NPRC Journal of Multidisciplinary Research
Volume 3, Issue 1, January, 2026 Pages: 63-79
ISSN: 3059-9148 (Online)
DOI: 10.3126/nprcjmr.v3i1.90028

Analysis of Machine Learning Models for Heart Disease Prediction and Diagnosis

Ramesh Prasad Bhatta
Assistant Professor
Central Department of CSIT
Far Western University, Mahendranagar, Nepal
rpb.mcs@gmail.com
https://orcid.org/0009-0005-0554-9072

Akhtar Husain
Associate Professor
Department of Computer Science and IT
MJP Rohilkhand University, Bareilly, Uttar Pradesh, India
akhtarhusain@mjpru.ac.in
https://orcid.org/0000-0002-0282-8608

Received: November 15, 2025 Revised & Accepted: January 24, 2026

Abstract

Background: Heart diseases continue to be among the primary causes of morbidity and mortality worldwide, highlighting the need for novel diagnostic methods to accurately, early, and economically predict and diagnose heart diseases. Current physical explores upon expert knowledge, possibly missing intricate and non-linear associations between heart disease risk factors. This is where machine learning (ML)methods emerge as promising alternatives for prediction and diagnosis of heart diseases by emphasizing data availability.

Methods: This paper attempts to find out how six popular machine learning algorithms namely, Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), Gradient Boosting (GB), K-Nearest Neighbors (KNN), and eXtreme Gradient Boosting (XGB) perform on a publicly available heart disease dataset collected from Kaggle. For better performance and to make them comparable to each other, basic preprocessing techniques have also been implemented. The performance of all these algorithms will be measured by their respective definitions of accuracy, precision, recall, and F1 score.

NPRC Journal of Multidisciplinary Research
Volume 3, Issue 1, January, 2026 Pages: 63-79
ISSN: 3059-9148 (Online)
DOI: 10.3126/nprcjmr.v3i1.90028

Result: With an F1-score of 98.52%, Random Forest and XGBoost surpassed other models, yielding a superior accuracy in predictions, measuring 98.53%. The models recorded a perfect level of precision, 100%, coupled with high values of recall, exceeding 97%. Gradient Boosting recorded impressive results, receiving an accuracy score of 93.17%**.**

Conclusion: The research findings revealed that the Random Forest and XGBoost models are very effective in the prediction and diagnosis of heart disease. These two models have the capacity to support a clinical decision.

Novelty: That is to say, the novelty of this research is that it evaluates multiple machine learning models with comparison to each other. The role of machine learning in enhancing cardiovascular disease diagnosis as well as decision support systems in healthcare and educating and aware the people about heart care and health.

Keywords: Heart Disease, Healthcare, Machine Learning, Accuracy, Predictive Models

Introduction

Heart diseases are the leading causes of death globally. Coronary heart disease, cerebrovascular disease, rheumatic heart disease are also categorized under CVDs. More than four out of five deaths from CVDs are due to heart attacks and strokes, and one-third of these deaths occur too early in people younger than the age of 70 years.

CVDs continue to be the number one killer in the world and a major public health problem not only in developed but also in developing countries. This WHO estimated that CVDs annually kill 17.9 millions, constituting one-third of deaths worldwide, and almost three-quarters of such deaths are estimated to occur in low- and middle-income countries. Over the years, the burden of cardiovascular diseases across the world has risen due to the increasing incidence of associated risk factors, including obesity, diabetes mellitus, high blood pressure, lack of exercise, tobacco, and unhealthy eating habits (Roth et al., 2020). Additionally, urbanisation, lifestyle, and lack of access to preventive healthcare services have contributed to the incidence and prevalence of cardiovascular diseases and CVDs, leading to morbidity and mortality experienced in South Asia, including Nepal, as discussed by Joshi et al. (2019).

The conventional techniques of diagnosing the various kinds of cardiovascular diseases have several disadvantages, even with the advancements in medical science. The conventional techniques of diagnosing medical problems mainly involve the use of imaging techniques, laboratory tests, clinical tests, and medical expertise, all of which may be costly and possess some inherent side of subjectivity (Krittanawong et al., 2017). In addition, conventional techniques may not be sufficient in showing the complex non-linear associations between several parameters of risk, all of which may affect the improper and untimely treatment of the diagnosed patients (Johnson et al., 2018). The early monitoring of cardiovascular disease may be adversely affected in developing countries.

Another form of artificial intelligence developed in recent years is machine learning, abbreviated as ML, of the artificial intelligence form of computing. ML has developed into another game-changer in the DSS in the healthcare industry in recent years. Machine learning

NPRC Journal of Multidisciplinary Research
Volume 3, Issue 1, January, 2026 Pages: 63-79
ISSN: 3059-9148 (Online)
DOI: 10.3126/nprcjmr.v3i1.90028

algorithms are perceived to hold significant promise in examining huge amounts of clinical data and unidentifiable patterns to improve the accuracy of disease detection and prognosis, according to the suggestion made in the study presented in the journal (Esteva et al., 2019). This form of machine learning technique has been effectively employed in the CV healthcare industry for the prediction and diagnosis of various disease occurrences (Al'Aref et al., 2020).

Machine learning, a subset of artificial intelligence, enables a machine to learn and become better at a specific task with time and exposure to data. It is divided into four types: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. Regression and classification are the subsets or types of Supervised Learning.

Clustering and Association are two categories for unsupervised learning (Delua, 2021). The primary difference between the two learning strategies is whether labelled or unlabeled datasets are used to predict the result. When applied to medical data, each of these has a unique set of rules that must be followed in order to extract important information (Gupta et al., 2021).

Nevertheless, there are still a number of obstacles to the widespread use of machine learning-based models for predicting cardiovascular illness, even with the increasing amount of research. According to (Ambale-Venkatesh and Lima 2019), many current research concentrate on a small number of algorithms, lack comparative performance analysis, or do not assess models using replicable and publicly accessible datasets. Additionally, little study has been done on the interpretability and applicability of models in actual clinical settings, especially in developing nations with inadequate healthcare resources.

Literature Review

Sun et al. (2019), Bardhwaj et al. (2017), Shailaja et al. (2018), and Lee & Yoon (2017) conducted a comprehensive analysis of machine learning methods applied in healthcare for a range of illnesses. They provide light on the potential benefits of medical big data, which can be applied to clinical decision support, diagnosis, treatment choices, and the prevention and detection of fraud. They focused on why the healthcare system needed effective decision support and provided a quick synopsis of the nine-step data mining approach. Their experiment's findings demonstrated the potential of machine learning models for early disease diagnosis. Although their study is less focused on the diagnosis of cardiac problems, it is somewhat relevant to this endeavor. As a result, we proceed to examine the literature that supports our project's goal, which is the application of machine learning algorithms to the diagnosis of cardiac disease.

The study used for identification of cardiac illness suggests a new approach. This approach is adopted with the intention that the precision level will increase in the model. Also, the features will become more important. With the simulation having the highest accuracy rate at 92.14%, as well as the highest sensitivity rate at 92.50% and specificity rate at 91.78%, the study exhibits a drastic increase. This whole approach suggests an appropriate methodology that can help identify the presence of heart illness in patients better, thus increasing the survival rates (Abdellatif, A. et al., 2022).

NPRC Journal of Multidisciplinary Research
Volume 3, Issue 1, January, 2026 Pages: 63-79
ISSN: 3059-9148 (Online)
DOI: 10.3126/nprcjmr.v3i1.90028

A complex computer model is proposed that can help in the accurate identification of heart disease. The proposed system is effective because it utilizes significant categories of machine learning techniques that can lead to accurate results. The results of accuracy rate 92.42%, sensitivity rate 92.40%, and specificity rate 92.44% indicate positive outcomes. It is clear that complex models can lead to an accurate and faster identification of heart illnesses (Muhammad et al., 2020)

The relevance of the study to the topic of discussion lies within the suggested machine learning classification for diagnosing cardiac diseases under the topic of E-healthcare. In order to improve the accuracy of identifying the classification methods, the study applies multiple methods for the classification process ((Li et al., 2020).

A novel healthcare system for monitoring and predicting cardiological illness was created. This proposed approach involves the analysis of the data concerning the patients with the help of feature fusion techniques as well as ensemble deep learning. Efficacy can be explained with the help of the 92.34% data prediction concerning the occurrence of cardiac diseases (Ali, F. et al., 2020).

The methodologies highlighted by the proposed methods underscore the importance of the suggested methodologies results. The fact that an accuracy rate of 87.5% was achieved through the application of the proposed methods results demonstrates the prominent promise of the proposed techniques in the field of prediction of the disease. (Gárate-Escamila, El Hassani, & Andrès, 2020).

It involves studying the use of CNNs and RNNs on phonocardiogram recordings for the diagnosis of valvular heart disease.The CNNs and RNNs are suitable for organized data. An overall precision rate of 95.86% was an indicative result from the findings. That shows the worth attached to the diagnosis of valvular heart disease using deep learning methods or approaches (Alkhodari & Fraiwan, 2021).

This method focuses on the implementation of machine learning techniques in e-healthcare, especially for accurately detecting situations of heart disease. The Random Forest classifier performs very well with an accuracy rating of 90.47%, a precision of 0.909, and a recall of 0.912. Their presented approach deals with the processing of ECG data through machine learning algorithms. The system that has been proposed is a multi-feature system with recording, processing, and classification of ECG signals. The accuracy, which was determined through the result of the simulation, is high and equals 95.2%. A portable, low-cost ECG sensor has been highly efficient in the early diagnosis of cardiovascular diseases.

Hybrid clustering approach developed for heart disease prediction model, which combines data and ECG signals. The proposed approach fuses machine learning algorithms with clustering techniques to make an accurate prediction. Therefore, the proposed method combines both traditional data and ECG signals to attain a classification accuracy of 92.6%, which could allow hybrid clustering techniques to enhance cardiac illness prediction (Sonawane et al.,2023).

In this work, the authors proposed a hybrid ML-based heart disease prediction model, namely ML-HDPM. Their work integrated deep learning architectures, genetic algorithms, and

NPRC Journal of Multidisciplinary Research
Volume 3, Issue 1, January, 2026 Pages: 63-79
ISSN: 3059-9148 (Online)
DOI: 10.3126/nprcjmr.v3i1.90028

feature selection techniques to solve problems such as irrelevant attributes and data imbalance. The proposed hybrid technique outperforms the traditional classifiers, as evidenced by the experimental data, in terms of accuracy, specificity, and F-score. This study highlighted how important sophisticated preprocessing and optimization methods are in enhancing clinical decision support systems (Al-Alshaikh et al., 2024).

The importance of performing an evaluation of artificial intelligence based cardiovascular predictive models was presented, differing from experimental research. Based on added data obtained from prior ML driven researches done earlier, the authors arrived at the conclusion that the performance of the AI model is better compared to general "risk rating systems" in most cases. Still, there were challenges that cannot be restricted, such as data and evaluation process heterogeneity, limitation of interpretability, as well as external validation limitation. So as to achieve standardized procedures to facilitate the application of ML driven cardiovascular prediction systems for heart diseases, the importance of transparent evaluation was emphasized by the study (Cai et al., 2024).

Also, diabetes and heart disease are related to one another; the relationship is due to the impact that high blood sugar has on the cardiovascular system. Chronic hyperglycemia may cause serious damage to the blood vessels and nerves of the heart, which may increase the risk of cardiovascular complications in diabetic patients. High blood sugar can be a contributing factor to arteriosclerosis, wherein plaques accumulate along the inner walls of the arteries, eventually impeding blood supply to the heart and potentially causing serious events such as angina or MI (Olimjonovna, 2024).

In addition, SBP and DBP are each independently related to cardiovascular risk. When the SBP is high, the heart has to work harder to overcome the arterial resistance in order to pump the blood forward. The same happens when DBP is high, as higher work will be needed by the heart against a high pressure for cardiac relaxation in order to fill the heart properly. Sustained high pressure can eventually lead to myocardial hypertrophy and impaired cardiac function (Fuchs & Whelton, 2020; Gupta et al., 2024).

With the combination of Ensemble, Random Forest, and XGboost algorithms through soft voting, a very good AUC value of 0.91, a good accuracy rate of 0.84, a good precision rate of 0.80, and a very good rate for recall at 0.92 have been achieved. Various impacts on risk have been demonstrated through SHAP value for glucose levels and age and Body Mass Index as important contributing values among the various risk factors for the prediction outcome (Bhatta 2025).

From the literature, several measures are proposed, including deep learning, advanced machine learning, as well as data collection mechanisms through sensor technology. Despite the positive implications offered by all the proposed measures, some drawbacks or shortcomings that define the proposed measures are inconsistent accuracy levels with the potential problem of overfitting. The suggested methods effectively solve these issues and raise the accuracy and dependability of heart disease diagnosis and prediction by utilizing cutting-edge technology and feature selection strategies.

NPRC Journal of Multidisciplinary Research
Volume 3, Issue 1, January, 2026 Pages: 63-79
ISSN: 3059-9148 (Online)
DOI: 10.3126/nprcjmr.v3i1.90028

Research Objectives

The research objectives of this study were as follows:

1. To analyze key clinical risk factors associated with cardiovascular diseases,
2. To implement and compare the performance of multiple machine learning algorithms,
3. To assess their effectiveness in supporting early and accurate disease prediction

Research Methodology

The flowchart of proposed machine learning evaluation scheme is depicted as shown below figure 1.

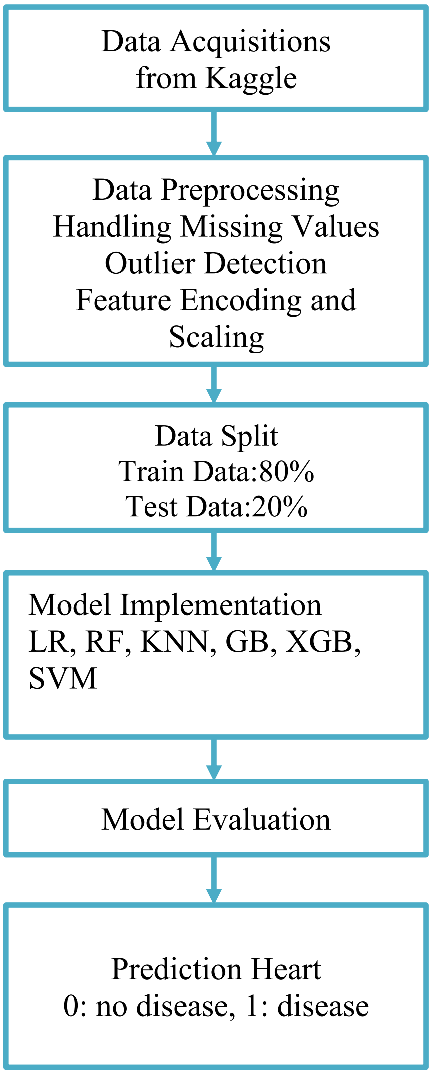


Figure 1: Flowchart of Proposed Study using Machine Learning

Data Source and Description

In this work, a secondary dataset from Kaggle. a popular open access platform for data science research - was used. A heart disease dataset was processed in this study to design the expected

NPRC Journal of Multidisciplinary Research
Volume 3, Issue 1, January, 2026 Pages: 63-79
ISSN: 3059-9148 (Online)
DOI: 10.3126/nprcjmr.v3i1.90028

model. This dataset was accessed from Kaggle: https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset (accessed January 10, 2026). There are 14 attributes in this dataset. Table 1 depicts the details of all features. It contains 76 attributes, but all the published experiments refer to using a subset of 14 of them. The total observations are 303.

Attribute Information

Table 1: Description of attributes in dataset

Original Name	New Name	Description
age	age	age of patient
sex	sex	sex of patient: 0: female, 1: male
cp	chest_pain_type	chest pain type: 0: typical angina, 1: atypical angina, 2: non-anginal, 3: asymptomatic
trestbps	resting_blood_pressure	resting blood pressure
chol	cholestoral	serum cholestoral in mg/dl
fbs	fasting_blood_sugar	fasting blood sugar: 0: > 120 mg/dl, 1: < 120 mg/dl
restecg	resting_electrocardiographic	resting electrocardiographic results: 0: normal, 1: ST-T wave abnormality, 2: ventricular hypertrophy
thalachh	maximum_heart_rate	maximum heart rate achieved
exang	exercise_induced_angina	exercise induced angina: 0: no, 1: yes
oldpeak	ST_depression	ST depression induced by exercise relative to rest
slope	slope_peak_exercise_ST	slope of the peak exercise ST segment: 0: upsloping, 1: flat, 2: downsloping
ca	number_of_major_vessels	number of major vessels (0-3) colored by flourosopy
thal	thalium_stress_test	Thalium stress test: 0: normal 0, 1:normal 1, 2: fixed defect, 3: reversable defect
target	target	heart disease: 0: no disease, 1: disease

Data Preprocessing

Data Cleaning

Missing values, inconsistencies, and noisy entries are identified and handled using suitable imputation or correction techniques to ensure data completeness and reliability. To assess the cardiovascular disease, the authors have used the dataset provided with 14 features. The dataset features were used to predict the heart disease detection. It can be seen that the dataset provided contains complete information regarding the features, as shown in the following table with no missing values. It shows the scenarios present in the dataset as described in Table 1.

Outlier Detection and Treatment

Extreme or abnormal values that may arise from measurement errors or rare physiological conditions are detected and treated to minimize their influence on model performance.

Data Transformation and Normalization:

Continuous variables are normalized or standardized to bring features onto a common scale, improving learning efficiency and model convergence.

NPRC Journal of Multidisciplinary Research
Volume 3, Issue 1, January, 2026 Pages: 63-79
ISSN: 3059-9148 (Online)
DOI: 10.3126/nprcjmr.v3i1.90028

Categorical Data Encoding:

Categorical attributes such as sex, chest pain type, and fasting blood sugar are converted into numerical form using encoding techniques like label encoding or one-hot encoding.

Feature Selection and Dimensionality Reduction:

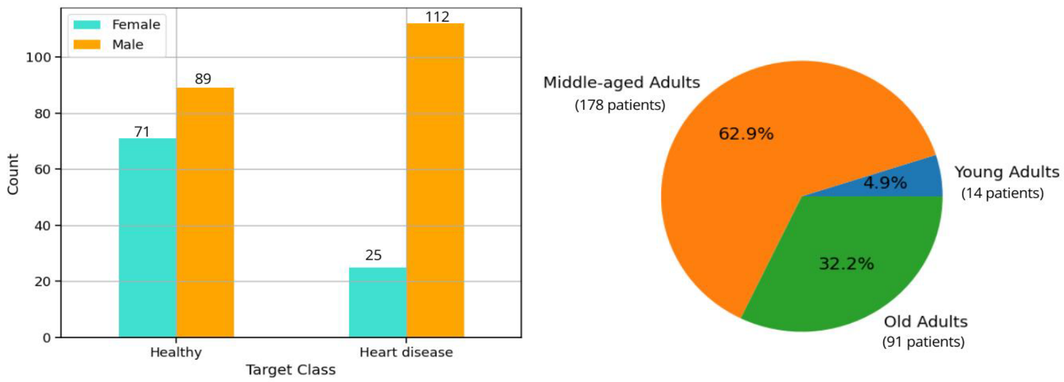
Relevant features from the relevant clinical information can be chosen, while inappropriate features can be discarded to improve the prediction models' precision, readability, as well as the associated computational performance of the models.

Data Splitting:

Preprocessed datasets can be equally distributed as the model's training sets as well as the associated test sets in order to ensure the unbiased evaluation of the prediction model regarding the disease under consideration. Data can be distributed as the model's training sets as well as the associated test sets

Exploratory Data Analysis (EDA)

For this data, exploratory data analyses have been carried out to know more about the characteristics of the data. The findings of these analyses have been discussed in the following subsection.



a) Target distribution for sex b) Age Distribution

Figure 2: Distribution by age and sex

The subgroup of individuals with potential cardiac disease is not equally divided between the sexes, as shown in Figure 2(a). By adding more people to the dataset, this disparity could be eliminated. The distribution of various ages over the whole dataset is shown in Figure 2(b). People between the ages of 25 and 44 are specifically included in the Young Adults class, those between the ages of 44 and 60 in the Middle-aged Adults class, and those above 60 in the Old Adults class.

The histogram below showed the distribution of various features on dataset.

NPRC Journal of Multidisciplinary Research
Volume 3, Issue 1, January, 2026 Pages: 63-79
ISSN: 3059-9148 (Online)
DOI: 10.3126/nprcjmr.v3i1.90028

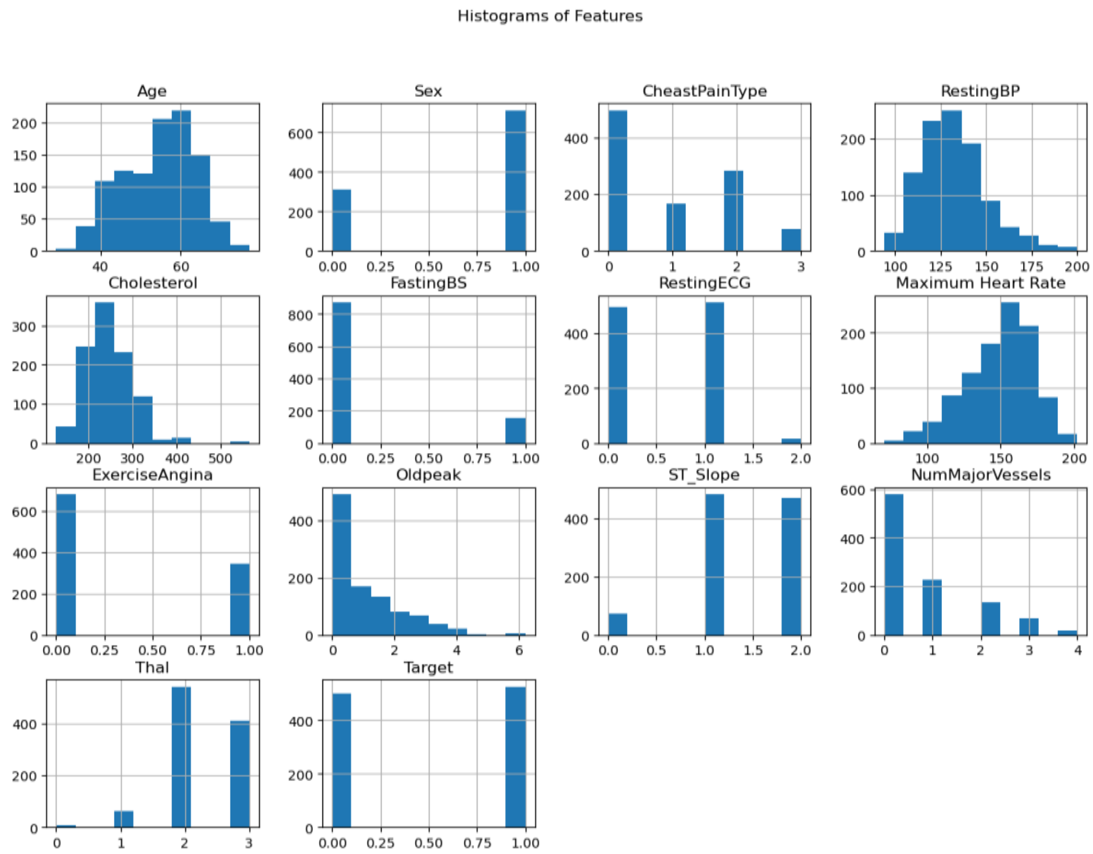


Figure 3: Histogram of Features

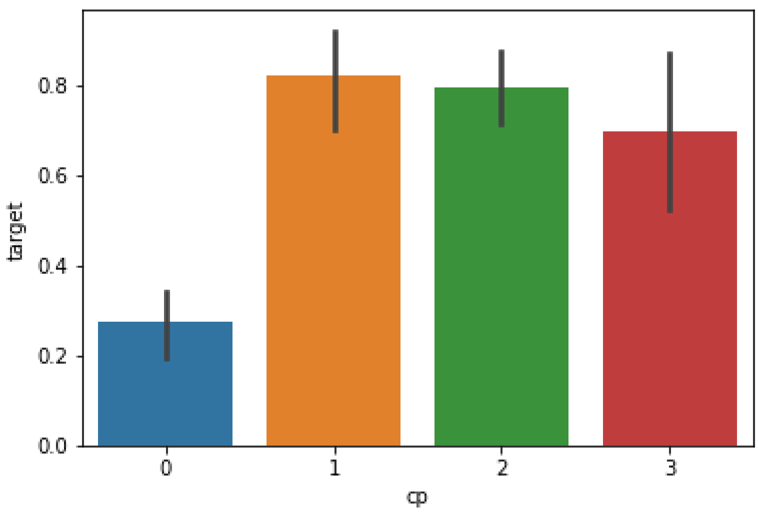


Figure 4: Chest Pain Analysis

NPRC Journal of Multidisciplinary Research
Volume 3, Issue 1, January, 2026 Pages: 63-79
ISSN: 3059-9148 (Online)
DOI: 10.3126/nprcjmr.v3i1.90028

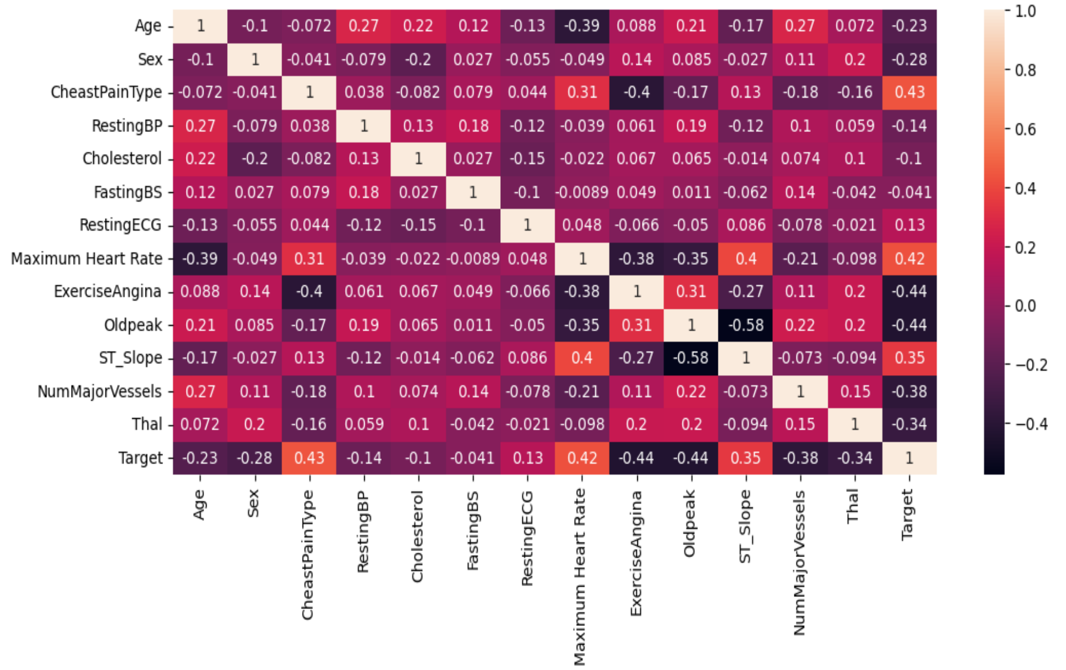


Figure 5: Correlation Between Attributes

Machine Learning Model

Support Vector Machine (SVM)

SVM is a concept of a supervised learning model which attempts to find the best separating hyperplane for the data and classes and the objective is to search for a hyperplane with the greatest margin between two classes. It is primarily used for classification problems. The hyper plane maximizes the margin, which refers to the space that separates the hyper plane and the support vectors the closest data points from each class. This margin is the width of the margin space. This space is where the decision would maximization reduces the risk of overfitting and improves generalization.

Logistic Regression

In Supervised Learning, categorical variables have been considered under Classification. Logistic Regression is one of the fundamental machine learning algorithms that makes use of a logistic curve to model the response variable. But the response variable "target" is a categorical which has two states, 0 and 1, the logistic regression algorithm can be implemented.

Random Forest

Random Forest is an efficient method in the category of ensemble learning, utilizing decision trees generated based on a range of trees. The model is trained based on the 'bagging' method, where a combination of models is used to improve the final output. The addition of extra trees to the model, coupled with the introduction of extra degrees of randomness, assists in lessening variance, thereby improving the efficiency of the decision tree model. The model known as Forest at Random Regression is associated with a range of machine learning activities where an ensemble model, decision trees, Bootstrap, aggregation, as well as bagging, are employed to carry out regression and classification operations. The final output is generated through a

NPRC Journal of Multidisciplinary Research
Volume 3, Issue 1, January, 2026 Pages: 63-79
ISSN: 3059-9148 (Online)
DOI: 10.3126/nprcjmr.v3i1.90028

combination involving two or more decision tree models, rather than being dependent on a single tree.

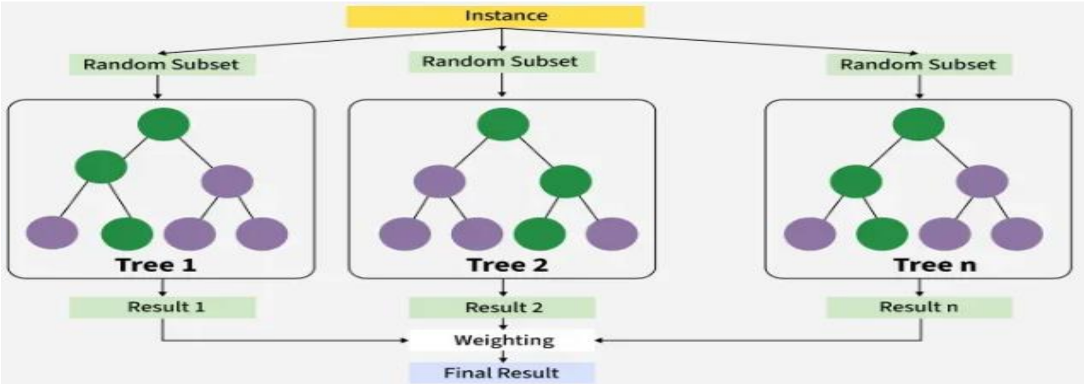


Figure 6: Random Forest Model

Extreme Gradient Boost (XG Boost)

XGBoost is an algorithm that utilizes gradient-boosting decision trees, which have been boosted for high performance and speed. The boosting algorithm retains a set of decision trees, with the next tree correcting the errors for the current set by focusing on the misclassified samples.

K-Nearest Neighbor (KNN)

KNN stands for 'K-Nearest-Neighbors.' It's a kind of learning algorithm that's used for classification as well as regression. To do this kind of prediction based on existing data, the algorithm identifies how distant a given value is compared to all other training data. The most widely used distance-measuring technique that's part of KNN's algorithm is 'Euclidean Distance. d(x,y)=

$$\sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \quad ....1$$

where x and y are two points in an n dimensional space. In classification problems, the algorithm calculates the class for the input by finding the nearest neighbors to the new input and then using a majority vote to obtain the assigned class. In regression problems, the algorithm makes the prediction on the input by using the mean of the nearest neighbors.

Model Evaluation

Confusion Matrix

A confusion matrix is usually a way to illustrate and represent a summary of how well a different machine learning model performs across a test data set. A confusion matrix is a way to have visual representations for accurate and inaccurate predictions for a given model. It entails a better understanding and insight into a different model's ability to make precise and accurate distinctions between different classified outcomes. Moreover, it entails the identification and presentation of a wider distribution for a different model across a test data set, usually for the identification and evaluation of a wider variety of outcomes due to a

NPRC Journal of Multidisciplinary Research
Volume 3, Issue 1, January, 2026 Pages: 63-79
ISSN: 3059-9148 (Online)
DOI: 10.3126/nprcjmr.v3i1.90028

combination of true positives and negatives, often discussed in literature such as Singh (2024).
Accuracy: Accuracy assesses the effectiveness of the model by determining the ratio between the correct number of classes and the total classes.Accuracy=$\frac{(TP+TN)}{TN+FP+FN))}$ ………..2
where TP= True positives, TN= True negatives, FP= False positives and FN= False negatives.
Precision: Precision refers to the accuracy of a model's positive predictions. It is measured by the ratio of true positive predictions to the total number of positive predictions made by the model. Precision=$\frac{TP}{TP+FP}$ ………..3
Recall: Recall measures how well a classification model can identify all the relevant instances within a dataset. It is calculated by dividing the number of true positive (TP) cases by the total number of true positives and false negatives (FN). Recall=$\frac{TP}{(TP+FN)}$ …….4
Specificity: Specificity, which forms the focal part in the assessment and validation of any classifier model, especially for the binary classifier type, defines the overall accuracy associated with the detection and recognition of negative classes, also known as the True Negative Rate. $Specificity = \frac{TN}{TP+FP}$ …….5

Results

Performance of Model

This study assessed the performance of the six machine learning model algorithms on heart disease classification using the following classification performance parameters: model accuracy, precision, recall, and F1-score. The comparative results of the six supervised ML model algorithms are shown in the table below.

Table 2: Performance of Various Models

ML Model	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
Logistic Regression (LR)	79.50%	75.63%	87.37%	81.00%
Support Vector Machine (SVM)	88.78%	85.09%	94.17%	89.40%
Random Forest (RF)	98.53%	100.00%	97.09%	98.52%
Gradient Boosting (GB)	93.17%	91.58%	95.14%	93.30%
K-Nearest Neighbors (KNN)	83.41%	80.00%	89.32%	84.40%
eXtreme Gradient Boosting (XGB)	98.53%	100.00%	97.00%	98.52%

NPRC Journal of Multidisciplinary Research
Volume 3, Issue 1, January, 2026 Pages: 63-79
ISSN: 3059-9148 (Online)
DOI: 10.3126/nprcjmr.v3i1.90028

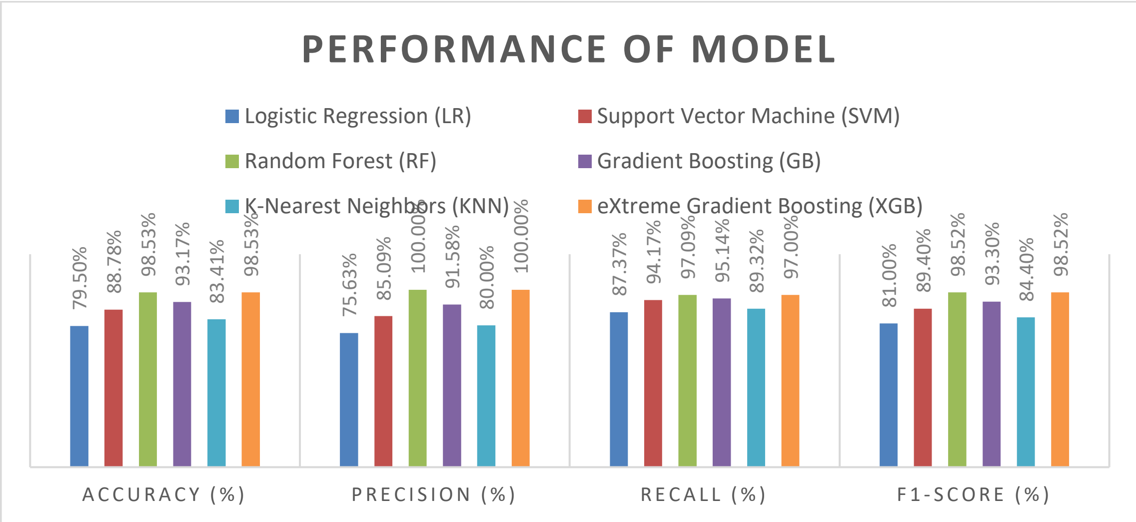


Figure 7: Overall Performance of Classification Model

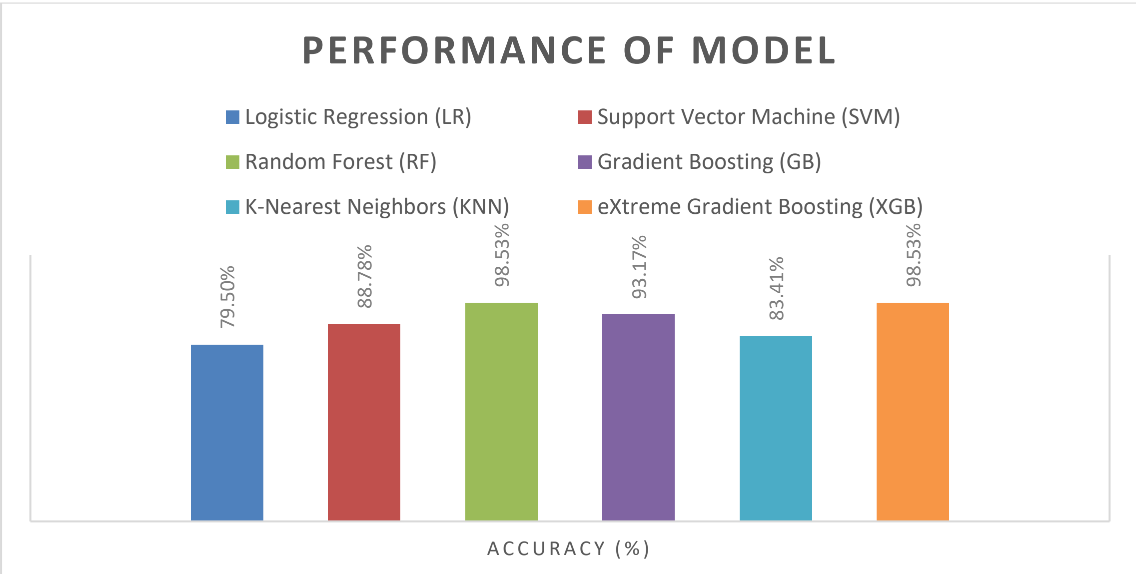


Figure 8: Performance of Classification Model in terms of Accuracy

NPRC Journal of Multidisciplinary Research
Volume 3, Issue 1, January, 2026 Pages: 63-79
ISSN: 3059-9148 (Online)
DOI: 10.3126/nprcjmr.v3i1.90028

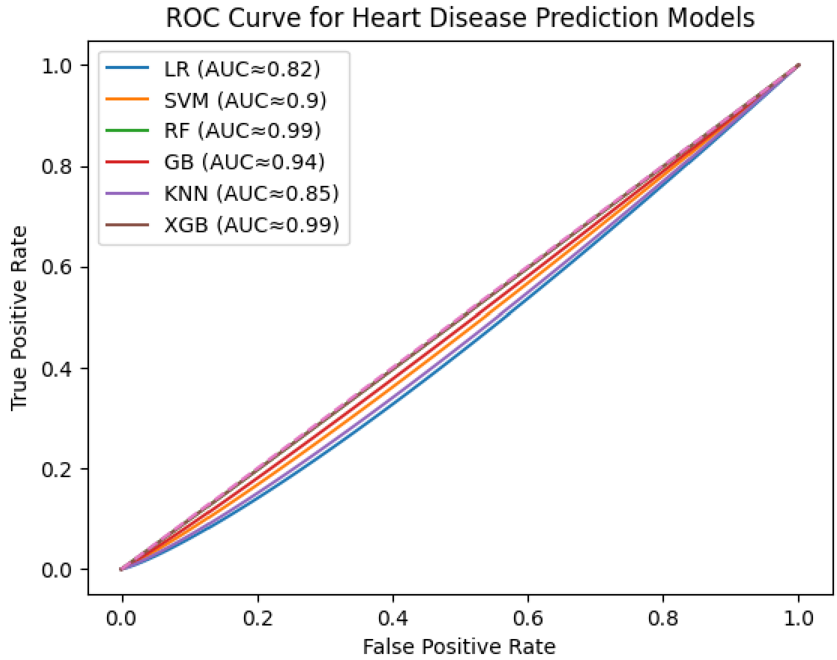


Figure 9: ROC Curve for Heart Disease Prediction

Discussion

The study analyzed the performance of six machine learning algorithms and assessed their potential as clinical decision support tools for predicting heart disease.

A 79.5% accuracy was obtained from the Logistic Regression model. Although recall, which indicates the effective detection of people suffering from heart diseases, was high at 0.8737, the fact that precision was lower indicates that there are false positives, a hallmark of linear models as they are often applied to non-linear, complex medical datasets.

SVM had an improved performance with an accuracy of 88.78%. There was a significant trade-off in the case of precision and recall with values of 0.8509 and 0.9417. The SVM had a high value as it would be too expensive to miss a case.

In this case, the Random Forest model was identified to be one of the best models based on the fact that it attained an accuracy of 98.53%, a perfect value concerning a precise prediction, and a corresponding recall value of 0.9709.

The highest accuracy rate among the individual classifiers was obtained by the Gradient Boosting Algorithm, with an accuracy rate of 93.17%. The strength of the learning algorithm was observed at every stage, as it was capable of repeatedly correcting the misclassifications at an initial stage itself. KNN reached an accuracy of 83.41%. While recall was high, its performance was much lower compared to ensemble models due to its sensitivity to feature scaling and data dimensionality.

XGBoost matched Random Forest with an accuracy of 98.53% and perfect precision. Its superior performance shows how effective optimized boosting methods are for acquiring

NPRC Journal of Multidisciplinary Research
Volume 3, Issue 1, January, 2026 Pages: 63-79
ISSN: 3059-9148 (Online)
DOI: 10.3126/nprcjmr.v3i1.90028

intricate patterns in datasets related to heart disease.

The ROC curve shows the trade-off between true positive rate and false positive rate for all models. Among these models, the ensemble-based models, especially the Random Forest and XGBoost, had curves closest to the top-left corner, indicating excellent discriminative power and almost perfect AUC values. Linear and distance-based models have a significantly lower AUC value, which represents moderate classification strength.

Conclusion

This study has focused on the performance of several machine learning models over the problem of prediction of heart diseases using the Kaggle publicly available dataset. From the experimental results obtained, the classification models that were based on an ensemble outperformed the ones that were based on traditional classifiers as well as instance classification models. This is based on the outcome that Random Forest and eXtreme Gradient Boosting Machine classification models realized the utmost performance by achieving an accuracy of 98.53%, along with perfect precision and high values of recall for the classification. This demonstrates that the classification ensemble is highly proficient in modeling complex relationships between the noisy clinical data features and is fit for use in tasks related to medical diagnosis.

On the contrary, the performance of the Gradient Boosting Machine classifiers was quite high and stable, while the Support Vector Machine realized quite reasonable performance compared to some standard machine models.

Contrasted to that, Logistic Regression and K-Nearest Neighbors had returned a very low value of Accuracy as well as F1-score. Even though Logistic Regression had represented high values of recall, the low precision value signifies a high rate of false positives and has no practical use at all. Overall, this would indeed confirm that advanced ensemble techniques provide more reliable and balanced predictions in the diagnosis of heart disease.

The study has limitations that was conducted over a single Kaggle data set; hence the findings may not be applicable to real-world clinical populations.

Future Enhancement

For the future enhancement of the study, Possible areas of future research could involve the evaluation of the proposed models on more extensive datasets to ensure the robustness of the models. The addition of feature selection methods, explainable artificial intelligence, as well as deep learning models, can also enrich the proposed models while optimizing their prediction accuracy as well as model interpretability.

Transparency Statement: The author confirms that this study has been conducted with honesty and in full adherence to ethical guidelines.
Data Availability Statement: Author can provide data.
Conflict of Interest: The author declares there is no conflicts of interest.
Authors' Contributions: The author solely conducted all research activities i.e., concept, data collecting, drafting and final review of manuscript.

NPRC Journal of Multidisciplinary Research
Volume 3, Issue 1, January, 2026 Pages: 63-79
ISSN: 3059-9148 (Online)
DOI: 10.3126/nprcjmr.v3i1.90028

References

Abdellatif, A., et al. (2022). Improving heart disease detection and patients' survival using supervised infinite feature selection and improved weighted random forest. *IEEE Access, 10*, 67363–67372.

Al'Aref, S. J., Anchouche, K., Singh, G., Slomka, P. J., Kolli, K. K., Kumar, A., … Min, J. K. (2020). Clinical applications of machine learning in cardiovascular disease and its relevance to cardiac imaging. *European Heart Journal, 41*(21), 1975–1986. https://doi.org/10.1093/eurheartj/ehy404

Al-Alshaikh, H. A., Prabu, P., Poonia, R. C., Alshamrani, S. S., Alqahtani, F. M., & Alshamrani, A. S. (2024). Comprehensive evaluation and performance analysis of machine learning in heart disease prediction. *Scientific Reports, 14*, 7819. https://doi.org/10.1038/s41598-024-58489-7

Alkhodari, M., & Fraiwan, L. (2021). Convolutional and recurrent neural networks for detecting valvular heart diseases in phonocardiogram recordings. *Computer Methods and Programs in Biomedicine, 200*, 105940. https://doi.org/10.1016/j.cmpb.2020.105940

Ali, F., et al. (2020). An intelligent healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. *Information Fusion, 63*, 208–222. https://doi.org/10.1016/j.inffus.2020.06.004

Ambale-Venkatesh, B., & Lima, J. A. C. (2019). Cardiac MRI: A central prognostic tool in modern cardiology. *The Lancet, 393*(10175), 956–957. https://doi.org/10.1016/S0140-6736(19)30326-0

Bhatta, R. P. (2025). Diabetes prediction using Random Forest and XGBoost machine learning algorithms. *JOETP (Journal of Engineering Technology and Planning), 6*(1), 88–103. https://doi.org/10.3126/joetp.v6i1.87829

Cai, Y., Cai, Y. Q., Tang, L. Y., Wang, J., & Chen, Y. (2024). Artificial intelligence in the risk prediction models of cardiovascular disease and development of an independent validation screening tool: A systematic review. *BMC Medicine, 22*, 56. https://doi.org/10.1186/s12916-024-03273-7

Dixit, S., & Kala, R. (2021). Early detection of heart diseases using a low-cost, compact ECG sensor. *Multimedia Tools and Applications, 80*, 32615–32637. https://doi.org/10.1007/s11042-021-11010-3

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., … Dean, J. (2019). A guide to deep learning in healthcare. *Nature Medicine, 25*(1), 24–29. https://doi.org/10.1038/s41591-018-0316-z

Fuchs, F. D., & Whelton, P. K. (2020). High blood pressure and cardiovascular disease. *Hypertension, 75*(2), 285–292. https://doi.org/10.1161/HYPERTENSIONAHA.119.14240

Gárate-Escamila, A. K., El Hassani, A. H., & Andrès, E. (2020). Classification models for heart disease prediction using feature selection and PCA. *Informatics in Medicine Unlocked, 19*, 100330. https://doi.org/10.1016/j.imu.2020.100330

Gupta et al 2022 J. Phys.: Conf. Ser. 2161 012013 https//doi.org/10.1088/1742-6596/2161/1/012013

NPRC Journal of Multidisciplinary Research
Volume 3, Issue 1, January, 2026 Pages: 63-79
ISSN: 3059-9148 (Online)
DOI: 10.3126/nprcjmr.v3i1.90028

Johnson, K. W., Torres Soto, J., Glicksberg, B. S., Shameer, K., Miotto, R., Ali, M., … Dudley, J. T. (2018). Artificial intelligence in cardiology. *Journal of the American College of Cardiology, 71*(23), 2668–2679. https://doi.org/10.1016/j.jacc.2018.03.521

Joshi, P., Islam, S., Pais, P., Reddy, S., Dorairaj, P., Kazmi, K., … Yusuf, S. (2019). Risk factors for early myocardial infarction in South Asians compared with individuals in other countries. *JAMA, 297*(3), 286–294. https://doi.org/10.1001/jama.297.3.286

Krittanawong, C., Zhang, H., Wang, Z., Aydar, M., & Kitai, T. (2017). Artificial intelligence in precision cardiovascular medicine. *Journal of the American College of Cardiology, 69*(21), 2657–2664. https://doi.org/10.1016/j.jacc.2017.03.571

Li, J. P., et al. (2020). Heart disease identification method using machine learning classification in e-healthcare. *IEEE Access, 8*, 107562–107582. https://doi.org/10.1109/ACCESS.2020.3000850

Muhammad, Y., Tahir, M., Hayat, M., & Chong, K. T. (2020). Early and accurate heart disease detection and diagnosis using an intelligent computational model. *Scientific Reports, 10*(1), Article 19747. https://doi.org/10.1038/s41598-020-76635-9

Olimjonovna, K. O. (2024). The link between diabetes and heart disease. *Biologiya va Kimyo Fanlari Ilmiy Jurnali, 2*(5), 29–35.

Roth, G. A., Mensah, G. A., Johnson, C. O., Addolorato, G., Ammirati, E., Baddour, L. M., … Murray, C. J. L. (2020). Global burden of cardiovascular diseases and risk factors, 1990–2019. *Journal of the American College of Cardiology, 76*(25), 2982–3021. https://doi.org/10.1016/j.jacc.2020.11.010

Singh, D. P. (2024). An extensive examination of machine learning methods for identifying diabetes. *Tuijin Jishu/Journal of Propulsion Technology, 45*(2), 1380–1394.

Sonawane, R., & Patil, H. (2023). A design and implementation of a heart disease prediction model using data and ECG signal through hybrid clustering. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 11*(4), 1532–1548. https://doi.org/10.1080/21681163.2022.2156927