WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Enhancing Heart Disease Diagnosis: A Machine Learning Prediction System for Real-Time Insights

***Sarah Jessica*[1], *Dr. Vanitha Kakollu*[2]**

[1]UG Student, GITAM Deemed to be University, Visakhapatnam
[2]Assistant Professor, GITAM Deemed to be University, Visakhapatnam

ABSTRACT:

Around the world, machine learning is utilized across various domains, and healthcare is no exception. Machine learning proves to be highly beneficial in predicting the likelihood of loco motor disorders, heart diseases, and other medical conditions. When identified well in advance, such predictions provide valuable insights to clinicians, enabling them to tailor diagnostic and treatment strategies for individual patients. This leverages machine learning algorithms to predict potential heart diseases in patients and compare multiple classifiers, including Ada Boost, Extra Trees, and Convolutional Neural Networks (CNN). Additionally, we propose an ensemble classifier that combines both strong and weak classifiers, enhancing performance by leveraging a large number of training and validation samples.

Keywords:Machine Learning, Deep Learning, AdaBoost, Extra Trees, CNN, ML Techniques, Evaluation.

I. INTRODUCTION

Heart disease is a leading cause of mortality worldwide, affecting millions of individuals annually. The early detection and accurate prediction of heart-related ailments are crucial for reducing fatalities and improving patient outcomes. Traditional diagnostic techniques primarily rely on medical expertise and extensive clinical tests, which can be time-consuming, expensive, and sometimes subjective. Recent advancements in artificial intelligence (AI) and machine learning (ML) have introduced data-driven approaches that significantly enhance diagnostic precision and decision-making in healthcare.
Machine learning algorithms analyze vast amounts of patient data, identifying patterns and risk factors associated with heart disease. By leveraging these techniques, healthcare professionals can move toward predictive models that aid in early diagnosis, timely intervention, and personalized treatment plans. This study explores the potential of machine learning in heart disease prediction by evaluating various classifiers, including AdaBoost, Extra Trees, and Convolutional Neural Networks (CNN). Furthermore, we introduce an ensemble learning approach that integrates multiple classifiers to improve predictive accuracy and robustness.

II.MACHINE LEARNING ALGORITHMS CLASSIFICATION

- ***ADABOOST:*** Ada boost is an ensemble learning technique that combines multiple weak classifiers to create a strong classifier. It assigns higher weights to misclassified instances and improves classification performance iteratively. Reduces bias and variance, leading to high accuracy. Works well on structured medical data. Effective even with small datasets.

- ***EXTRA TREES CLASSIFIER*:** Extra Trees is an ensemble learning method that creates multiple decision trees by randomly selecting features and thresholds for splits. Unlike Random Forest, it does not calculate the optimal split, making it computationally efficient. Reduces over fitting compared to traditional decision trees. Works efficiently on large datasets. Performs well in high-dimensional data.

- ***CONVOLUTIONAL NEURAL NETWORK*:** CNN is a deep learning algorithm primarily used for image and feature-based classification tasks. It is applied in this project to extract complex patterns from structured medical data. Captures complex relationships between risk factors. Highly accurate due to deep learning capabilities. Effective in large datasets with multiple features.

- ***ENSEMBLE MODEL:*** The algorithms Ada Boost, Extra Trees, and CNN are combined to form an ensemble classifier. By leveraging the strength of each model, the ensemble approach aims to improve overall accuracy and robustness in heart disease prediction.

III. DATA SET AND ATTRIBUTES

The data set used in this study contains multiple patient records with *14 key attributes* that help in predicting heart disease. These attributes include *demographic factors, clinical measurements, and test results.* The data set consists of features such as *age, sex, chest pain type (cp), resting blood pressure (trestbps), serum cholesterol level (chol), fasting blood sugar (fbs), resting electrocardiographic results (restecg), maximum heart rate achieved (thalach), exercise-induced angina (exang), ST depression induced by exercise (old peak), slope of the peak exercise ST segment (slope), number of major vessels colored by fluoroscopy (ca), and thalassemia type (thal).* The target variable, *"target,"* indicates the presence (1) or absence (0) of heart disease. The data set is preprocessed to handle missing values, normalize numerical features, and encode categorical variables. For model evaluation, the data set is split into *75% training data and 25% testing data*, ensuring robust performance assessment of machine learning models.

75%X_Train (Trainning Data)	75%Y_Train (Trainning Data)
25%X_Test (Testing Data)	25%Y_Test (Testing Data)

***Fig.1*. Diagrammatic representation of Trained and Test data**

IV. RESULTS AND ANALYSIS

The models are trained based on data set and evaluated their accuracy using training and test data. The results of each classifier are summarized below

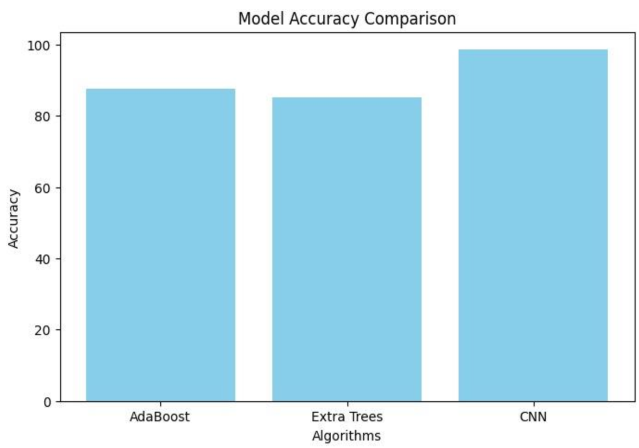

***Fig.2*. Accuracy Comparision**

The data set results are with comparison with the other algorithms. The training data accuracy and testing data accuracy of each algorithm is generated in order to get exact and effective results.

Algorithm	Training Data Accuracy	Test Data Accuracy
AdaBoost	0.92	0.91
Extra Trees	0.96	0.93
CNN	0.98	0.96
Ensemble Model	0.99	0.98

***Fig.3.* Comparison values**

V. CONCLUSION

The Heart Disease Prediction System successfully leverages machine learning and deep learning techniques to predict the likelihood of heart disease in patients. By implementing Ada Boost, Extra Trees, and Convolutional Neural Networks (CNN), the system achieves high accuracy and reliability, providing a valuable tool for early detection and diagnosis.

The study demonstrates that machine learning-based models outperform traditional diagnostic methods by processing large datasets efficiently, identifying critical risk factors, and providing real-time predictions. The integration of feature selection, data preprocessing, and ensemble learning techniques enhances the model's precision and robustness.

Additionally, the system is designed with a user-friendly interface, scalability, security, and performance optimization, making it adaptable for real-world medical applications. The feasibility analysis confirms that the system is technically viable, economically feasible, and socially acceptable, ensuring its practical implementation in healthcare environments.

VI. FUTURE SCOPE

The Heart Disease Prediction System has shown promising results in accurately predicting heart disease risk using machine learning and deep learning techniques, but there are several areas for future improvement and expansion. One potential enhancement is the integration of real-time data from wearable devices, such as smartwatches and ECG monitors, to continuously assess heart health and provide instant alerts. Additionally, exploring more sophisticated deep learning architectures, such as transformers, auto-encoders, or hybrid models, could further improve predictive accuracy and interpretability. Implementing explainable AI (XAI) techniques would also help clinicians better understand model decisions, fostering greater trust and transparency in automated predictions. To ensure scalability and accessibility, deploying the system on cloud platforms like AWS, Google Cloud, or Azure is essential, along with developing user-friendly mobile or web-based applications for patients and healthcare providers worldwide. Incorporating genetic factors, lifestyle habits, and family medical history would allow for personalized risk assessments, leading to better disease prevention. Moreover, integrating the system with electronic health records (EHRs) would streamline data analysis, aiding doctors in making informed, data-driven decisions. Finally, collaborating with hospitals and research institutions could facilitate clinical validation and real-world testing, improving the system's credibility and effectiveness.

REFERENCES

[1] Jayaprakash, S., Nagarajan, M.D., Prado, R.P.D., Subramanian, S. and Divakara Chari, P.B., 2021. A systematic review of energy management strategies for resource allocation in the cloud: Clustering, optimization and machine learning. Energies, 14(17), p.5322.

[2] Mr. Santhana Krishnan.J, Dr. Geetha.S," Prediction of Heart Disease Using Machine Learning Algorithms",2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT), doi:10.1109/ICIICT1.2019.8741465.

[3] Avinash Golande, Pavan Kumar T," Heart Disease Prediction Using Effective Machine Learning Techniques", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277- 3878, Volume-8, Issue-1S4, June 2019.

[4] V.V. Ramalingam, Ayantan Dandapath, M Karthik Raja," Heart disease prediction using machine learning techniques: a survey", International Journal of Engineering & Technology, 7 (2.8) (2018) 684-687.

[5] Hazra, A., Mandal, S., Gupta, A. and Mukherjee, "A Heart Disease Diagnosis and Prediction Using Machine Learning and Data Mining Techniques: A Review" Advances in Computational Sciences and Technology, 2017.

[6] Soni J, Ansari U, Sharma D & Soni S (2011). Predictive data mining for medical diagnosis: an overview of heart disease prediction. International Journal of Computer Applications, 17(8), 43-8.

[7] Dangare C S & Apte S S (2012). Improved study of heart disease prediction system using data mining classification techniques. International Journal of Computer Applications, 47(10), 44-8.

[8] Ordonez C (2006). Association rule discovery with the train and test approach for heart disease prediction. IEEE Transactions on Information Technology in Biomedicine, 10(2), 334-43.

[9] Dangare Chaitrali S and Sulabha S Apte. "Improved study of heart disease prediction system using data mining classification techniques." International Journal of Computer Applications 47.10 (2012): 44-8.

[10] Shinde R, Arjun S, Patil P & Waghmare J (2015). An intelligent heart disease prediction system using k-means clustering and Naive Bayes algorithm. International Journal of Computer Science and Information Technologies, 6(1), 637-9.