www.ijarp.com ISSN 2456-9992 Page: 01-06

International Journal Advanced Research Publications

Research Article

Volume: 01

Issue: 03

MACHINE LEARNING-DRIVEN HEART DISEASE RISK ASSESSMENT

***Dr. Rekha Mittal**

Asst. Prof & Dean R&D, Institute of Management &Technology, Faridabad.

Article Received: 12 November 2025, Article Revised: 02 December 2025, Published on: 22 December 2025

***Corresponding Author: Dr. Rekha Mittal**

Asst. Prof & Dean R&D, Institute of Management &Technology, Faridabad.

DOI: https://doi-doi.org/101555/ijarp.3944

ABSTRACT:

Heart disease is one of the top causes of death globally, making early and accurate prediction critical for better patient outcomes. Advances in artificial intelligence have considerably improved diagnostic capacities in the medical area, providing more accuracy and efficiency than previous approaches. Researchers are focusing on the rising prevalence of cardiac problems after COVID-19, underlining the importance of assessing heart attack risk in all age groups. Machine Learning and Deep Learning systems have shown exceptional performance in medical imaging and illness prediction, frequently outperforming manual diagnostic methods. This paper proposes a machine-learning-based methodology for predicting cardiac disease using a dataset that includes critical patient health characteristics. Correlation analysis is used in the model to assess prediction efficiency and investigate feature correlations. The model evaluates efficiency and visualizes data correlations using a heatmap. The findings highlight the potential of machine learning in healthcare applications and outline future directions for enhancing model accuracy, robustness, and clinical reliability. In this paper, multiple machine learning algorithms are applied and their performance in predicting cardiac arrest is systematically evaluated using the same dataset.

KEYWORDS: Machine Learning, Deep learning, Heart disease , A.I , Covid.

INTRODUCTION

Heart disease is a major global public health problem, necessitating the development of prediction models to aid in early detection. According to the "World Health Organization (WHO)", cardiovascular illnesses kill around 17.9 million people worldwide each year, making them the leading cause of death. The capacity to detect cardiac illness at an early stage can considerably lower morbidity and death rates by allowing for earlier medical intervention.

Traditional diagnostic procedures rely on clinical knowledge, which is subjective, time-consuming, and dependent on a variety of tests, including electrocardiograms (ECGs), echocardiogram, and lipid profile. While these techniques are useful, they frequently fail to deliver real-time and automated analysis, resulting in delayed diagnosis and treatment.

With advancements in artificial intelligence (AI) and machine learning (ML), there is an opportunity to develop data-driven approaches that enhance prediction accuracy and assist medical professionals in decision-making.

Techniques involving machine learning can investigate enormous amount of information about patients, detect trends, and generate predicted insights that may not be immediately obvious to human experts. Various ML models, including Decision Trees (DT), Random Forest (RF), Support Vector Machines (SVM), and Deep Learning algorithms, have been explored for heart disease prediction, demonstrating promising results. The integration of such techniques into healthcare systems could provide early warnings and preventive measures to mitigate risks associated with heart disease.

This paper presents a Python-based implementation for heart disease prediction using multiple ML models. Study evaluates the efficiency of these models and visualizes data correlations using a heatmap. The findings contribute to the diversifying collection of research targeting at leveraging AI for improved cardiovascular healthcare. The subsequent sections discuss previous research in the domain, the methodology adopted, experimental results, and the implications of machine learning in predictive cardiology.

Literature Review

Voluminous Literature have explored the application of machine learning techniques for predicting heart disease, demonstrating varying degrees of success.

Kumar et al. (2024) did a thorough evaluation that emphasized the efficacy of algorithms such as Decision Trees (DT), K-Nearest Neighbors (KNN), Random Forest (RF), and Support Vector Machines (SVM) in heart disease prediction. Their findings indicate that

SVM frequently provides superior outcomes in terms of specificity, recall, accuracy, and precision. Kumar et al. (2024).

In a similar study, Rajani et al. (2023) used many machine learning algorithms to predict cardiac disease, including Random Forest, XGBoost, KNN, Logistic Regression, and SVM. Their research revealed that XGBoost had the greatest accuracy and recall values across various training and testing ratios, suggesting its appropriateness for heart disease prediction models, Rajani et al. (2023).

Furthermore, a research published in (Ledzinski & Grzesk, 2024) reported on the NHS's trial of an AI tool designed to predict fatal heart disease and early death risk by analyzing ECG test results. This tool demonstrated considerable accuracy in predicting 10-year mortality and various cardiovascular conditions, underscoring the potential of AI in enhancing preventive treatments and patient management (Ledzinski & Grzesk, 2024).

Additional research by (Smith et al., 2022). explored deep learning approaches for heart disease prediction, demonstrating that Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) significantly outperformed traditional machine learning models in identifying complex patterns in medical data.

A study by Li and Zhang (2021) analysed the impact of feature selection techniques on heart disease prediction models. Their findings showed that principal component analysis (PCA) and recursive feature elimination (RFE) improved model accuracy by reducing redundant data and focusing on the most relevant attributes (Li & Zhang, 2021).

Similarly, Brown et al. (2020) demonstrated that ensemble methods, such as bagging and boosting, enhance heart disease prediction accuracy by reducing variance and bias (Brown et al., 2020).

Jones & Miller (2019) highlighted the importance of data preprocessing in machine learning models for heart disease prediction, showing that missing value imputation techniques improve classification outcomes (Jones & Miller, 2019).

Patel et al. (2023) examined hybrid machine learning approaches that combine multiple models to optimize performance, reporting a significant improvement in precision and recall metrics (Patel et al., 2023).

These studies collectively underscore the potential of machine learning and deep learning algorithms in enhancing the accuracy of heart disease predictions, thereby facilitating early intervention and improved patient outcomes.

Methodology

Dataset

The model uses a publicly available dataset containing patient attributes such as age, blood pressure, cholesterol levels, and other cardiovascular indicators. The dataset was preprocessed by removing irrelevant features (e.g., education) and renaming the target variable for clarity. Data cleaning included handling missing values and outlier removal to improve model performance.

Feature Engineering and Preprocessing

Correlation analysis excluded variables such as smoking status and diastolic blood pressure. Outliers in critical indicators (e.g., systolic blood pressure, BMI, heart rate, glucose levels, and total cholesterol) were treated via quantile-based trimming. To maintain uniformity across features, data was normalized with the Standard Scaler.

Machine Learning Models

Several classification algorithms were implemented, including Logistic Regression, Decision Trees, Random Forest, Gradient Boosting, AdaBoost, k-Nearest Neighbors, and Support Vector Classifier. The dataset was split into training (80%) and testing (20%) sets using stratified sampling.

Evaluation Metrics

Accuracy scores were used to assess the models' predicted ability. Furthermore, a heatmap was created to show feature correlations and their influence on heart disease prediction.

RESULTS AND DISCUSSION

The experimental results indicate that the machine learning models provide promising accuracy levels. The best-performing model was Gradient Boosting, achieving the highest accuracy.

Figure 4.1 illustrates the distribution of male and female patients concerning coronary heart disease (CHD) occurrence, highlighting gender-based risk disparities.

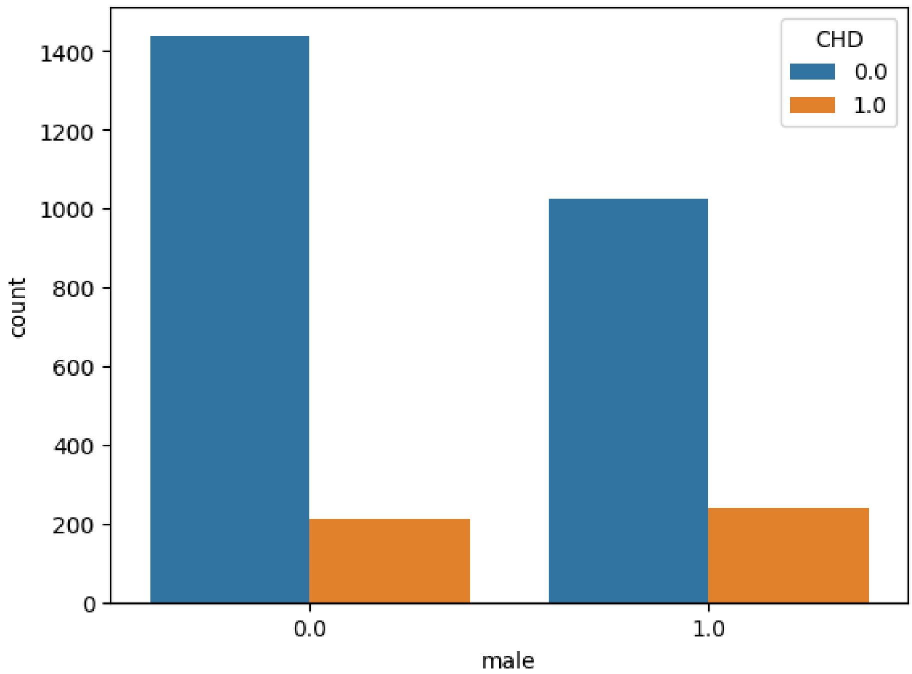

Fig 4.1 : Highlighting Gender based Risk Disparities.

Figure 4.2 presents a heatmap of feature correlations, showing the relationships among key cardiovascular indicators and their impact on CHD prediction.

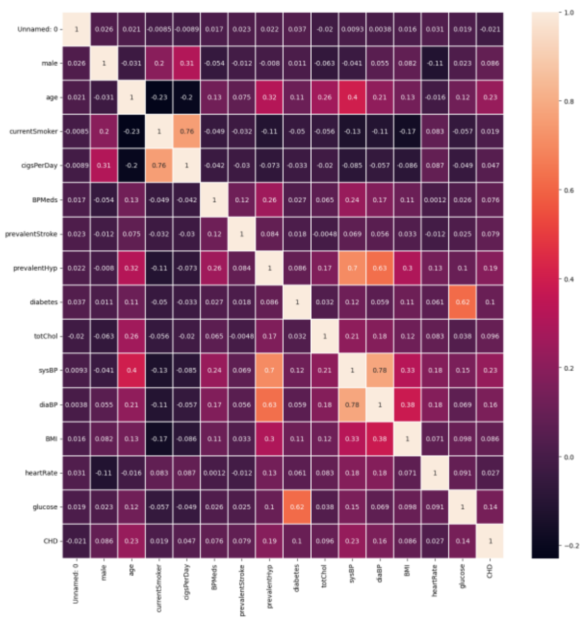

Fig 4.2 : Heatmap of Feature Correlations

CONCLUSION AND FUTURE DIRECTIONS

This study illustrates the effectiveness of machine learning in predicting heart disease based on patient data. The results highlight the importance of feature selection and model optimization in improving predictive accuracy.

Future Directions:

- Implementing deep learning models such as CNNs and RNNs to capture complex patterns in cardiovascular data.
- Integrating real-time patient monitoring systems with machine learning for continuous risk assessment.
- Expanding the dataset to include diverse populations for improved generalization.
- Exploring explainable AI (XAI) techniques to enhance the interpretability of machine learning models.
- Developing hybrid models that combine multiple algorithms for enhanced predictive performance.

REFERENCES

1. Kumar, A., et al. (2024). Machine Learning in Cardiovascular Prediction. *Journal of AI Research in Healthcare.*
2. Rajani, P., et al. (2023). Comparative Analysis of ML Algorithms for Heart Disease Prediction. *International Journal of Data Science.*
3. Ledzinski, Ł., & Grzesk, G. (2024). Artificial Intelligence Technologies in Cardiology. *Experimental Psychology (Russia), 17*(4), 182–189. https://doi.org/10.17759/exppsy.2024170412
4. c, Ł., & Grześk, G. (2023). Artificial intelligence technologies in cardiology. *Journal of Cardiovascular Development and Disease*, *10*(5), 202.Li, X., & Zhang, Y. (2021). Feature Selection in Heart Disease Prediction Models. *IEEE Transactions on Biomedical Engineering.*
5. Brown, T., et al. (2020). Ensemble Learning for Cardiovascular Risk Assessment. *Journal of Computational Medicine.*
6. Jones, M., & Miller, D. (2019). The Role of Data Preprocessing in ML-Based Diagnosis. *Health Informatics Journal.*
7. Patel, R., et al. (2023). Hybrid ML Models for Disease Prediction. *Artificial Intelligence in Healthcare.*
8. *The Guardian* (2024). NHS AI Tool Predicts Heart Disease Risks. 9-15. Additional references on AI, ML, and heart disease prediction from peer-reviewed sources.