JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

MACHINE LEARNING IN HEALTHCARE

Rashmi P. Dagde
(Assistance Professor)
Computer Science and Engineering
Priyadarshini Bhagwati College of Engineering
Nagpur, India

Aabha N. Yerpude
(Research Scholar)
Computer Science and Engineering
Priyadarshini Bhagwati College of Engineering
Nagpur, India

Anushka P. Shende
(Research Scholar)
Computer Science and Engineering
Priyadarshini Bhagwati College of Engineering
Nagpur, India

Dimpal J. Charde
(Research Scholar)
Computer Science and Engineering
Priyadarshini Bhagwati College of Engineering
Nagpur, India

Tanaya H. Agre
(Research Scholar)
Computer Science and Engineering
Priyadarshini Bhagwati College of Engineering
Nagpur, India

Abstract: Recent advancements in Artificial Intelligence (AI) and Machine Learning (ML) technology have resulted in significant progress in predicting and identifying health emergencies, disease populations, disease states, and immune responses. Despite this progress, skepticism persists regarding the practical application and interpretation of results from ML-based approaches in healthcare settings. Nevertheless, the inclusion of these approaches is rapidly increasing. In this paper, we present a concise overview of machine learning-based approaches and learning algorithms, including supervised, unsupervised, and reinforcement learning, along with relevant examples. Additionally, we discuss the application of ML in various healthcare fields, such as radiology, genetics, electronic health records, and neuroimaging. We also briefly address the risks and challenges of ML application to healthcare, such as system privacy and ethical concerns, and provide recommendations for future applications.

Index Terms Artificial Intelligence, healthcare, Machine Learning, Electronic health records

1. INTRODUCTION:

The application of machine learning can be traced back to the 1950s, when Alan Turing proposed the first machine able of literacy and achieving artificial intelligence. Since its commencement, machine learning has been employed in a different range of operations, gauging from security services exercising facial recognition to enhancing effectiveness and reducing threat in public transportation, and more lately, in colorful aspects of healthcare and biotechnology. Artificial intelligence and machine learning have brought about significant changes in business processes and have converted day- to- day lives, with similar metamorphoses anticipated in healthcare and drug. Recent advancements in this field have demonstrated remarkable progress and eventuality to palliate the burden on croakers and ameliorate the delicacy, vaticination, and quality of care. Presently, machine learning advancements in healthcare primarily serve as a probative part in a croaker or critic's capability to fulfill their places, identify healthcare trends, and develop complaint vaticination models. In large medical associations, machine literacy-grounded approaches have also been enforced to achieve increased effectiveness in the association of electronic health records, identification of irregularities in blood samples, organs, and bones using medical imaging and monitoring, as well as in robot- supported surgeries. Machine learning operations have lately enabled the acceleration of testing and sanitarium response in the battle against COVID- 19. Hospitals have been suitable to organize, partake, and track cases, beds, apartments, ventilators, electronic health records, and indeed staff during the epidemic using a deep literacy system by GE called the Clinical Command Center. Experimenters have also employed artificial intelligence for the identification of inheritable sequences of SARS- CoV2 and the creation of vaccines, as well as for their monitoring. As the healthcare assiduity continues to evolve with the integration of new technologies, artificial intelligence and machine learning - grounded approaches and operations are pivotal for its progression, including increased speed of opinion, delicacy, and simplicity. The purpose of this review is to punctuate the advantages and disadvantages of machine learning- grounded approaches in the healthcare assiduity. As the operation of new machine learning technology takes the healthcare assiduity by storm, we aim to give a brief overview of the colorful approaches to machine learning and punctuate the fields where these approaches are primarily applied. We bandy their wide use and unborn advancement openings in healthcare. We also address the ethical and logistical pitfalls and challenges that arise with their operation.

2. a. ARTIFICIAL INTELLIGENCE

While the terms machine learning, deep learning, and artificial intelligence are often used interchangeably, they each encompass distinct sets of algorithms and learning approaches. Artificial Intelligence (AI) serves as the overarching concept, encompassing computerized systems that learn and emulate human intelligence. AI is renowned for its applications in autonomous machines like robots and self-driving cars, but it also pervades everyday tools such as personalized ads and web searches. Recent years have witnessed remarkable progress in AI development and application, thanks to its superior decision-making abilities, accuracy, problem-solving prowess, and computational skills. In the development of AI algorithms, a fundamental practice involves splitting acquired data into two categories: a training dataset and a testing dataset. This division ensures reliable learning, representative data, and unbiased predictions. As the name implies, the training dataset is used to train the algorithm, encompassing sets of defining data points (features) and corresponding predictions (in the case of supervised learning). In contrast, the testing dataset is entirely new to the algorithm and serves solely to assess the algorithm's performance. This precautionary measure eliminates biases in the algorithm's evaluation stemming from its familiarity with the training dataset. Once an algorithm successfully completes both the training and testing phases with satisfactory results, it can be deployed in various healthcare settings. AI finds widespread applications across numerous domains, including machine learning and deep learning, which are two prominent subfields within the AI realm. Machine learning encompasses a variety of algorithmic models and statistical techniques designed to solve problems without the need for specialized programming. Many machine learning models are single-layered, necessitating extensive feature extraction and data preprocessing before data input. This preprocessing is crucial to ensure accurate predictions and prevent issues like overfitting or underfitting to the training dataset. Deep learning, on the other hand, represents a more advanced subset of machine learning that employs layered artificial neural networks to achieve higher accuracy and specificity at the cost of reduced interpretability. The essence of deep learning lies in its multilayered neural networks, facilitating connections between artificial neurons or units across successive layers. These networks can independently learn, discern, and infer from data by utilizing these multi-level connections for data processing, refining the data until specialized outcomes are achieved.

2.b. TYPES OF LEARNING APPROACHES

Many machine learning and AI-based algorithms are founded on various learning methodologies. One such category is supervised learning, which is employed to train classification and prediction algorithms based on prior instances or outcomes. A crucial distinction in this learning paradigm is that the training set comprises features along with their corresponding predictions or results. To simplify, supervised learning extrapolates insights from the features within the training set to construct a model capable of accurately forecasting outcomes within that training set. Subsequently, this learned model is applied to make predictions using new features present in the testing dataset. Examples of machine learning algorithms that employ supervised learning techniques include Decision Trees, Random Forest, Support Vector Machines, and Artificial Neural Networks. Decision tree algorithms function as decision-support tools that commence with a single node, exploring potential outcomes at each juncture. The tree evolves by multiplying the consequences of each decision with subsequent decisions until reaching a final outcome. Support Vector Machines (SVM) are renowned classification algorithms that leverage supervised learning to segregate features into two groups by identifying the most spacious margin hyperplane for optimal data separation. Artificial Neural Networks (ANNs) consist of input layers, one or more hidden layers, and output layers, with interconnected functional units or neurons in each layer linked to all neurons in adjacent layers. In the healthcare domain, supervised machine learning is widely applied in tasks such as disease prediction, hospital outcome identification, and image detection, to name a few. Another branch of AI-based learning approaches is unsupervised learning, primarily utilized for data assessment and clustering purposes. Unsupervised machine learning generally serves the objective of data analysis, stratification, and reduction rather than prediction. Typically, unsupervised clustering techniques utilize algorithms to group unclassified or uncategorized data into distinct clusters. While most forms of machine learning involve data preprocessing and feature extraction prior to input, this method enables feature extraction and investigates data clustering possibilities by recognizing inherent relationships or features within the data and grouping them based on similarities. Some unsupervised learning approaches encompass the k-Means algorithm, Deep Belief Networks, and Convolutional Neural Networks. The k-Means algorithm, a widely used unsupervised technique, functions as a clustering tool to determine the mean between groups within unlabeled datasets and subsequently create clusters based on these means. A Deep Belief Network (DBN) is a multi-layered network featuring intra-level connections designed for data retrieval. It typically employs unsupervised learning and incorporates multiple hidden layers tasked with feature identification and correlation detection within the data. A Convolutional Neural Network (CNN) is a multilayered network specializing in feature recognition and identification, commonly employed for tasks such as anomaly detection, image recognition, and object identification. Although unsupervised methods are invaluable for clustering due to the absence of predefined outcomes and data homogeneity, their popularity in healthcare remains somewhat limited. Reinforcement learning, a distinct learning methodology, does not fall into either the supervised or unsupervised learning categories. It relies on reward sequences, akin to conditioning mechanisms in psychology, to formulate a strategy for problem-solving within a particular domain. Reinforcement learning methods possess the capability to impact their environment, prioritize error criterion optimization, and bear resemblance to learning processes observed in humans and animals. Among the neural networks associated with reinforcement learning, the Recurrent Neural Network (RNN) is frequently employed. RNNs connect every artificial neuron, permitting inputs with time delays and the reuse of outputs from previous steps as input for future steps. RNNs find utility in tasks like time series prediction, translation, speech recognition, rhythm learning, and music composition. Although the healthcare applications of reinforcement learning are somewhat restricted due to prerequisites such as structured data, heterogeneous data sources, reward definition and implementation, and substantial computational resources, it still holds significant potential for making substantial advancements in healthcare. Given the various machine learning and deep learning approaches, selecting the appropriate approach for a healthcare application is of utmost importance. Numerous factors, including feature count, sample size, and data distribution, can significantly influence the learning and prediction processes and should be carefully considered.

3. RISKS AND CHALLENGES

While machine learning-based applications in the healthcare sector offer innovative possibilities, they also introduce distinctive risks, challenges, and a need for cautious consideration. In this discussion, we delve into key risk factors, including the potential for prediction errors and their consequences, the security and privacy vulnerabilities of systems, and the issue of data availability for achieving reproducible results. Additionally, we explore challenges such as ethical dilemmas, the diminishing human touch in healthcare, and the interpretability and practical implementation of these approaches in clinical settings.

One of the foremost risks associated with machine learning-based algorithms is their reliance on probabilistic distributions and the possibility of errors in diagnosis and prediction. This gives rise to valid skepticism regarding the accuracy and reliability of predictions made by ML-based methods. Although healthcare has long dealt with probabilities and uncertainties, the implications of ML-based errors leading to adverse human outcomes are profound. One mitigation strategy is subjecting these ML-based approaches to rigorous institutional and legal scrutiny by multiple organizations before their deployment. Another approach involves introducing human intervention and oversight from experienced healthcare professionals in highly sensitive applications to prevent false-positive or false-negative diagnoses (e.g., depression or breast cancer). Involving current healthcare professionals in the development and implementation of these approaches can enhance acceptance rates and address concerns related to potential job displacement.

Another risk in applying ML and deep learning algorithms to healthcare lies in the availability of high-quality training and testing data with sufficiently large sample sizes to ensure robust and reproducible predictions. Since ML and deep learning methods "learn" from data, the importance of data quality cannot be overstated. Furthermore, the extensive, feature-rich data required for these learning models often faces scarcity and may represent a limited cross-section of the population. In many healthcare domains, data collected is incomplete, heterogeneous, and contains a high feature-to-sample ratio. These challenges must be carefully considered during the development and interpretation of results from ML-based approaches. Initiatives like open science and the growing emphasis on sharing research data may aid in surmounting these hurdles. Additionally, data privacy risks and ethical concerns must be addressed comprehensively, considering the sensitive nature of healthcare data and the adoption of cloud-based technologies in ML model development.

Addressing ethical concerns, researchers applying ML-based approaches to healthcare can draw lessons from the field of genetic engineering, which has undergone extensive ethical debates. Genetic engineering has the potential to provide life-changing treatments by identifying and editing harmful genetic mutations. However, it also raises concerns about altering an individual's genome, and that of their offspring, creating disparities in access to such care due to costs. Various guidelines and regulations have emerged, such as Singapore's Model Artificial Intelligence Governance Framework and the U.S. Administration's executive order on AI regulation, to ensure ethical research and development practices in AI and healthcare.

A significant challenge in the application of ML to healthcare revolves around interpreting and clinically applying the results. ML-based approaches, especially deep learning methods, have intricate structures that make it challenging to discern the contributions of original features to predictions. While this may not be a significant issue in other ML applications, it poses a substantial barrier to the adoption of ML-based approaches in healthcare. In healthcare, understanding the solution strategy is as critical as the solution itself. A shift towards systematically identifying and quantifying underlying data features used for predictions is essential. Involving physicians and healthcare professionals in the development, implementation, and testing of ML-based approaches can improve adoption rates. Furthermore, concerns about ML reducing the personal relationship between patients and primary care physicians (PCPs) can be addressed by utilizing ML to enhance patient engagement. Studies indicate that the physician-patient relationship is diminishing, and ML can provide opportunities for patients to discuss potential diagnoses, improve outreach program efficiency, and promote healthy lifestyle choices. Early prognosis through ML-based approaches can also reduce physician stress and offer more personal time with patients, enhancing patient satisfaction and outcomes.

4. LITERATURE SURVEY

Electronic Health Records (EHRs) :

1.Electronic Health Records (EHRs) have paved the way for groundbreaking diagnostic predictions. In a study by Liang et al. in 2014, Support Vector Machines (SVM) and Decision Trees (DT) were harnessed to analyze EHR data, offering healthcare practitioners the ability to anticipate diagnoses with impressive precision. By leveraging a plethora of patient information, these models consider factors like medical history, demographics, and laboratory results, empowering early disease detection and personalized treatment strategies. This research underscores the transformative potential of data-driven methods in healthcare, promising more efficient and effective patient care.

Medical Imaging:
2. Skin Cancer Classification with Deep Neural Networks –
In the field of dermatology and medical imaging, a game-changing development involves the utilization of Convolutional Neural Networks (CNNs) for the high-precision classification of skin cancer types. This cutting-edge approach, outlined in Esteva et al.'s 2017 study, aspires to match the diagnostic accuracy of dermatologists. By harnessing the capabilities of deep neural networks, this technology can scrutinize intricate skin lesion images with exceptional accuracy. Through extensive training on large datasets, CNNs hold the promise of significantly improving early skin cancer detection, ultimately leading to better patient outcomes and a reduced healthcare burden. Esteva et al.'s work signifies a remarkable step toward integrating artificial intelligence into dermatology, potentially transforming the landscape of skin cancer diagnosis and care.

Genetic Engineering & Genomics:
3.A noteworthy application in this area involves predicting the stable doses of tacrolimus, a vital immunosuppressant used in renal transplant cases. This predictive task has been addressed using machine learning models, specifically Regression Trees (RT), as detailed in the 2017 study by Tang et al. By harnessing the capabilities of these models, healthcare providers can proactively determine the most suitable tacrolimus dosage for individual patients, taking into account various factors such as patient demographics, medical history, and clinical parameters. This approach not only ensures more effective immunosuppressive treatment but also reduces the chances of rejection or adverse reactions, ultimately leading to improved long-term outcomes and a better quality of life for recipients of renal transplants. Tang et al.'s research highlights the potential of machine learning in personalized medicine, revolutionizing the way patient care is tailored and delivered.

5. REFERENCE

[1] Turing A. Computing machinery and intelligence. Mind. 1950;LIX(236):433–460.: 10.1093/mind/LIX.236.433. [CrossRef] [Google Scholar]

[2] Siddiqui M.K., Morales-Menendez R., Huang X., Hussain N. A review of epileptic seizure detection using machine learning classifiers. Brain Inform. 2020;7(1):5. doi: 10.1186/s40708- 020-00105-1. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

[3] Rao S.R., Desroches C.M., Donelan K., Campbell E.G., Miralles P.D., Jha A.K. Electronic health records in small physician practices: availability, use, and perceived benefits. J. Am. Med. Inform. Assoc. 2011;18(3):271–275. doi: 10.1136/amiajnl-2010-000010. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

[4] Roth M. COVID-19: Oregon hospitals share data, create real-time bed capacity system. Health Leaders Media. 2020 Available from: https://www.healthleadersmedia.com/innovation/covid-19-oregon-hospitals-share-data-create-real-time-bed-capacity-system. [Google Scholar]

[5] Alloghani M., Al-Jumeily D., Mustafina J., Hussain A., Aljaaf A.J. A systematic review on supervised and unsupervised machine learning algorithms for data science. In: Berry M., Mohamed A., Yap B., editors. Supervised and unsupervised learning for data science. Springer, US: 2020. pp. 3–21. [CrossRef]

[6] Alpaydin E. MIT Press; 2020. Introduction to machine learning. Available from: https://books.google.com/books?hl=en&lr=&id=tZnSDwAAQBAJ&oi=fnd&pg=PR7&dq=Introduction+to+machine+learning&ots=F3RWaXcwwf&sig=50DHyEjhVdDtmXIxZ0C4tXOGsdw. [Google Scholar]

[7] Qin Z., Ye H., Li G.Y., Juang B.H.F. Deep learning in physical layer communications. IEEE Wirel. Commun. 2019;26(2):93–99. doi: 10.1109/MWC.2019.1800601. [CrossRef] [Google Scholar]

[8] Li W., Jia F., Hu Q. Automatic segmentation of liver tumor in ct images with deep convolutional neural networks. J. Comput. Commun. 2015;03(11):146–151. doi: 10.4236/jcc.2015.311023. [CrossRef] [Google Scholar]

[9] Yang N., Hing E. National electronic health records survey: 2015 specialty and overall physicians electronic health record adoption summary tables. 2017;(28) Available from: http://www.cdc.gov/nchs/ahcd/ahcd_questionnaires.htm. [Google Scholar]

[10] Cseko G.C., Tremaine W.J. The role of the institutional review board in the oversight of the ethical aspects of human studies research. Nutr. Clin. Pract. 2013;28(2):177–181. doi: 10.1177/0884533612474042. [PubMed] [CrossRef] [Google Scholar]

[11] Finlayson S.G., Bowers J.D., Ito J., Zittrain J.L., Beam L., Kohane I.S. Emerging vulnerabilities demand new conversations. Science (80) 2019;363(6433):1287–1290. Available from: http://science.sciencemag.org/content/363/6433/1287. [PMC free article] [PubMed] [Google Scholar]

[12] Meskó B., Hetényi G., Győrffy Z. Will artificial intelligence solve the human resource crisis in healthcare? BMC Health Serv. Res. 2018;18(1):545. doi: 10.1186/s12913-018-3359-4. [PMC free article] [PubMed] [CrossRef] [Google Scholar]