straightforward and need to be guarded against, such as concerns that algorithms may mirror human biases in decision making. Recent advancements in Artificial Intelligence (AI) and Machine Learning (ML) technology have brought on substantial strides in predicting and identifying health emergencies, disease populations, and disease state and immune response, amongst a few. Although skepticism remains regarding the practical application and interpretation of results from ML-based approaches in healthcare settings, the inclusion of these approaches is increasing at a rapid pace [2].

Keywords: Machine Learning (ML); Drug Discovery; Artificial intelligence (AL)

INTRODUCTION

The ML is a branch of Artificial Intelligence (AI) that aims to develop and apply computer algorithms that learn from raw, unprocessed data, to later perform a specific task. The main tasks performed by the AI algorithms are classification, regression, clustering or pattern recognition within a large data set [3]. There is a great variety of ML methods that have been used in the pharmaceutical industry for the prediction of new molecular characteristics, biological activities, interactions and adverse effects of drugs. Machine learning (ML), a field of computer science and a part of artificial intelligence, refers to the science and engineering by which machines (i.e., computer systems) can analyze data and "learn" information from that data. "Learned" information is then used to explain processes and gain additional knowledge from data (e.g., prediction of outcomes). Alternatively, ML denotes an algorithm or set of algorithms often implemented in software that takes as inputs extracted features from data and returns as an output acquired knowledge [4]. However, consideration of the ethical challenges inherent in implementing machine learning in health care is warranted if the benefits are to be realized. Some ethical challenges are straightforward and need to be guarded against, such as concerns that algorithms may mirror human biases in decision making [5].

Applications of Machine Learning in Healthcare

The following are a few main areas of Machine Learning which can be included in healthcare systems [6].

Unsupervised Learning: Unsupervised learning approaches are Machine Learning techniques that use unlabeled data. Unsupervised learning techniques include clustering sets of data using a similarity attribute and dimensionality reductions to allocate high dimensional data to low dimensional feature space. Unsupervised learning can also be used to identify anomalies, such as grouping. Prediction of cardiovascular problems using grouping and diagnosis of gastrointestinal disorder using Principal factor analysis (PFA), a dimensional reduction technique, are two good examples of unsupervised learning approaches in healthcare.

Supervised Learning: Supervised learning approaches are those that use marked testing data to construct or model the relationship between the sources and outcomes. The task is referred to as identification if the performance is temporal and correlation if the output is consistent. The detection of various forms of respiratory illnesses (lumps) and the identification of different internal organs from medical data are two good examples of supervised learning approaches in medical care. Machine Learning methods cannot be supervised or unsupervised when the training set contains both categorized and unlabeled samples. Semi-supervised learning approaches make use of such results.

Semi supervised Learning: If both labelled and unlabeled data models are obtainable for practice such as a minor number of labelled images and a substantial quantity of unlabeled data, semi-supervised learning approaches are useful. Due to the difficulty of collecting a sufficient volume of labelled data for the training process of medical care, semi-supervised learning methods can be particularly beneficial for several healthcare applications. Different dimensions of semi-supervised learning utilizing different learning approaches have been proposed in the research [7]. For example, it presents a semi-supervised clustering approach for health information and presents a semi supervised Machine Learning method for person identification consuming sensor information.

Reinforcement Learning: Reinforcement learning (RL) approaches are those that acquire a decision function from a series of findings, behaviors, and incentives in addition to behavior performed over time. Many healthcare applications could benefit from RL, and it is generally utilized for disease detection using context-aware condition analysis. Furthermore, the Go test, in which a computer utilizing RL and a mixture of supervised and unsupervised learning methods defeated a professional champ player, exemplifies the potential of using RL in intelligent healthcare.

International Research Journal of Education and Technology
Peer Reviewed Journal, **ISSN: 2581-7795**

Data Source	ML Process	Final Output
Medical Images [cite: 68, 69]	Pattern Recognition	Cancer Diagnosis [cite: 68, 69]
Patient Profiles [cite: 56, 67]	Causal Inference	Personalized Treatment [cite: 86, 55]
Historical Records [cite: 93]	Regression/Classification [cite: 9]	Readmission Probability [cite: 93, 94]
Unlabeled Samples [cite: 19, 79]	Clustering	Identification of Anomalies [cite: 20]

Table 1 : ML Process

PREDICTIVE MODELS FOR EARLY DIESEASE DETECTION

This paper reviews applications of machine learning (ML) predictive models in the diagnosis of chronic diseases. Chronic diseases (CDs) are responsible for a major portion of global health costs. Patients who suffer from these diseases need lifelong treatment. Nowadays, predictive models are frequently applied in the diagnosis and forecasting of these diseases. In this study, we reviewed the state-of-the-art approaches that encompass ML models in the primary diagnosis of CD. The first section, titled" Predictive Power [8]: Machine Learning's Role in Early Disease Detection," introduces the overall theme and significance of leveraging machine learning for proactive healthcare strategies. Subsequent sections delve into particulars, highlighting the complex relationship between machine learning and early disease detection. The article" Unleashing the Potential: How Machine Learning Enhances Early Disease Detection" analyzes the multidimensional capabilities of machine learning in analyzing complex data to identify correlations that underlie early disease symptoms [9]. The article" A Primer on Predictive Models: Understanding the Core Concepts in Disease Detection" explains the fundamental principles of predictive models and their function in identifying patterns within data." From Pixels to Diagnoses: The Role of Imaging Data in Machine Learning-Driven Disease Detection" demonstrates how machine learning algorithms excel at analyzing medical images to detect subtle anomalies, thereby improving diagnostic accuracy." Challenges and Opportunities: Navigating Ethical and Technical Considerations in Predictive Disease Detection" delves into the ethical implications of data privacy, bias, interpretability, and accountability, while also addressing technical obstacles such as data quality and model validation. The following sections highlight the convergence of clinical expertise and machine learning. Weisenberg Data Predictive Analytics Model for Disease Prediction using Naive Bayes Technique (BPA-NB). It provides probabilistic classification based on Bayes' theorem with independent assumptions between the features. Naive Bayes approach suitable for huge data sets especially for bigdata [10]. The Naive Bayes approach trains the heart disease data taken from UCI machine learning repository. Then, it was making predictions on the test data to predict the classification. The results reveal that the proposed BPA-NB scheme provides better accuracy about 97.12% to predict the disease rate. The proposed BPA-NB scheme used Hadoop-spark as a big data computing tool to obtain significant insight on healthcare data. The experiments are done to predict different patients' future health condition. It takes the training dataset to estimate the health parameters necessary for classification. The results show the early disease detection to figure out future health of patients.

PERSONALIZED TREATMENT RECOMMANDATIONS USING

MACHINE LEARNING ALGORITHMS

Advancements in artificial intelligence (AI) and machine learning (ML) have sparked a paradigm shift in healthcare, offering unprecedented opportunities to revolutionize patient care. This research paper explores the transformative potential of AI-driven solutions in the healthcare landscape, specifically focusing on the utilization of machine learning techniques to deliver personalized and tailored medical interventions. The paper begins by highlighting the current challenges within traditional healthcare systems, emphasizing the need for more individualized approaches to patient care [11]. It delves into the application of machine learning algorithms in healthcare settings, elucidating how these technologies enable the analysis of vast amounts of patient data to derive actionable insights for personalized treatment strategies. Various case studies and examples illustrate the practical implementation of machine learning in healthcare, showcasing its efficacy in disease prediction, diagnosis, treatment planning, and outcome forecasting.[12] Conversely, real-world observational data, such as electronic health records (EHRs), contain large amounts of clinical information about heterogeneous patients and their response to treatments. In this paper, we introduce the main opportunities and challenges in using observational data for training machine learning methods to estimate individualized treatment effects and make treatment recommendations. We describe the modeling choices of the state-of-the-art machine learning methods for causal inference, developed for estimating treatment effects both in the cross-section and longitudinal settings. The rapid advancements in healthcare technologies and the increasing availability of patient data have paved the way for the development of intelligent medication recommendation systems. These systems leverage the power of artificial intelligence and machine learning to provide personalized medication recommendations, thereby improving treatment outcomes and patient well-being [13]. This research paper explores the design, development, and evaluation of such a system, emphasizing its ability to harness patient-specific information, medical knowledge, and historical treatment data to generate tailored medication recommendations. The proposed system employs sophisticated algorithms to analyze patient profiles, medical histories, and relevant clinical guidelines to offer evidence-based and patient-centered medication suggestions.

AUTOMATED DIAGNOSIS SYSTEM BASED ON MEDICAL IMAGING DATA

Artificial intelligence (AI) continues to garner substantial interest in medical imaging. The potential applications are vast and include the entirety of the medical imaging life cycle from image creation to diagnosis to outcome prediction. The chief obstacles to development and clinical implementation of AI algorithms include availability of sufficiently large, curated, and representative training data that includes expert labeling (eg, annotations). Current supervised AI methods require a curation process for data to optimally train, validate, and test algorithms. Cancer is one of the deadliest diseases diagnosed among the population across the globe so far. The number of cases is increasing at a high pace each year that subsequently leads to the advancement in different diagnosis tools and technologies to handle this pandemic [14]. A significant increase in the mortality rate worldwide leads tremendous scope to devices and implement latest computer aided diagnostic systems for its early detection. The one among such techniques is machine learning coupled with medical imaging modalities. This combination has proven to be efficient in diagnosing various medical conditions in cancer diagnosis. Current study presents a review of different machine learning techniques applied on imaging modalities for cancer diagnosis from 2008 to 2019. This study focuses on diagnosis of five most prevalent and deadly cancers i.e., cervical cancer, oral cancer, breast cancer, brain cancer and skin cancer. Extensive and exhaustive review was carried out after going through different research papers, research articles and book chapters published by reputed international and national publishers such as Springer Link, Science Direct, IEEE Xplore Digital library and PubMed. Machine learning is a technique for recognizing patterns that can be applied to medical images. Although it is a powerful tool that can help in rendering medical diagnoses, it can be misapplied. Machine learning typically begins with the machine learning algorithm system computing the image features that are believed to be of importance in making the prediction or diagnosis of interest [15]. The machine learning algorithm system then identifies the best combination of these image features for classifying the image or computing some metric for the given image region. There are several methods that can be used, each with different strengths and weaknesses. There are open-source versions of most of these machine learning methods that make them easy to try and apply to images.

MACHINE LEARNING OF DRUG DISCOVERY

The popularity of machine learning (ML) across drug discovery continues to grow, yielding impressive results. As their use increases, so do their limitations become apparent. Such limitations include their need for big data, sparsity in data, and their lack of interpretability. It has also become apparent that the techniques are not truly autonomous, requiring retraining even post deployment. In this review, we detail the use of advanced techniques to circumvent these challenges, with examples drawn from drug discovery and allied disciplines. In addition, we present emerging techniques and their potential role in drug discovery. [16] The techniques presented herein are anticipated to expand the applicability of ML in drug discovery.

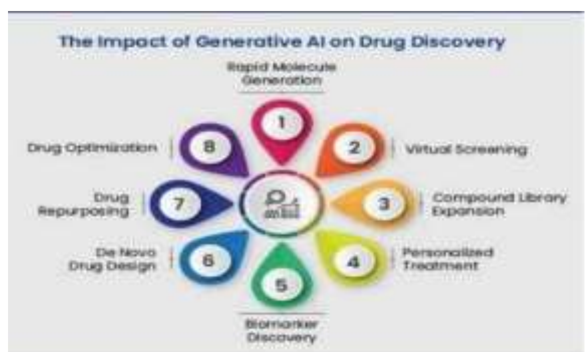

Fig 1 : The impact of generative AI on Drug Discovery

Conventional MLTs have been thoroughly explored in drug discovery.[18] Such techniques include both supervised and unsupervised MLTs, including k-Nearest Neighbour (kNN), decision tree, random forest, support vector machines (SVM), artificial neural networks (ANN), principal component analysis (PCA), and k-means. Their appeal stems from their simplicity, computationally undemanding, yet improved prediction accuracy compared with traditional predictive algorithms. Equally, the underlying mechanisms for conventional techniques can be cognitively comprehended by noncomputer scientist researchers. For example, for kNN, the user has one parameter to control, the k value, which in turn determines the classification search space based on a plurality vote.[17] Another example is SVM, which delineates categories using a hyperplane in conjunction with support vectors to maximize the distance between the different categories.

PREDICTING PATIENT OUTCOMES AND HOSPITAL READMISSIONS USING ML

The number of critically ill patients has increased globally along with the rise in emergency visits. Mortality prediction for critical patients is vital for emergency care, which affects the distribution of emergency resources. Traditional scoring systems are designed for all emergency patients using a classic mathematical method, but risk factors in critically ill patients have complex interactions, so traditional scoring cannot as readily apply to them. As an accurate model for predicting the mortality of emergency department critically ill patients is lacking, this study's objective was to develop a scoring system using machine learning optimized for the unique case of critical patients in emergency departments. Injury is the leading cause of mortality for persons ages 1 to 44 in the United States.

Hospital readmission prediction is a study to learn models from historical medical data to predict probability of a patient returning to hospital in a certain period, e.g. 30 or 90 days (about 3 months), after the discharge. The motivation is to help health providers deliver better treatment and post-discharge strategies, lower the hospital readmission rate, and eventually reduce the medical costs.[19] Due to inherent complexity of diseases and healthcare ecosystems, modeling hospital readmission is facing many challenges. By now, a variety of methods have been developed, but existing literature fails to deliver a complete picture to answer some fundamental questions, such as what the main challenges and solutions in modeling hospital readmission what typical features are used for readmission prediction; how to achieve meaningful and transparent predictions for decision making. In this paper, we systematically review computational models for hospital readmission prediction, and propose a taxonomy of challenges featuring four main categories:

(1) data variety and complexity

(2) data imbalance, locality and privacy

(3) model interpretability

(4) model implementation.

The review summarizes methods in each category, and highlights technical solutions proposed to address the challenges. In addition, a review of datasets and resources available for hospital readmission modeling also provides firsthand materials to support researchers and practitioners to design new approaches for effective and efficient hospital readmission prediction.[20]

MACHINE LEARNING ON ANALYSING ELECTRONIC HEALTH RECORDS AND PATIENT DATA

Advances in precision medicine will require increasingly individualized prognostic assessments for patients to guide appropriate therapy. Conventional statistical methods for prognostic modelling require significant human involvement in selection of prognostic variables (based on an understanding of disease aetiologia), variable transformation and imputation, and model optimization on a per-condition basis.[21]

Electronic health records (EHR) contain large amounts of information about patients' medical history including symptoms, examination findings, test results, prescriptions and procedures. The increasing quantity of data in EHR means that they are becoming more valuable for research. Although EHR is a rich data source, many of the data items are collected in a non-systematic manner according to clinical need, so missingness is often high [22]. The population of patients with missing data may be systematically different depending on the reason that the data are missing tests may be omitted if the clinician judges they are not necessary, the patient refuses, or the patient fails to attend. Multiple imputation to handle missing data can be computationally intensive and needs to include sufficient information about the reason for missingness to avoid Bia.

Accordingly, our primary focus is to provide a comprehensive overview of the types of ML models and techniques used by researchers, as well as types of data on which these models were trained, how the models were validated and, where done, how they were then tested.

Electronic health records (EHR) consist of broad, numerous and erratic access through self-authorizations and "break the glass" scenarios.[25] This is to fulfil the availability aspect of the CIA (confidentiality, integrity) due to the time sensitive nature in healthcare especially during health emergency situations. Adversaries can use this as an opportunity to illegitimately access patient's records, thereby compromising the entire EHR system.

CONCLUSION:

ML has resulted in important contributions to a few disciplines in recent years, including vision and natural language processing. In these fields, more complex models can take advantage of the large amount of existing training data (e.g., images in vision or sentences in natural language). Similarly, we are on the verge of a major shift in HE.[26] Through the appropriate application of ML to increasingly available electronic health data including genomic data healthcare epidemiologists will be able to better understand the underlying risk for acquisition of infectious diseases and transmission pathways, develop targeted interventions, and reduce HAIs. While powerful, it is important to remember that ML cannot identify relationships that are not present in the data. Moreover, ML does not replace the need for standard statistical analyses or randomized, control trials. Instead, ML can serve as a tool to augment HE's current toolbox. Going forward, the greatest impact will come from interdisciplinary teams that work together to make sense of the data.

While the overview demonstrates how much progress has been achieved with machine learning, there continues to be potential for widescale advancement in the future. Many of the current machine learning advancements in healthcare aim to support the physician’s or specialist's ability to provide a more effective treatment to patients with increased quality, speed, and precision. The challenges of developing ML algorithms can be solved by developing and implementing improvements in data collection, storage, and dissemination or by creating algorithms to process unstructured data to address the lack of data availability.[27] Future applications can also bring forth inexpensive forms of medical imaging and affordable medical examinations, potentially ending health disparities and creating more accessible services for countries and lower-income populations. Scientists expect advancement in the prediction of personalized drug response, optimization of medication selection and dosage, and an application of genetic modification to provide treatment for genetic disorders and mutations. With its application, ML can augment the role of physicians and redefine patient care. While the risks and challenges of the future application are addressed and corrected, the current ML algorithms can provide an excellent framework for future advancements and applications of ML in healthcare.

References

1. Char, D. S., Shah, N. H., & Magnus, D. (2018). Implementing machine learning in health care—addressing ethical challenges. The New England journal of medicine, 378(11), 981.

2. Habehh, H., & Gohel, S. (2021). Machine learning in healthcare. Current Genomics, 22(4), 291.

3. Callahan, A., & Shah, N. H. (2017). Machine learning in healthcare. In Key advances in clinical informatics (pp. 279-291). Academic Press.

4. Wiens, J., & Shenoy, E. S. (2018). Machine learning for healthcare: on the verge of a major shift in healthcare epidemiology. Clinical infectious diseases, 66(1), 149-153.

5. Jain, V., & Chatterjee, J. M. (2020). Machine learning with health care perspective. Cham: Springer, 1-415.

6. Bala, P. M., Usharani, S., Rajmohan, R., & Leemaroselin, G. (2020). Applications of Machine Learning and Deep Learning for maintaining Electronic Health Records. Global Journal of Internet Interventions and IT Fusion [ISSN: 2582-1385 (online)], 3.

7. Vellido, A. (2020). The importance of interpretability and visualization in machine learning for applications in medicine and health care. Neural computing and applications, 32(24), 18069-18083.

8. Battineni, G., Sagaro, G. G., Chinatalapudi, N., & Amenta, F. (2020). Applications of machine learning predictive models in the chronic disease diagnosis. Journal of personalized medicine, 10(2), 21.

9. Aghili, M., Tabarestani, S., Adjouadi, M., & Adeli, E. (2018). Predictive modeling of longitudinal data for alzheimer’s disease diagnosis using rnns. In PRedictive Intelligence in MEdicine: First International Workshop, PRIME 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Proceedings 1 *(pp. 112-119). Springer International Publishing.*

10. Lee, Y. H., Bang, H., & Kim, D. J. (2016). How to establish clinical prediction models. Endocrinology and Metabolism, 31(1), 38-44.