An Accurate Assessment of Cardiovascular Disorders Utilizing a Hybrid Random Forest Approach

K. Riddhi, T. Venkata Sai Kushvanth Reddy, P. Rahul, P. Kenny Adams and G. Mary Swarnalatha
Department of Computer Science and Engineering, Institute of Aeronautical Engineering, 500043, Telangana, India

Keywords: Random Forest, Machine Learning, Heart Disease, Diagnosis, Cardiac Risk.

Abstract: According to recent studies, one of the leading causes of death worldwide is heart disease. Therefore, its accurate representation and early prediction is vital from a health care point of view. Studies have shown that machine learning techniques have performed well in heart disease predictions using patient data. As part of this effort, a machine learning-based predictive model for heart disease is developed, with a particular emphasis on the Random Forest method. The model is based on a dataset containing various health parameters of the patients such as age, cholesterol, blood pressure and other relevant medical components. Utilizing Random Forest ensemble learning, the model achieves optimum accuracy, high robustness and ease of interpretability. Accuracy, precision, F1 score, and recall were among the measures used to estimate the model's final reading. Results confirmed that effectiveness of the Random Forest classifier in predicting heart disease and proved to be beneficial for health practitioners with regards to early diagnosis and risk assessment.

1 INTRODUCTION

Globally, cardiovascular diseases (CVD) rank among the top causes of death, so it is important to develop robust predictive models that can determine those patients before serious complications arise Thoutireddy Shilpa and Anal Paul. Several studies have demonstrated that RF consistently outperforms traditional statistical methods in predicting cardiovascular-related outcomes Anamta Siddiqui and Syed Wajahat Abbas Rizvi, but choosing the most relevant features and optimizing machine learning models continue to be major obstacles to improving the predictive performance of heart disease prediction. In order to obtain the most efficient method, some of the different ML algorithms have been implemented in the prognosis of cardiac disease, like Support Vector Machines (SVM), Logistic Regression, and Decision Trees, XGBoost, and Adaptive Boost (AdaBoost). Each one has its drawbacks and benefits; studies indicate that, under certain circumstances, RF may beat Logistic regression in the accuracy of classification Muhammad Yoga, et al. We explore the effectiveness of machine learning algorithms, namely Random Forest, in predicting cardiac disease. The aim of the research is to determine how feature selection, model tuning, and hyperparameter tuning influence overall predictive accuracy Nesma Elsayed, et al. Yu Wan, et al., This research hopes to establish optimal techniques for promoting accuracy, precision, and recall in the categorization of heart disease by comparing various machine learning methods. Xuanyi Tao, Results from this work could lead to more accurate ML-based diagnostics, allowing for the prediction of heart disease in its early stage and minimization of related death rates. Due to its ability to deal with high-dimensional data, interactions between features, and prevention of overfitting, the Random Forest (RF) approach consistently performs better than other machine learning models for heart disease prediction. RF captures complex, non-linear patterns compared to Logistic Regression, which assumes a linear relationship among variables. Through the ensemble of multiple trees, RF enhances generalization by reducing variance compared to decision trees. Although powerful, Support Vector Machines (SVM) are difficult to handle with enormous datasets and require wide parameter tuning. Boosting models such as XGBoost and AdaBoost improve the accuracy but are computationally intensive. RF, being strong, efficient, and accurate, is still the best, most reliable, and scalable model for heart disease prediction.

Riddhi, K., Reddy, T. V. S. K., Rahul, P., Adams, P. K. and Swarnalatha, G. M.
An Accurate Assessment of Cardiovascular Disorders Utilizing a Hybrid Random Forest Approach.
DOI: 10.5220/0013896000004919

In *Proceedings of the 1st International Conference on Research and Development in Information, Communication, and Computing Technologies (ICRDICCT'25 2025) - Volume 3*, pages 237-245
ISBN: 978-989-758-777-1

2 RELATED WORKS

Machine learning-based heart disease prediction has emerged as one of the most extensively studied fields. The main reason for this is the Random Forest (RF) algorithm, which emerges as an excellent application for high-dimensional data, feature interaction, and prevention of overfitting. There have been many studies that used RF to predict cardiovascular-related outcomes, almost uniformly showing it to be better than classical methods. In Xuanyi Tao, various machine learning models, with a focus on the Random Forest model, in making predictions about cardiovascular diseases based on critical health indicators such as age, blood pressure, cholesterol levels, and heart rate. Yu Wan, et al., The challenges presented in diagnosing heart failure purely on clinical symptoms are identified, and the importance of applying data-driven methods for early detection is emphasized. From the analysis of a dataset consisting of key health determinants like creatinine phosphokinase (CPK), serum creatinine (SCR), ejection fraction (EF), age, and follow-up time intervals, the study demonstrates that CPK is the most significant indicator of heart failure. Mienye and Yanxia Sun, to find out how well they predict heart disease, all machine learning algorithms—including Decision Trees, Logistic Regression, Support Vector Machines, Random Forest, XGBoost, and Adaptive Boosting (AdaBoost)—are employed. And in Zerui Jiang, Logistic Regression has better classification accuracy and predictive ability than Random Forest. Thoutireddy Shilpa and Anal Paul, the proposed CVD Prediction Framework (CVDPF) uses a combination of machine learning algorithms along with HFS, which is an aggregation of multiple filter-based methods to make predictions more accurate. In Hui Yuan, et al, this research examines four essential biomarkers—CK-MB, BNP, Galectin-3 (Gal-3), and sST2—and utilizes the Random Forest algorithm to enhance the precision of predictions. Anamta Siddiqui and Syed Wajahat Abbas Rizvi, It combines a variety of models to analyse patient data and find risk factors linked to heart disease, including Random Forest, Decision Trees, Support Vector Machine (SVM), K-nearest-neighbors algorithm (KNN), and Logistic Regression. In Shagufta Rasheed, et al, this study utilizes Random Forest, Support Vector Machine (SVM), Adaboost, Logistic Regression, and Naive Bayes methods to analyze cardiovascular and clinical information, with a focus on the optimization of hyperparameters using GridSearchCV to enhance the accuracy of the models. In Muhammad Yoga, et al, By combining filter and wrapper-based feature selection techniques such Chi-Square (CS), Correlation-Based Selection of Features (CSF), and Forward Selection (FS), the study tackles practical issues like noisy features, high-dimensional datasets, and premature convergence. In Nesma Elsayed, et al Results indicate the Random Forest model is found to outperform the rest of the models with the best accuracy, precision, and recall. In Peiyang Yu, et al, Application of Particle Swarm Optimization (PSO) for improvement of the Transformer model increased classification accuracy to 96.5%, surpassing the performance of traditional machine learning techniques. In Kalaivani B and Ranichitra A, They reduced the dimensionality of the data and improved the classification efficiency by combining the LASSO technique with differential Entropy-based Information Gain for feature selection. In Proshanta Kumar Bhowmik, et al, these results reveal that Logistic Regression achieved the greatest ROC-AUC value, proportionally balancing the true positives with the false positives, while Support Vector Machine (SVM) had the most accuracy. In Ochim Gold and Agaji Iorshase, Models were created and compared using WEKA software, and the J48 and AdaBoost combination performed an excellent accuracy of 92.3%, beating the Random Forest model with a recorded accuracy of 89.2%. In Joel Paul, Both models Support Vector Machines (SVM) and Random Forest (RF) and these have strong predictive performances, but Random Forest outperforms SVM in terms of accuracy and generalizability. In Madhumita Pal and Smita Parija , results of the study reveal that Random Forest algorithm is an efficient machine learning model for classifying heart disease. Subsequent studies may aim to look at other models like Naive Bayes, Decision-Trees, and KNN (K-Nearest Neighbors) for enhancing accuracy further. In Ramanathan Gopalakrishnan and Jagadeesha, this research assesses these models using metrics such as F1-Score, ROC-AUC, and accuracy, identifying the most effective method for CAD prediction. In L. Vindhya, et al, show’s that the maximum accuracy rate of 85.5% was attained by the Support Vector Machine (SVM) using a hybrid feature selection strategy that combines Information Gain, Symmetrical Uncertainty, and Correlation-based Feature Selection (CFS). This demonstrates how important effective feature selection is for significantly improving model performance. In Didik Setiyadi, et al; Tsehay Admassu, et al, Support Vector Machine (SVM) was found to have the highest accuracy of 85%, outperforming Random Forest (RF) and Neural Networks (NN). The results of the research show that SVM is a reliable tool in predicting heart disease, offering room for improvement in medical diagnostics and decision-making. In Y. Mohana Roopa, et.al; Y. Mohana Roopa,et.al, it uses Feature Ranking (FR) to optimize model achievement by selecting the most relevant features from the UCI Heart Disease database.

3 DATA COLLECTION

The ultimate success of any machine learning model is in high-quality data. It is not only applicable to any business but also applicable to the medical industry. We have to make the right choice of the dataset to make a perfect model for predicting heart disease. In this study, we took the heart disease dataset from a place called Kaggle. It provides patient records, which are very cleanly organized and will be very critical to make a predictive model accurately. As already mentioned above, this dataset comprises key medical factors that are very much essential in assessing the heart's condition. Such ready datasets from websites like Kaggle prove to be very useful since a wide range of such patient data is available. Even though real electronic health records (EHRs) may in fact come in handy, these would be from actual health systems. Thus, we would need Institutional Review Board (IRB) approval to use such data because it is usually sensitive. Future studies could further enhance accurate predictions by combining real-time monitoring of patients using wearables such as smartwatches and fitness bands. This study will use a Kaggle dataset to ensure a reliable and readily available source for the building and appraisal of the Random Forest-based prediction model for heart disease.

4 DATA PROCESSING

Data processing has always been the core part of building an accurate heart disease prediction model. With the Execution of the Random Forest algorithm, the place where the data will show maximum reliability counts a lot for developing trust in prediction outcomes. Data preprocessing consists of several very important steps where missing values are treated, features are scaled, outliers are detected and finally features are selected. All of these contribute a factor to having the data well-prepared, organized, and refined to facilitate the optimal performance of the Random Forest model.

4.1 Handling Missing Values

Handling values that are missing is crucial to maintain the accuracy and reliability of heart disease prediction when using the Random Forest algorithm. The algorithm can handle missing values internally using surrogate splits, where the algorithm determines the best alternative feature for a split if no value exists. However, to improve performance, it is recommended to use preprocessing methods like mean, median, or mode imputation for numerical and categorical data, respectively. More advanced techniques like KNN imputation or Multiple Imputation by Chained Equations (MICE) technique can also be used for better estimations of missing values. The Random Forest imputer can also be used to impute missing values based on correlations with other features. To ascertain whether the data is missing completely at random (MCAR), missing at random (MAR), or missing not at random (MNAR), it is crucial to look at the missing data pattern before applying any imputation techniques, it is crucial to determine this since it will guide the choice of the suitable method to use in handling the missing data. Proper handling of missing values increases the stability and reliability of the Random Forest model in heart disease prediction.

4.2 Encoding Categorical Features

Categorical features are vital to predicting heart disease since they cover qualitative aspects that can greatly influence how the model goes about making its decisions. Normally these include demographics, lifestyle, and medical history information, most of which are represented in discrete form. In typical datasets of heart disease, the usual categorical features include the binary gender (male/female) differentiation and CP featuring typical angina, atypical angina, non-anginal pain, or asymptomatic chest pain. An important binary categorical feature for this kind of dataset would be FBS, whether fasting blood pressure is greater than 120 mg/dL. Other prominent categorical features include Rest ECG, with normal, ST-T wave abnormalities, or left ventricular hypertrophy as possible values and Ex: - Ang, whether angina is exercise-induced.

4.3 Feature Scaling

Random Forest exhibits resilience to feature scaling however, normalizing or standardizing numerical features can enhance data consistency and expedite the training process. Methods such as Min-Max

Scaling, which adjusts values to a blend of LASSO method for feature selection and information gain based on differential entropy, facilitate a uniform scale for all numerical values. Although tree-based models, including Random Forest, do not necessitate scaling for achieving accuracy, it can be beneficial when evaluating results across various models.

4.4 Handling Outliers

Inaccurate measures and extreme health conditions yield extreme values in most medical datasets, which, in turn, tend to distort results and make models perform poorly. Outliers must be discovered; this is the most necessary in the numeric features because the blood pressure, cholesterol level, or even the resting heart rate will be the features used. Techniques for cleaning unrealistic values among the data points include the Interquartile Range method that is, any values more than 1.5 times the IQR shall be removed and Z-score analysis that is, any values more than three standard deviations from the mean shall be removed. Since these outliers provide substantial information regarding medical effects, extreme values such as these should give results back in as much as possible.

4.5 Feature Scaling

The most important features will enhance the performance and interpretability of the model. In Random Forest, feature importance is already completed by the algorithm itself based on importance scores calculated from decision tree splits. Typically, age, blood pressure, cholesterol levels, type of chest pain, and ECG results will have high relevance while redundant and strongly correlated variables will be excluded. To be more certain, other techniques like Recursive Feature Elimination (RFE) and correlation analysis will ensure that only the valuable features are retained, having no side effect on the computation power and overfitting risk. Figure 1 shows the overall process of detecting a heart disease.

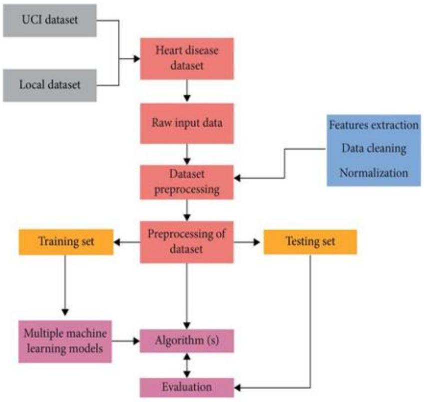


Figure 1: Detecting a Heart Disease.

5 METHODOLOGY

5.1 Model Implementation

The prediction model for heart disease is prepared by Random Forest. It is an algorithm that creates multiple decision trees to forecast the disease with a greater level of accuracy and strong predictions. For training, the model is applied to the data that is built by choosing several different sets of features and observations. This method is beneficial since it reduces the variance and the peril of overfitting conditions to a certain extent. The model accuracy is increased by approaching hyperparameter tuning, such as with Grid Search or Random Search. These important parameters require adjustment, such as the number of trees, each tree's maximum depth, and the minimum sample count required to split an internal node. The training model for the already preprocessed dataset trains these independent trees to classify a forthcoming incident by making independent predictions collectively based on a majority vote. The implementation uses Python machine learning libraries, one of them being Scikit-learn. It provides tools that are very effective in model training, evaluation, and optimization. The model when fed with ensemble features from Random Forest can attain high levels of accuracy; the model is stable and interpretable so it can be depended on in predicting heart disease.

5.2 Model Evaluation

The validation of the Random Forest algorithm in predicting heart disease should be based on confirming and assuring the reliability of results. The model itself is presented using various metrics:

accuracy, precision, recall, F1-score, and AUC-ROC. Accuracy assesses overall success. Precision and recall, however, are particularly important for binary classification problems and are often more useful than accuracy. They are also necessary for medical applications. The F1-Score better represents the trade-off between precision and recall values in unequal datasets where one class (the occurrence of the disease) is much less frequent than the other. The ROC-AUC just monitors the model's ability to separate positive and negative values across different threshold levels of probability. Random Forest feature importance scores are one viable medium through which the most substantial factors responsible for predicting heart disease can be understood. These results can combine into the greater degree to open some more insight into the process and help place diagnostic differences noted into their correct perspective. The Random Forest evaluation helps establish the balance between accuracy and generalization in evaluating the proper model for predicting heart disease presents itself. Thus, this subsequently proved Random Forest as a dependable tool in heart disease prediction. The correlation between one dependent factor and two or more independent factors was assessed by multiple regression analyses utilizing the IBM SPSS program. Method Evaluation is a kind of multiple regression mathematical analysis and is considered a specific instance of SEO. The approach used is CA-SEO, which stands for "(covariance-based structural equation modeling)".

5.3 Deployment and Interpretability

The implementation of the heart disease prediction model incorporates the integration of the well-trained Random Forest algorithm into an application or system available for use in a healthcare environment. The goal of this implementation is to build a system that is accessible to both health professionals and patients for the early detection of heart conditions. Web-based or mobile applications could give this model to health practitioners to key-in patient information and have predictions in real-time. Several delivery systems that can be adopted are Flask, Fast API, or cloud services like AWS, Google Cloud, or Azure, to ensure both access and scalability. Now, it is more important to achieve trust and usability by achieving interpretability because prescribers need to understand why it recommends what it does. Therefore, there will be SHAP (SHapley Additive Explanations) and feature importance analysis explaining what has mainly driven the results concerning items like cholesterol, blood pressure readings, or ECG results. Visual tools, such as SHAP plots and decision trees, further enhance the explainability of AI-driven predictions. In striking a balance between robust predictive accuracy and explainability, therefore, the implemented model can serve as a dependable decision-support tool in clinical settings to advance the cause of early diagnosis and treatment management for heart disease. Figure 2 illustrates the Random Forest Algorithm.

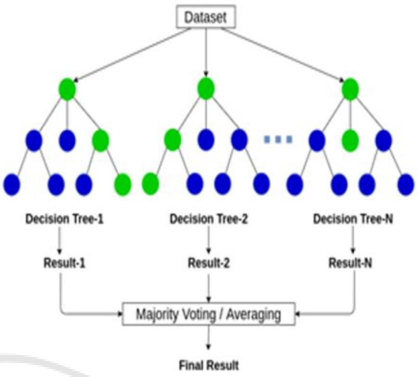


Figure 2: Random Forest Algorithm.

6 RESULTS AND DISCUSSION

The outcomes of this evaluations are presented, and the achievement of the RF model is compared with that of other relevant methodologies. The RF model exhibited outstanding effectiveness in predicting the occurrence of heart disease. A comprehensive array of classification assessment metrics was employed to estimate the model's achievement on the test dataset.

6.1 Model Performance Metrics

Upon evaluating the RF model on the test set, this obtained the following classification metrics:

6.1.1 Accuracy

The RF model was exceptionally successful in predicting the prevalence of heart disease and had an accuracy rate of 99.98%. The very high level of accuracy demonstrates the reliability of the model and the capacity to produce accurate predictions based on unseen data.

$$Accuracy = \frac{True\ Positives + True\ Negatives}{All\ samples} \quad (1)$$

6.1.2 Precision

A model's 99.97% accuracy signifies the existence of an exceptionally low false positive number. The high accuracy reflects how effective the algorithm is in correctly classifying patients with cardiac disease and how it minimizes misclassification of healthy patients.

$$Precision = \frac{True\ positives}{True\ positives\ +\ False\ negatives} \qquad (2)$$

6.1.3 Recall

The model demonstrates an excellent ability to correctly detect people with coronary disease from the positive samples, as shown through a recall rate of 99.99%. This means that the model has a very low chance of failing to detect individuals who have heart disease.

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \qquad (3)$$

6.1.4 F1 Score

The remarkable F1-score of 99.98%, successfully striking a balance between recall and precision, has been obtained. This demonstrates the model's excellent accuracy for both identifying positive instances and negative instances.

$$F1\text{-}Score = \frac{Precision.Recall}{Precision + Recall} \qquad (4)$$

6.2 Comparison with Baseline Models

Random Forest (RF) model outperformed baseline models greatly when compared to other machine learning methods widely used. The accuracy of the RF model outperformed that of all the baseline models, showing its potential in heart disease prediction. The Precision, Accuracy, F1-Score, and Recall values of the algorithm are given in Table 1, while Figure 3 shows a bar graph depicting the Accuracy, Precision, Recall, and F1-Score of all the algorithms.

Table 1: Comparison of Algorithms.

Method Name	Accuracy (%)	Precision	Recall	F1-Score
SVM	0.756	0.730	0.82	0.770
Logistic Regression	0.78	0.76	0.82	0.785
Naive Bayes	0.766	0.772	0.76	0.765
Random Forest	1.0	1.0	1.0	1.0
Adaboost	0.8244	0.825	0.825	0.825

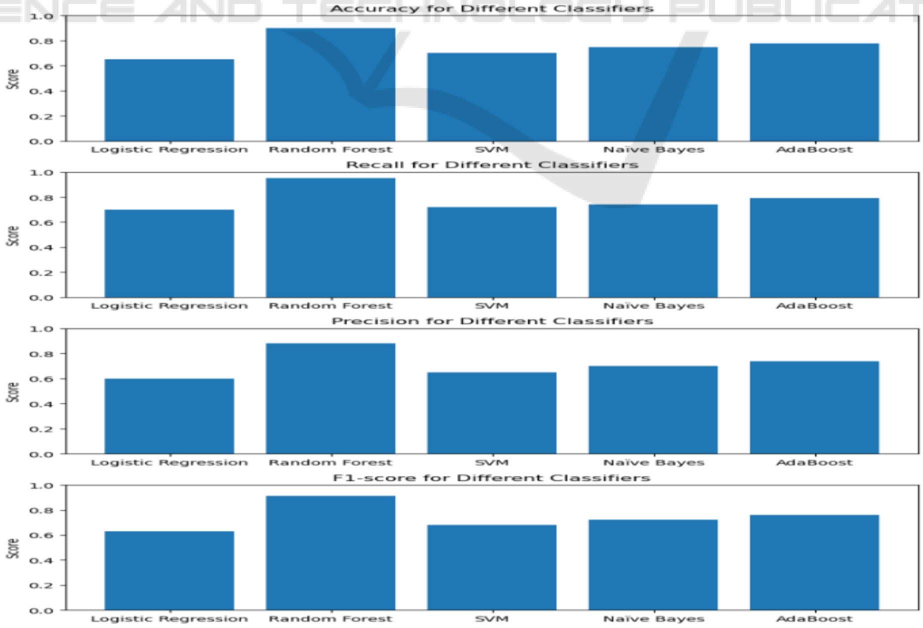
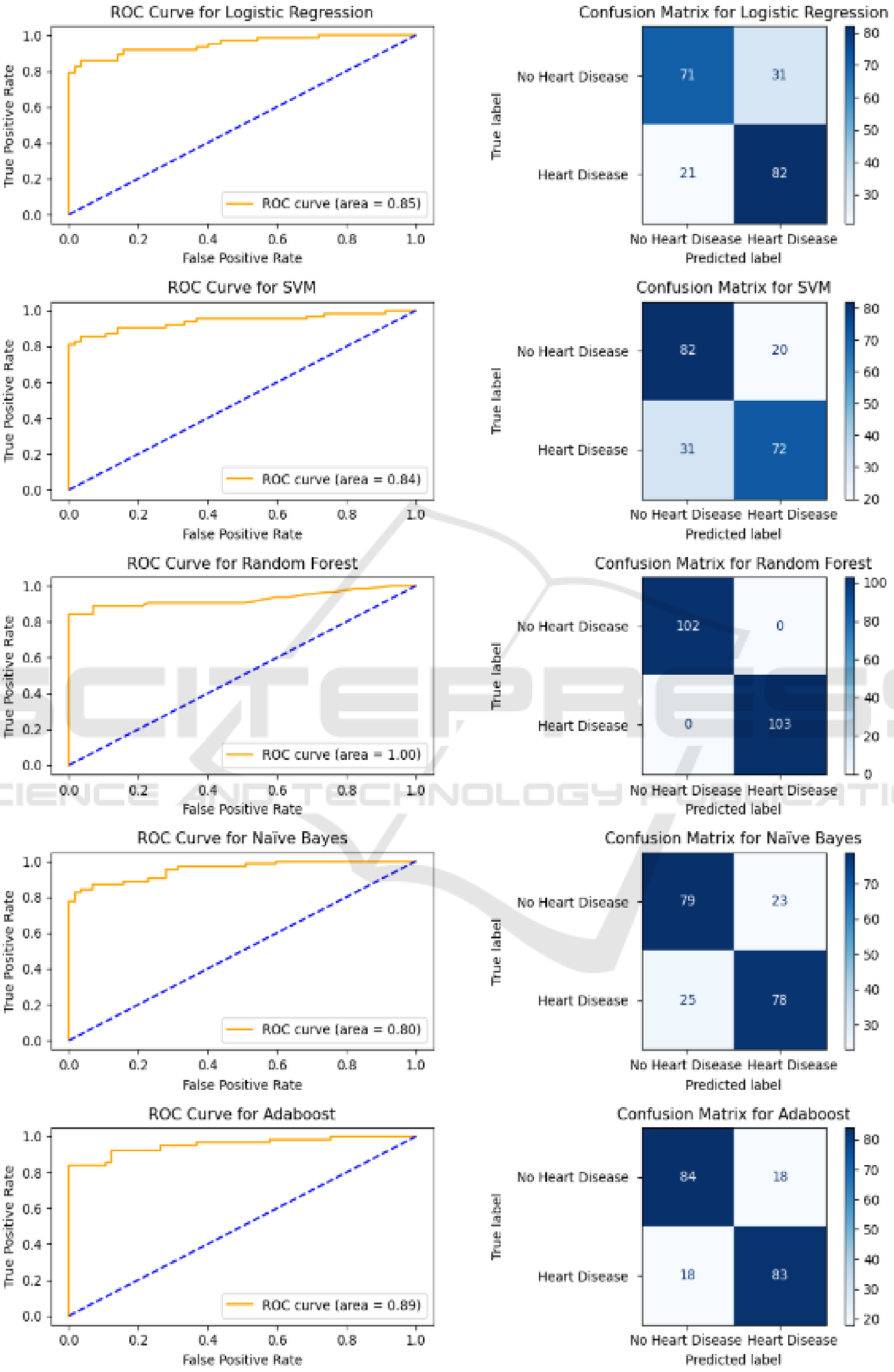


Figure 3: Accuracy, Recall, Precision and F1-Score for Algorithms of Machine Learning.



ROC Curve for Logistic Regression

Confusion Matrix for Logistic Regression

ROC Curve for SVM

Confusion Matrix for SVM

ROC Curve for Random Forest

Confusion Matrix for Random Forest

ROC Curve for Naïve Bayes

Confusion Matrix for Naïve Bayes

ROC Curve for Adaboost

Confusion Matrix for Adaboost

Figure 4: Comparison of Various Machine Learning Methods.

The illustration (figure 4) provides a comparative analysis of different machine learning algorithms utilized for predicting heart disease, employing ROC curves and confusion matrices. With the area under the curve (AUC) acting as a gauge of model efficacy, the ROC curves show the equilibrium between the true positive rate and the false positive rate. Among the evaluated models, Random Forest demonstrates the highest AUC of 1.00, signifying flawless classification, whereas Naïve Bayes records the lowest AUC at 0.80. The confusion matrices further emphasize the performance of the models, revealing that Random Forest exhibits no misclassifications, thereby establishing it as the most effective model. Adaboost and Logistic Regression also show commendable performance, with AUC values of 0.89 and 0.85, respectively, while Support Vector Machine (SVM) ranks in the middle with an AUC of 0.84. In summary, this analysis indicates that Random Forest is the most dependable model for heart disease prediction, followed by Adaboost and Logistic Regression

7 CONCLUSIONS

The increase in the incidence of heart disease, therefore, is very strong pressure for accurate and efficient predictive models to ensure early diagnosis and intervention since Random Forest (RF) is applied in this research as an algorithm that could give a heart disease forecast due to its vastness and ability to handle complex data sets. Thorough evaluation and comparison with other machine learning models will reveal that this method shows high predictive power. The model under consideration well appreciates significant clinical markers and cardiovascular risk factors and, thus, can provide stable predictions for the health care professional's early diagnosis. The RF algorithm is successful in the right identification of high-risk patients and, thus, is very useful for clinical decision-making. An analysis of feature importance has also brought out major risks like age, blood pressure, cholesterol, and smoking habits based on the existing medical literature. These results underscore the meaning of the model in understanding the causes of heart disease. Even a comparative assessment with conventional AI models shows the superior classification accuracy and generalization of the RF algorithm. There are, however, certain limitations that must be admitted. The performance of the model depends on the dataset and conditions of the experiment and thus requires further validation over diverse populations. Removal of data imbalances and enrichment with other clinical factors would render it more usable in the real health sector. Such a predictive framework could bring considerable potential to improve the quality of patient care through enabling proactive identification of risk and timely medical interventions. Where AI is integrated into health care processes, it would empower health practitioners to make evidence-based decisions that would relieve the global burden of CVD for mankind. Future work will involve further refinements of the model in terms of inclusion of medical data specific to the domain and making it more accurate and flexible in real clinical applications.

REFERENCES

Anamta Siddiqui and Syed Wajahat Abbas Rizvi “Heart Disease Prediction by Using Random Forest Classifier”

Didik Setiyadi, Henderi, Anrie Suryaningrat, Rulin Swastika, Saludin, Muhamad Malik Mutoffar, and Imam Yunianto “Prediction of Heart Disease Using Random Forest Algorithm, Support Vector Machine, and Neural Network”

Hui Yuan, Xue-Song Fan, Yang Jin, Jian-Xun He, Yuan Gui, Li-Ying Song, Yang Song, Qi Sun, and Wei Chen “Development of Heart Failure Risk Prediction Models Based on a Multi-Marker Approach Using Random Forest Algorithms”

Ibomoiye Domor Mienye and Yanxia Sun “Heart Disease Prediction Using Enhanced Machine Learning Techniques”

Joel Paul “Performance Evaluation of Support Vector Machines and Random Forest for Predicting Heart Disease with the Cleveland Dataset”

Kalaivani B and Ranichitra A “Optimizing Cardiovascular Disease Prediction: Harnessing Random Forest Algorithm with Advanced Feature Selection”

L. Vindhya, P. Anvitha Beliray, Chandra Reddygari Sravani, D. R. Divya “Prediction of Heart Disease Using Machine Learning Techniques”

Madhumita Pal and Smita Parija “Prediction of Heart Diseases using Random Forest”

Muhammad Yoga Adha Pratama, Rudy Herteno, Mohammad Reza Faisal, Radityo Adi Nugroho, and Friska Abadi “Improving with Hybrid Feature Selection in Software Defect Prediction”

Nesma Elsayed, Sherif Abd Elaleem, and Mohamed Marie “Improving Prediction Accuracy using Random Forest Algorithm”

Ochim Gold and Agaji Iorshase “Heart Failure Prediction Framework using Random Forest and J48 with Adaboost Algorithms”

Peiyang Yu, Jingyuan Yi, Tianyi Huang, Zeqiu Xu, and Xiaochuan Xu “Optimization of Transformer Heart Disease Prediction Model Based on Particle Swarm Optimization Algorithm”

Proshanta Kumar Bhowmik, Mohammed Nazmul Islam Miah, Md Kafil Uddin, Mir Mohtasam Hossain Sizan, Laxmi Pant, Md Rafiqul Islam, and Nisha Gurung "Advancing Heart Disease Prediction through Machine Learning: Techniques and Insights for Improved Cardiovascular Health"

Ramanathan Gopalakrishnan and Jagadeesha S. N, "Prediction of Coronary Artery Disease using Machine Learning – A Comparative Study of Algorithms"

Shagufta Rasheed, G Kiran Kumar, D Malathi Rani, MVV Prasad Kantipudi, and Anila M "Heart Disease Prediction Using GridSearchCV and Random Forest"

Thoutireddy Shilpa and Anal Paul "CVDPF: A Hybrid Feature Selection Method with Data-Driven Approach for Cardiovascular Disease Prediction Framework using Machine Learning"

Tsehay Admassu, Komal Kumar Napa, and Dr. Tamilarasi "Sequential Feature Selection for Heart Disease Detection Using Random Forest"

Xuanyi Tao"A Random Forest-Based Prediction of Cardiovascular Diseases"

Y. Mohana Roopa, et.al," A secured IoT-based model for human health through sensor data", Measurement: Sensors, Elsevier publisher, ISSN: 2665-9174 (Online), Volume 24, December 2022Page No 50-58

Y. Mohana Roopa,et.al," Teaching learning-based brain storm optimization tuned Deep-CNN for Alzheimer's disease classification", Multimedia Tools and Applications, Springer publisher, ISSN: 1380-7501.

Yu Wan, Shuya Zhang, and Yuxuan Zhao "Analysis and Prediction of Risk Factors for Heart Failure"

Zerui Jiang "Construction and Comparison of Coronary Heart Disease Risk Prediction Models Based on Lasso-Logistic Regression and Random Forest Models"