ISSN: 2581-3404 (Online) IF: 5.86 (SJIF)
IJIRTM, Volume-10, Issue-1, January-2026

A Comparative Review on Role of Machine Learning Techniques in Healthcare

Neetiraj Singh Lodhi[1], Dr. Rajneesh Choubey[2]
M. Tech, Research Scholar[1], Associate Professor[2]
Dept. of CSE, Bansal Institute of Science and Technology, Bhopal, India[1]
Dept. of MCA, Bansal Institute of Science and Technology, Bhopal, India[2]
nrsklodhi@gmail.com[1], choubey.rajnish@gmail.com[2]

Abstract. *The aim of this research is to develop a mechanism that assists medical practitioners in predicting and diagnosing liver disease. Various systems have been proposed to aid medical experts by reducing errors and enhancing accuracy in the diagnosis and prediction of diseases. Machine Learning is employed to unearth patterns in extensive datasets, facilitating decision-making and allowing machines to learn through various processes such as supervised, unsupervised, semi-supervised, or reinforcement learning. Historical and categorized patient data are inputted into several algorithms for predicting future patient outcomes. The predictive algorithms examined include Logistic Regression, Decision Tree, Random Forest, K-Nearest Neighbor, and Artificial Neural Network. The aim of this research work is to study the performance of various Machine Learning algorithms to mitigate the high costs associated with diagnosing disease through prediction.*

Keywords: Classification, Prediction, Machine learning, Healthcare, Liver disease.

Introduction

As per the World Health Organization's 2017 survey, liver disease accounts for 2.95% of total deaths, placing India in the 63rd position globally [1]. The liver, weighing approximately 3 pounds and resembling a reddish-brown color, is the largest internal organ in the human body, composed of left and right lobes. Situated beneath the liver, the gallbladder has a key role in the liver's function of filtering toxic and harmful substances from the blood before its distribution to different body parts. Liver disease is recognized as one of the most dangerous and fatal health issues worldwide [5]. Major causes of liver disease include liver fibrosis, fatty liver, liver cirrhosis, hepatitis infections, excessive alcohol consumption, drug and toxin exposure, and genetic abnormalities. Early diagnosis and treatment of liver steatosis, characterized by abnormal fat accumulation in over 5% of liver cells, is essential for preventing the progression of liver diseases such as hepatocirrhosis and hepatocellular carcinoma. Ultrasound (US) is the primary imaging technique for diagnosing liver steatosis; however, US images often suffer from issues like speckle noise, blurring, and shading, which can negatively impact diagnosis accuracy [4]. The quality of US images is significantly influenced by the effectiveness of speckle noise reduction. Numerous methods have been developed to minimize speckle noise and enhance image quality, but these approaches often face challenges, such as sensitivity to kernel selection and image blurring. Furthermore, considerable research has focused on improving the accuracy of liver steatosis assessment through advanced algorithms, statistical models, and image-processing techniques applied to US images.

Many researchers in the medical field have experimented with various strategies to increase data classification accuracy. Techniques that offer better accurate classification will give more evidence to find possible patients and enhance diagnosis accuracy. Machine learning is making strides in the biomedical field, particularly for predicting and diagnosing liver diseases, enhancing detection, and improving objectivity in medical decision-making. This study focuses on efficiently predicting liver disease outcomes and reducing diagnostic costs by employing various classification techniques: Logistic Regression (LR), K-Nearest Neighbors (KNN), Decision Trees (DT), Support Vector Machines (SVM), Naive Bayes (NB), and Random Forest (RF) [8, 9].

Machine Learning

Machine learning is a subset of Artificial Intelligence that enables computer systems to autonomously learn from their environment through iterative processes, refining their performance based on experience. It involves algorithms that organize and analyze data to derive insights and make predictions without additional programming [11]. Core to machine learning is the training of models on existing data, which can subsequently predict outcomes for new data. The classification of machine learning algorithms includes supervised, unsupervised, semi-supervised, and reinforcement learning. The advancement of machine learning has significantly enhanced data description, while also offering engineering solutions and establishing crucial benchmarks in the field.

Machine learning is a highly regarded programming tool for predicting and classifying vast amounts of data, serving as a pivotal aspect of data mining [13]. Classification, a key method in data mining, aims to group similar data or patterns into a single category. Unlike traditional programming languages, machine learning emphasizes pattern analysis and behavioral analysis, organizing data based on properties such as shape, size, and attributes. The classification process is inherently challenging due to the dynamic nature of continuously updated data [10]. Various classification types exist, including supervised classification, unsupervised classification, and semi-supervised classification. The effectiveness of classification techniques is enhanced through a training process that improves the accuracy and performance of classifiers in recognizing patterns.

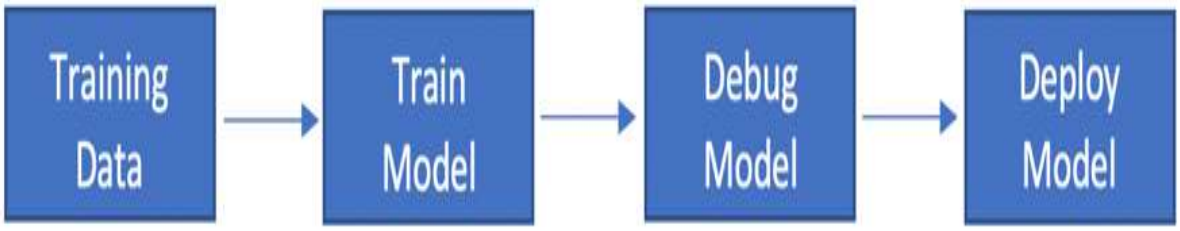

Fig. 1: Steps involved in ML pipeline [15].

Role of Machine Learning in Healthcare

Machine learning, a subset of computer science, plays a crucial role in enhancing computer innovation, especially within healthcare [6]. Key aspects of machine learning include feature extraction, selection, algorithm selection, training, and testing, which collectively enable robust data analysis. Its application in

healthcare aims to expedite and enhance the accuracy of disease prediction and diagnosis, particularly for heart and liver diseases, which are significant health concerns globally. These diseases are leading causes of death and even affect newborns, underscoring the necessity for timely and accurate diagnosis to save lives [14, 16]. Supervised machine learning techniques, such as classification, are employed to categorize illnesses, with methods like decision trees, artificial neural networks (ANN), support vector machines (SVM), and naive Bayes (NB) used for various healthcare objectives.

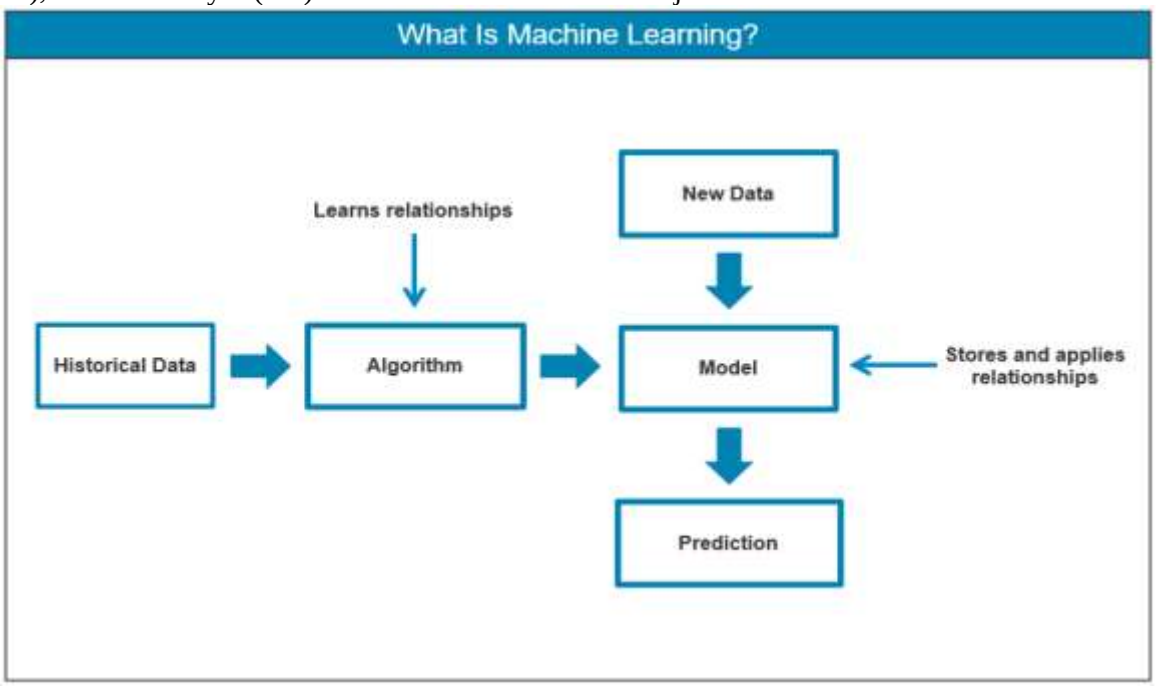

Fig. 2: Machine learning algorithms implementation model [7].

Machine Learning Techniques

Artificial Neural Networks

Artificial Neural Networks (ANN) are advanced algorithms designed to replicate the learning processes observed in biological neural networks. They excel at estimating functions that depend on a multitude of inputs through a structured design that includes multiple layers: an input layer, an output layer, and one or more hidden layers. The primary purpose of ANNs is to train models capable of understanding non-linear decision boundaries to effectively classify different classes of input data [17].

Visually, an ANN can be conceived of as a weighted directed graph composed of nodes (representing artificial neurons) and edges (which are weighted connections between these neurons). In this framework, the output produced by one neuron serves as the input for other neurons. Each neuron processes incoming data as a vector, which can represent patterns or images, enhancing the network's interpretative capabilities. During the training phase, the weights of connections between neurons are fine-tuned to improve classification accuracy on given tasks [9].

Decision tree (DT)

A decision tree (DT) is a predictive model primarily employed in supervised learning environments to tackle machine learning classification challenges. DTs are especially suited for scenarios involving target variables that can take discrete values. The architecture of a decision tree is composed of leaves, branches, and nodes. The leaves denote the class labels, while the branches encapsulate the attributes that contribute to determining those class labels. These branches can handle both discrete and continuous data types effectively.

The DT algorithm operates by partitioning data samples into two or more homogeneous subsets based on a key splitter derived from the input factors. Despite its utility in classification tasks, decision trees are prone to overfitting, a common problem where the model learns noise in the training data instead of the underlying pattern [18]. To mitigate this issue, techniques such as bagging and boosting can be implemented. Overall, decision trees exhibit remarkable efficiency, particularly when applied to discrete datasets, making them a popular choice in the field of machine learning.

Random Forest (RF)

Random Forest (RF) is an ensemble learning technique that constructs multiple decision trees by utilizing random subsets of features. The final class for an instance is determined by a majority vote from these trees, resulting in varying accuracies across iterations. RF is notably robust against overfitting compared to standalone decision trees [14]. It has shown considerable success in Intrusion Detection Systems (IDS) at the network level. Additionally, AdaBoost, another ensemble method, has been applied for intrusion detection in IoT device traffic, which typically suffers from inadequate security measures. A hybrid approach combining Decision Tree, Naive Bayes, and Artificial Neural Networks was employed to detect anomalies, achieving acceptable performance; however, it faced challenges with false positives and exhibited longer processing times than individual algorithms like Decision Tree, Naive Bayes, and Artificial Neural Networks [16].

Logistic Regression

Logistic Regression is a supervised machine learning algorithm that is used for classification. It is utilised for predictive modelling and helps to calculate the possibility of a particular event taking place. It mainly deals with the prediction of binary outcomes for a given set of independent variables and determines the discrete values. It performs the binary classification and predict the future outcomes based on training from the previous output [11].

Conclusion

This work presents a comparative study of various machine learning techniques to develop a efficient model for healthcare services. It emphasizes the use of multiple datasets for enhanced comparison and effectiveness. The study discusses the generation of classification rules and disease identification techniques through various efficient algorithms. The review focuses on different machine learning techniques for early diagnosis in patients, and it proposes an artificial intelligence-based model for predicting more accurate values by utilizing various performance parameters.

References

[1] Harshita Rajpoot, Aaditi Raj, Sachin Bhoite, "Leveraging Machine Learning for Early Detection and Prevention of Liver Disease Mortality in India", Grenze International Journal of Engineering and Technology, 2025, pp. 1-8.

[2] Nidhi Arora, Shilpa Srivastava, "An enhanced predictive modelling framework for highly accurate non-alcoholic fatty liver disease forecasting", Indonesian Journal of Electrical Engineering and Computer Science, 2025, pp. 214-222.

[3] Nilam Solanki, "Machine Learning and Ensemble Learning Based Predictive Modeling for Cirrhosis Stage Classification", Procedia Computer Science, 2025, pp. 1452-1462.

[4] Gangfeng Zhu, Yipeng Song, Zenghong Lu, Qiang Yi, "Machine learning models for predicting metabolic dysfunction-associated steatotic liver disease prevalence using basic demographic and clinical characteristics", Journal of Translational Medicine, 2025, pp. 1-14.

[5] Yaran Chen, Xueyu Chen, "Multimodal Learning-based Prediction for Nonalcoholic Fatty Liver Disease", Machine Intelligence Research, 2025, pp. 871-887.

[6] Pankaj Kawadkar, Deepak Kumar Rathore, Devendra Singh Kushwaha, B Rebecca, G Priyanka Jeeva Karunya, PM Yohan, "Performance analysis of image compression techniques for MRI image using machine learning techniques", AIP Conference Proceedings, 2024, 10.1063/5.0239127.

[7] Ketan Gupta, Nasmin Jawani, Neda Afreen, "Liver Disease Prediction using Machine learning Classification Techniques", IEEE, 2022.

[8] A.K.M Sazzadur Rahman, F. M. Javed Mehedi Shamrat, Zarrin Tasnim, Joy Roy, Syed Akhter Hossain, "A Comparative Study On Liver Disease Prediction Using Supervised Machine Learning Algorithms", International Journal of Scientific & Technology Research Volume 8, Issue 11, November 2019.

[9] Deepak Kumar Rathore, and Sunil Phulre, "Machine Learning Based Efficient Prediction of Human Heart Disease by Identifying the Features", International Journal of Innovative Research in Technology and Management, (2022), pp. 21-27.

[10] Jingwei Wei,Hanyu Jiang,Dongsheng Gu,Meng Niu,Fangfang Fu, Yuqi Han, Jie Tian, "Radiomics in liver diseases: Current progress and futureopportunities", Liver International, 2020.

[11] Mandakini Priyadarshani Behera, Archana Sarangi, Debahuti Mishra, Shubhendu Kumar Sarangi, "A Hybrid Machine Learning Algorithm for Heart and Liver Disease Prediction using Modified Particle Swarm Optimization", International Conference on Machine Learning and Data Engineering, 2023.

[12] Se-Yeol Rhyou and Jae-Chern Yoo, "Cascaded Deep Learning Neural Network for Automated LiverSteatosis Diagnosis Using Ultrasound Images", Journal of Sensors, 2021.

[13] Monica Lupsor-Platon , Teodora Serban, Alexandra Iulia Silion, George Razvan Tirpe, Alexandru Tirpe and Mira Florea, "Performance of Ultrasound Techniques and the Potential of Artificial Intelligence in the Evaluation of HepatocellularCarcinoma and Non-Alcoholic Fatty Liver Disease", Journal of Cancer Treatment, MDPI.

[14] D Rathore, P Mannepalli, "A Review of Machine Learning Techniques and Applications for Health Care ", International Conference on Advances in Technology, Management & Education, 2021, IEEE proceeding, 978-1-7281-8586-6/21.

[15] D.Devikanniga, Arulmurugan Ramu and Ananda kumar Haldorai, " Efficient Diagnosis of Liver Disease using Support Vector Machine Optimized with Crows Search Algorithm", EAI Endorsed Transactions on Energy Web, 2020.

[16] D Rathore, Praveen Kumar Mannepalli, "Recent Trends in Machine Learning for Health Care Sector", International Journal of Innovative Research in Technology and Management, Vol-5, Issue-2, 2021.

[17] Jaime Lynn Speiser, Constantine J. Karvellas, Bethany J. Wolf, Dongjun Chung, "Predicting daily outcomes in acetaminophen-induced acute liver failure patients with machine learning techniques", Computer Methods and Programs in Biomedicine, 2019.

[18] Nazim Razali1, Aida Mustapha, Mohd Helmy Abd Wahab, SalamaA Mostafa1, Siti Khadijah Rostam, "A Data Mining Approach to Prediction of LiverDiseases", Journal of Physics, 2020.