Smart Machine Learning-Based Heart Disease Prediction with Random Forest Classifier

Sujatha Krishna
Department of Computing and Information Sciences, College of Computing and Information Sciences, University of Technology and Applied Sciences, AI-Aqr, Shinas, Oman
Sujatha.krishna@utas.edu.om (corresponding author)

Pandian Vaidhyanathan
Department of Supportive Requirements, Sciences and Mathematics Unit, University of Technology and Applied Sciences, AI-Aqr, Shinas, Oman
pandian.vaidhyanathan@utas.edu.om

Amina Salim Mohammed AlJabri
Department of Computing and Information Sciences, College of Computing and Information Sciences, University of Technology and Applied Sciences, AI-Aqr, Shinas, Oman
amina.aljabri@utas.edu.om

Amna Salim Rashid Al Kaabi
Department of Computing and Information Sciences, College of Computing and Information Sciences, University of Technology and Applied Sciences, AI-Aqr, Shinas, Oman
amna.alkaabi@utas.edu.om

Received: 7 February 2026 | Revised: 6 March 2026, 24 March 2026, 27 March 2026, 29 March 2026, 1 April 2026, and 3 April 2026 | Accepted: 4 April 2026

DOI: https://doi.org/10.48084/etasr.18029

ABSTRACT

Heart diseases remain one of the most common causes of death worldwide, indicating a demand for reliable and accurate early prediction systems. In contrast to established comparative studies, this one develops a unique data-driven predictive framework that encompasses multiple Machine Learning (ML) algorithms and an optimized Random Forest (RF)-based classification strategy for enhanced clinical decision-making support. A systematic heart disease prediction model was designed, based on Logistic Regression (LR), Gaussian Naïve Bayes (GNB), Support Vector Classifier (SVC), K-Nearest Neighbors (KNN), Decision Tree (DT), Extra Trees (ET), Bagging, and an optimized RF classifier aimed at increasing predictive performance stability and generalizability. Experimental results show that RF achieves better performance with 97.50 accuracy, 97.51 precision, 97.50 recall, and 97.50 F1-score on the validation dataset. The proposed model integrates ensemble-based optimization and effective feature learning and generalization ability for medical risk prediction. Its better performance is due to the ensemble learning mechanism of RF, which increases model robustness, helps to decrease overfitting, and increases classification reliability. With its ability to accurately direct clinicians and health organizations to unknown patients at risk, the proposed model can offer great value in clinical and health economy perspectives, through timely intervention for both diagnosis of imminent heart disease complications and varied data-informed prescription steps.

Keywords-heart disease prediction; Random Forest (RF); Machine Learning (ML); medical diagnostics

I. INTRODUCTION

Cardiovascular disease is one of the most common causes of death worldwide and poses a great burden on healthcare systems, highlighting the importance of prompt diagnosis with high precision [1]. Recent global health statistics and epidemiological scenarios show a persistent increase in the prevalence and mortality of cardiovascular disease, highlighting the need for efficient early prediction and prevention. Late discovery of heart disease leads to no effective treatment and increases the risk of severe complications, indicating that early warning and prevention are paramount. Data-driven systems have been increasingly investigated to translate clinical data into actionable diagnostic information to support the medical decision-making process [2]. Recent research demonstrates that sensor fusion using Machine

Learning (ML) can improve the early detection of heart-related conditions by mining complex patterns in clinical data [3].

In the clinical field of heart disease detection, the diagnosis is based on several interrelated factors such as age, gender, type of chest pain, resting blood pressure, serum cholesterol, and fasting blood sugar results. Resting Electrocardiographic (ECG) testing and the maximum heart rate achieved during a thallium test study while exercising can also be used. These clinical features in combination are informative for cardiovascular health, but standard statistical models cannot find any meaningful relationship among them. This is because conventional methods are incompetent to reflect complicated and nonlinear patterns of multidimensional medical data, indicating that intelligent computational techniques need to be applied for accurate prediction [4].

ML has been gaining the spotlight in predictive healthcare due to its ability to learn from complex medical data sets and, therefore, perform automatic processing. Several ML approaches, such as Logistic Regression (LR), Gaussian Naïve Bayes (GNB), Support Vector Machines (SVM), Decision Trees (DT), and hybrid systems, have been effectively used for heart disease prediction, improving diagnostic accuracy compared to conventional methods [5, 6]. In addition, it is crucial to investigate a set of algorithms to choose the most reliable model for clinical use [7]. Furthermore, recent developments in AI-enabled cloud computing techniques have led to more scalable and efficient health information processing for disease diagnosis [8]. These trends highlight the increasing importance of ML in smart health systems.

Ensemble methods have attracted significant attention for their ability to enhance the prediction accuracy, increase model stability, and decrease overfitting. Methods like Bagging, Extra Trees (ET), and Random Forest (RF) ensemble several weak learners to form a strong predictive model that is more effective on difficult medical classification problems [9]. RF has been proven to be effective in numerous health prediction problems for its ability to handle high-dimensional data and model complex interactions among variables. In addition, recent works have discussed that the models should be transparent and interpretable to ensure trustworthiness and reliability of clinical decision-support systems [10].

Table I provides a comparative summary of previous works on ML approaches in heart disease prediction. Many studies have reported relatively high predictive accuracy, but lack a consistent framework to benchmark generalizability, interpretability, and robustness. This demonstrates the importance of a unified and systematic approach for benchmarking, highlighting the importance of this study in evaluating several ML models in a uniform experimental design aimed at identifying the most robust for the prediction of heart disease.

The main objective was to develop a robust and accurate ML model capable of predicting the existence of heart disease based on certain clinical features. Figure 1 summarizes the proposed framework that focuses on a transition from delayed diagnosis of heart diseases to early prediction based on ML. Multiple ML models are systematically benchmarked to determine which provides the utmost accuracy and reliability for early heart disease prediction. The purpose of the system is to aid in early diagnosis and risk assessment, helping clinicians make informed decisions, enhancing patient outcomes, and ultimately reducing heart disease-related mortality.

TABLE I. COMPARATIVE STUDY OF MACHINE LEARNING TECHNIQUES IN HEART DISEASE PREDICTION

Ref.	ML methods used	Dataset type	Evaluation metrics	Key contributions	Limitations
[11]	Multiple ML techniques (SVM, DT, LR, ANN)	Public heart disease datasets	Accuracy, Sensitivity, Specificity	Systematic review of ML-based heart disease prediction approaches	Focused mainly on accuracy comparison
[12]	SVM, KNN, LR, DT	Clinical heart datasets	Accuracy, Precision, Recall	Comparative implementation of ML algorithms	Limited dataset size
[13]	NB, RF, Gradient Boosting, AdaBoost	Heart disease datasets	Accuracy, Cross-validation	Demonstrated that ensemble models outperform single classifiers	Limited interpretability analysis
[14]	Neural Network + PSO Optimization	Cardiovascular dataset	Accuracy, F1-score	Integrated optimization with neural networks	Higher computational complexity
[15]	Explainable AI methods (SHAP, LIME)	Experimental datasets	Model interpretability measures	Focus on explainability and trust in AI models	Not specifically about heart disease
[16]	Multiple ML algorithms	Multi-disease datasets	Accuracy	Disease prediction using Applied ML	Limited disease-specific optimization
[17]	SVM, ANN, DT	Heart datasets + image fusion	Accuracy, Sensitivity	Fusion image processing with ML	Complex implementation
[18]	ML with normalization and cross-validation	Clinical dataset	Accuracy, ROC-AUC	Tuning and validating for better prediction	Dependent on preprocessing quality
[19]	ML algorithms in a cloud environment	Heart disease dataset	Accuracy	Scalable cloud-based ML healthcare system	Limited model comparison
[20]	Multiple ML classifiers	MLHeartDisPrediction dataset	Accuracy, Precision, Recall	Developed an ML-based heart prediction framework	Small dataset and limited generalization
[21]	ML with class imbalance optimization	Imbalanced heart dataset	F1-score, Recall	Used optimization techniques for addressing class imbalance	Examined only four models
[22]	Statistical & ML techniques	Simulation dataset	Error metrics	ML applied in engineering prediction	Not specific to the healthcare domain
[23]	XGBoost + SMOTE	Heart disease dataset	Accuracy, ROC-AUC	Performance was improved with boosting and imbalance mitigation	Computational cost considerations

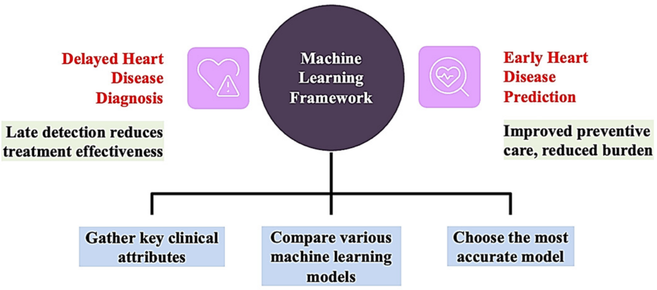


Fig. 1. ML for early heart disease prediction.

II. RESEARCH METHODOLOGY

This section discusses the methodology taken to create a realistic and authentic heart disease forecasting model based on ML. The process aims to turn clinical data into predictive insight through structured data preprocessing, feature analysis, and model building. A detailed comparative analysis of several ML algorithms is used to determine the best one for heart disease prediction. The whole procedure ensures more accurate results for better prediction, robustness, and reliability to promote early diagnosis in clinics.

Figure 2 represents the complete workflow of the proposed ML heart disease prediction system, which takes as input a raw clinical dataset (including patient health attributes). The dataset is then partitioned into two: training (80%) and testing (20%). Prior to model building, the training data is subject to data preprocessing, where feature scaling by utilizing the standard scaler normalizes input features for enhanced accuracy. The preprocessed data are used to train multiple classifiers, such as LR, GNB, SVM, KNN, DT, and ensembles like Bagging and RF. These models learn the characteristics from example training cases to differentiate between people with heart disease and those without. The trained models are evaluated on the testing dataset to obtain predictions on heart disease. Performance analysis considers popular metrics such as Accuracy, Precision, Recall, and F1 score for measuring the efficiency and robustness of each classifier. Model optimization is then performed according to the outcomes of the evaluation to enhance predictive performance and minimize prediction errors. The proposed framework provides a disciplined and structured methodology for creating predictive risk scores based on clinical data, which can help identify heart diseases at an early stage to support clinicians in their decision-making process.

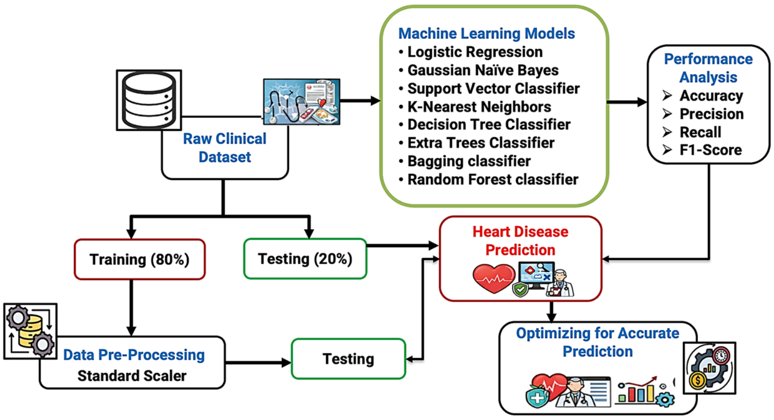


Fig. 2. An ML framework for predicting heart disease.

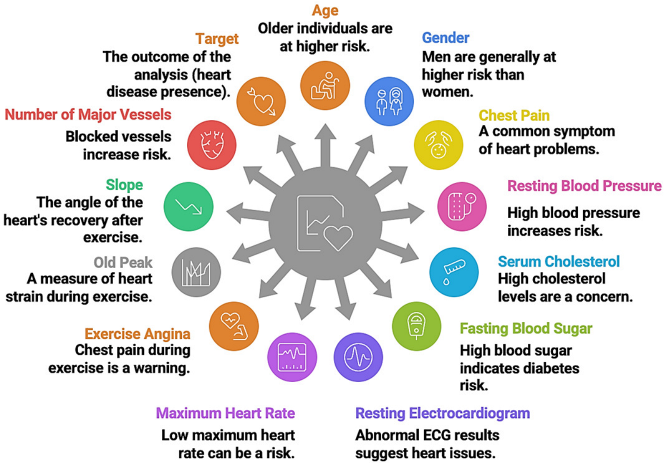


Fig. 3. Dataset features and description of heart disease risk factors.

A. Dataset

This study employed the Cardiovascular Disease Dataset [24], which contains 1000 patient records. Each instance contains 13 features, many of which are numeric and correspond to potential clinical, demographic, and physiologic predictors. These features carry information relevant to the patient's health status and are used to diagnose the presence of heart disease [24]. Figure 3 shows the features of the dataset and their clinical relevance.

The dataset has age and gender features that are influential demographic risks, as men have a higher risk than women of a certain age. Cardiac symptoms of chest pain and exertional angina suggest a possible cardiac origin. Physiological measures such as resting blood pressure, serum cholesterol, and fasting blood sugar are important indicators of cardiovascular health and metabolic risk. Quantitative features of the resting ECG and the achieved maximum heart rate offer the potential to assess multiple hierarchical levels of heart function in both the rest and the activity states. In addition, the old peak (ST depression) is representative of the strain on the heart during exercise, and the angle of the ST segment indicates the heart's recovery response. A greater number of major vessels indicates how blocked the arteries are and strongly relates to the risk of heart disease. The dependent features show the absence or presence of the disease in a patient. Together, these features create a well-rounded dataset intended to allow the development of ML models for prediction and risk inference on heart disease.

B. Exploratory Data Analysis

Table II shows the distribution of target classes in the dataset, showing that 58% of the records belong to patients with heart disease, while 42% represent patients without. The dataset is relatively balanced with a slight inclination in favor of the heart disease class. Such a distribution reflects reliable model training, with the caveat that this study wants to focus on positive cases, which are of extreme importance in medical prediction.

TABLE II. HEART DISEASE CLASS DISTRIBUTION

Class label	Description	Percentage
0	No Heart Disease	42.0%
1	Heart Disease	58.0%

Figure 4 shows the distribution of heart disease cases by type of chest pain. Under Typical Angina, there are 108 patients with heart disease and 312 without. For Atypical Angina, it has 154 patients with heart disease and 70 without. However, in the Non-anginal Pain group, 279 patients are diagnosed with heart disease and 33 without it, which is the highest proportion of diseased from all categories. The Asymptomatic category had 39 patients with heart disease and only 5 were without. These values suggest that the type of chest pain is crucial in identifying whether a patient has heart disease.

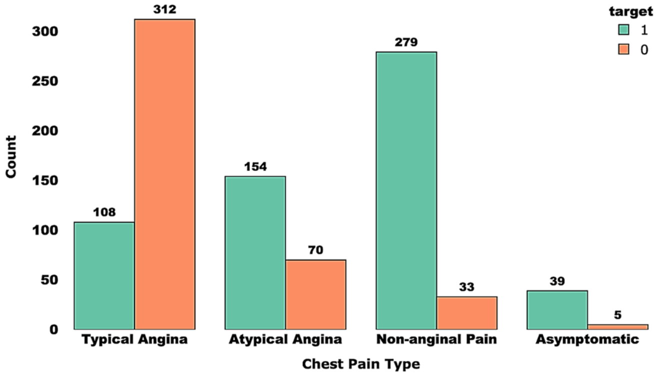


Fig. 4. Chest pain type versus heart disease.

Figure 5 demonstrates that as maximum heart rate increases, the prevalence of heart disease decreases. Blue bars represent patients with heart disease, and orange bars represent those without it. A maximum heart rate range between 130 and 170 indicates increased prevalence of heart disease, as the counts of affected patients far outnumber the non-affected. Patients without heart disease, at lower heart rate values, are relatively comparable or a little higher than those with heart disease. At the higher end, where maximum heart rate approaches the maximal value, the gap between groups shrinks. This indicates that the maximum heart rate is a significant physiological condition for heart disease prediction.

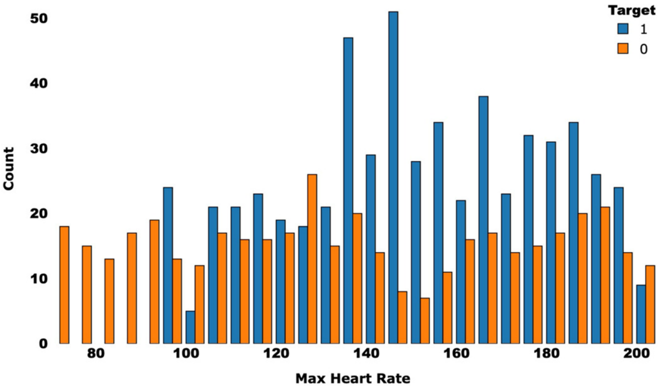


Fig. 5. Heart disease and maximum heart rate.

Figure 6 shows the relationship between ST-segment slope and heart disease. Patients with heart disease are in the blue bars, while those without are in orange. There were 310 patients with heart disease, and only 12 without heart disease in the flat slope category, which shows a good association of having a flat ST slope with the presence of disease. For the upsloping category, 21 patients had heart disease and 218 were at low risk. The downsloping column represents 199 people with heart disease and almost no values in the non-disease category. As can be observed, the ST slope is a very important feature for heart disease prediction.

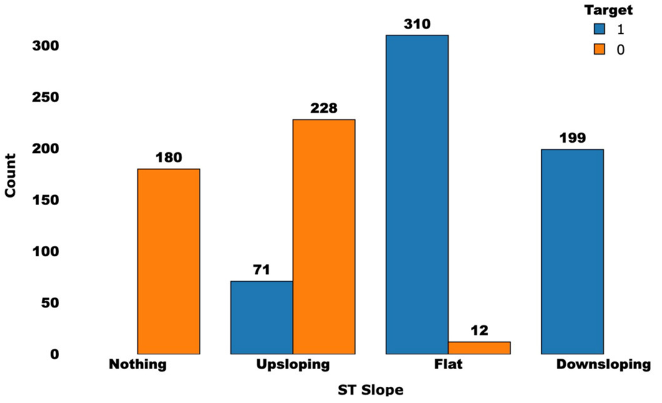
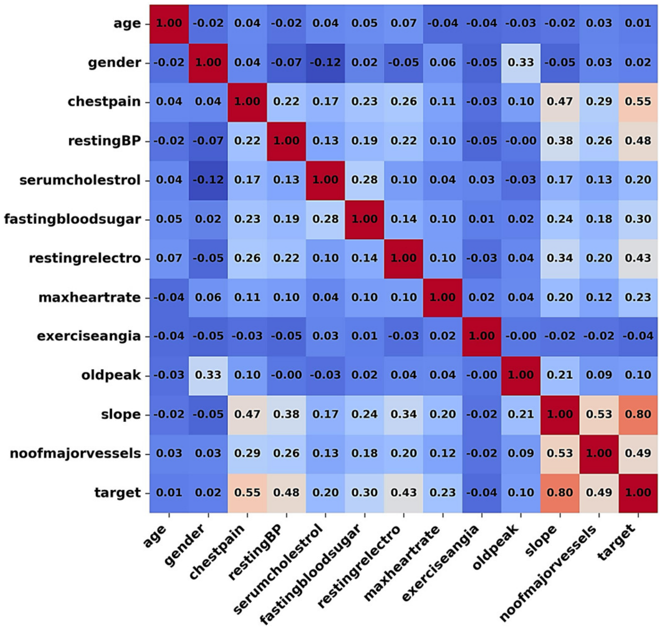


Fig. 6. ST slope vs heart disease.

Figure 7 shows the correlation matrix between clinical features and the target variable. Color intensity indicates the magnitude and direction of correlation, and values tending to 1 reflect a strong positive association. The output finally has a very high association with the slope of the ST segment and the number of major vessels, which reflects their significance in heart disease prediction. Chest pain, resting blood pressure, and resting electrocardiogram demonstrate fair correlation with the diagnosis. The matrix serves to determine the most important features for feature selection and model performance improvement.



	age	gender	chestpain	restingBP	serumcholestrol	fastingbloodsugar	restingelectro	maxheartrate	exerciseangia	oldpeak	slope	noofmajorvessels	target
age	1.00	-0.02	0.04	-0.02	0.04	0.05	0.07	-0.04	-0.04	-0.03	-0.02	0.03	0.01
gender	-0.02	1.00	0.04	-0.07	-0.12	0.02	-0.05	0.06	-0.05	0.33	-0.05	0.03	0.02
chestpain	0.04	0.04	1.00	0.22	0.17	0.23	0.26	0.11	-0.03	0.10	0.47	0.29	0.55
restingBP	-0.02	-0.07	0.22	1.00	0.13	0.19	0.22	0.10	-0.05	-0.00	0.38	0.26	0.48
serumcholestrol	0.04	-0.12	0.17	0.13	1.00	0.28	0.10	0.04	0.03	-0.03	0.17	0.13	0.20
fastingbloodsugar	0.05	0.02	0.23	0.19	0.28	1.00	0.14	0.10	0.01	0.02	0.24	0.18	0.30
restingelectro	0.07	-0.05	0.26	0.22	0.10	0.14	1.00	0.10	-0.03	0.04	0.34	0.20	0.43
maxheartrate	-0.04	0.06	0.11	0.10	0.04	0.10	0.10	1.00	0.02	0.04	0.20	0.12	0.23
exerciseangia	-0.04	-0.05	-0.03	-0.05	0.03	0.01	-0.03	0.02	1.00	-0.00	-0.02	-0.02	-0.04
oldpeak	-0.03	0.33	0.10	-0.00	-0.03	0.02	0.04	0.04	-0.00	1.00	0.21	0.09	0.10
slope	-0.02	-0.05	0.47	0.38	0.17	0.24	0.34	0.20	-0.02	0.21	1.00	0.53	0.80
noofmajorvessels	0.03	0.03	0.29	0.26	0.13	0.18	0.20	0.12	-0.02	0.09	0.53	1.00	0.49
target	0.01	0.02	0.55	0.48	0.20	0.30	0.43	0.23	-0.04	0.10	0.80	0.49	1.00

Fig. 7. Correlation matrix of clinical features and the target.

C. *Data Preprocessing*

In this stage, the clinical dataset was preprocessed to eliminate errors and optimize the performance of the ML models. The dataset has several other numerical and categorical attributes, such as type of chest pain, fasting blood sugar, resting electrocardiogram result slope, and number of major vessels. Categorical attributes were encoded into numeric values to facilitate the use of the data by ML algorithms. This transformation retains the categorical value distribution with enhanced model learning and prediction.

Z-score normalization was used to achieve equal weight for all features in the input vectors. This feature scaling algorithm standardizes the data so that features have zero-mean and unit-variance, eliminating numerical-feature discrepancies, such as one feature having a larger numerical range than others. For a feature x, its standardized value x' is given by

$$x' = \frac{x-\mu}{\sigma} \tag{1}$$

where μ is the feature mean and σ is the standard deviation. Standard scaling is especially significant for learning algorithms that are distance-based and gradient-based, as it helps in the convergence speed, numerical stability, and model performance overall.

D. *Random Forest (RF) Classifier*

The RF classifier, with a specific implementation, was the best-performing model in the heart disease prediction task, manifesting good generalization ability. RF is an ML technique in which several DTs are constructed, and their predictions are aggregated to generate the overall prediction. Let the training data be expressed as:

$$D = \{(x_{i,}y_i)\}_{i=1}^{M} \tag{2}$$

where $x_i \in R^d$ denotes each patient record and $y_i \in \{0,1\}$ represents a class label indicating whether the patient has heart disease or not. Every tree $h_j(x)$ is trained on a bootstrapped sample of the dataset D, and, at each node split, from a random subset of features, one chooses how to split that node. The quality of a split is measured by the Gini Index, which is given by:

$$Gini = 1 - \sum_{k=1}^{K} p_k^2 \tag{3}$$

where p_k denotes the probability of class k at a node. The final predicted class $\hat{y}$ of a new patient record x can be achieved by majority voting of all N decision trees as:

$$\hat{y} = \arg\max_{c\in\{0,1\}} \sum_{j=1}^{N} \mathbb{I}\left(h_j(x) = c\right) \tag{4}$$

where $\mathbb{I}(.)$ is the indicator function. The RF classifier diminishes variance by combining many DTs, detecting nonlinear associations between clinical features, and models simultaneously those complex interactions in the data. This strength also makes it well-suited for robust prediction of heart disease.

E. *Model Evaluation Metrics*

To evaluate the performance of the heart disease prediction model, some standard metrics were calculated, namely Accuracy, Precision, Recall, and F1-score. These metrics are computed from a confusion matrix made of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) predictions. Performance was estimated using stratified 5-fold cross-validation to calculate all evaluation metrics while ensuring a balanced representation of the data within each fold.

Accuracy is the ratio of correctly predicted samples and is used to measure of the overall correctness of the classification model.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{5}$$

Precision determines how correct positive predictions are, indicating how many predicted disease cases are true ones. Precision is high when the number of false positives is low, implying that the model raises fewer alarms and provides more reliable positive predictions.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{6}$$

Recall quantifies how well a model identifies positive cases, those suffering from heart disease. A larger recall shows that most actual heart disease patients are correctly classified as positive, and fewer patients are misdiagnosed, which is important in medical prediction systems.

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{7}$$

F1-score is the harmonic mean of Precision and Recall, being an important metric when aiming to minimize both false positives and false negatives.

$$\mathrm{F1-score} = 2 \times \frac{\mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}} \quad (8)$$

III. RESULTS AND ANALYSIS

This section describes the empirical evaluation of the proposed ML approach for heart disease prediction. The predictive power and reliability of various classification models were evaluated with standard performance indices and metrics. Evaluation was carried out to investigate the best model, which is superior in terms of classification performance and stability. The results offer some practical evidence on the efficacy of this approach in a proper and reliable prediction of heart disease.

Figure 8 shows the confusion matrices of all considered ML classifiers, demonstrating their classification performance in correctly and incorrectly classifying subjects with heart disease or without. The RF classifier performed best among all tested methods, leading to maximum accuracy over cases and minimum misclassification over errors, which means high predictive power. SVC and GNB also demonstrate good discrimination ability with significantly low false predictions. LR, DT, and Bagging classifiers obtained competitive performances with many correctly classified cases and a low error rate. KNN and ET classifiers exhibit higher misclassification rates, especially when discriminating borderline cases. The confusion matrices clearly show that the RF offers well-balanced and accurate classification, which can be the best option for consistent heart disease prediction.

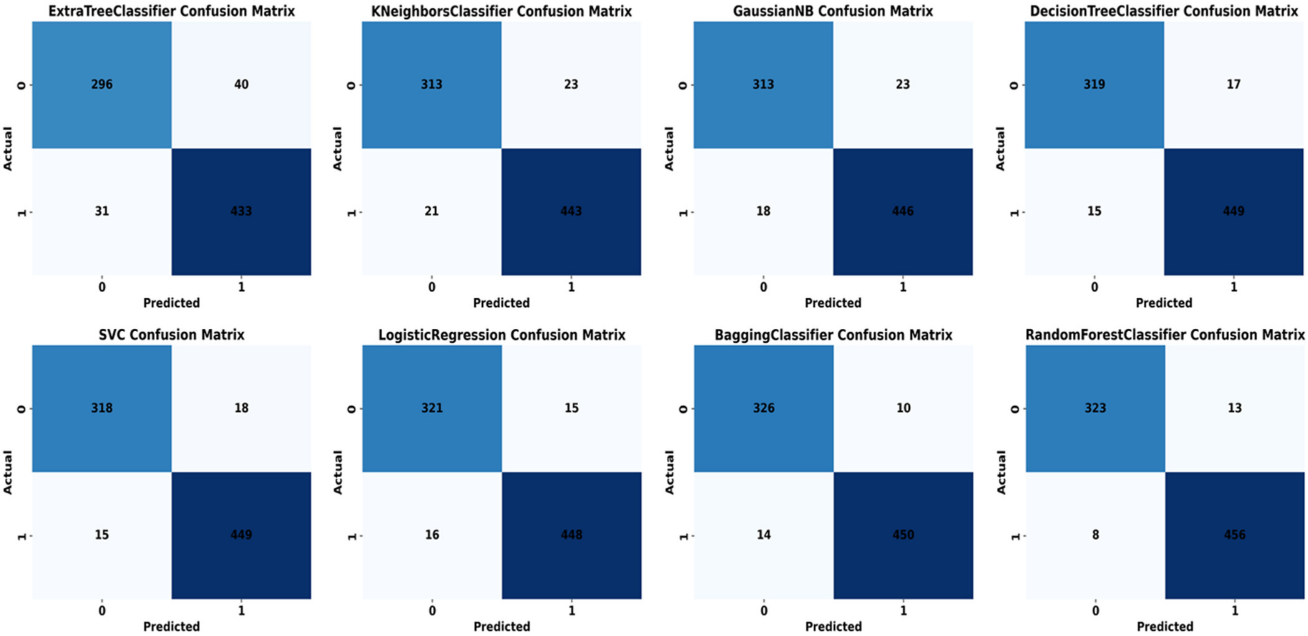


Fig. 8. Confusion matrices of ML models on heart disease prediction.

Figure 9 presents a Receiver Operating Characteristic (ROC) curve comparison of ML classifiers, demonstrating the trade-off between TP (sensitivity) and FP (1–specificity) rates for different classification thresholds. The nearer the curve to the top-left corner, the better the classification performance and the more discriminative ability. The AUC values indicate that both RF (AUC = 0.999) and LR (AUC = 0.998) achieve the best performance, with the Bagging (AUC = 0.992) and SVC (AUC = 0.997), respectively, closely following behind. Although DT and GNB show good predictive power, KNN and ET have comparatively lower (but still reasonable) performance. This figure confirms that the RF model can provide higher classification accuracy along with reliable heart disease prediction.

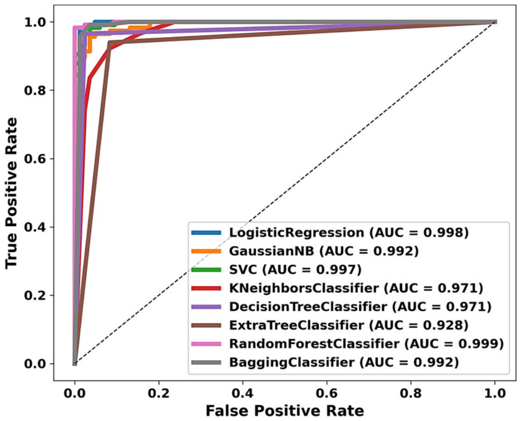


Fig. 9. ROC-curve comparison of ML classifiers for heart disease prediction.

Figure 10 compares the performances of different ML classifiers in predicting heart disease using a Precision-Recall curve, which is an important performance metric for medical datasets when accurate prediction of positive cases is more crucial than negative ones. The closer the curve is to the top-right corner, the better the performance, indicating high Precision and high Recall. RF demonstrated the maximum Area Under the Curve (AUC = 0.999), which means it correctly identifies heart disease cases with minimum FPs. Both LR and Bagging performed well, followed by SVC and GNB. On the other hand, KNN and ET demonstrated a relatively lower Precision-Recall performance.

Figure 11 compares the classification accuracy attained by the ML models. All classifiers achieved high classification accuracy, suggesting good learning of the clinical data. The best performance was observed by RF, which achieved 97.5%, revealing better predictivity and generalization ability. Bagging also displayed high accuracy (96.25%), and LR followed (96.12%). The SVC and DT models were quite comparable in performance (95.88%, 96.12%), followed by the GNB model (94.88%). KNN and ET demonstrated less accuracy than the other models.

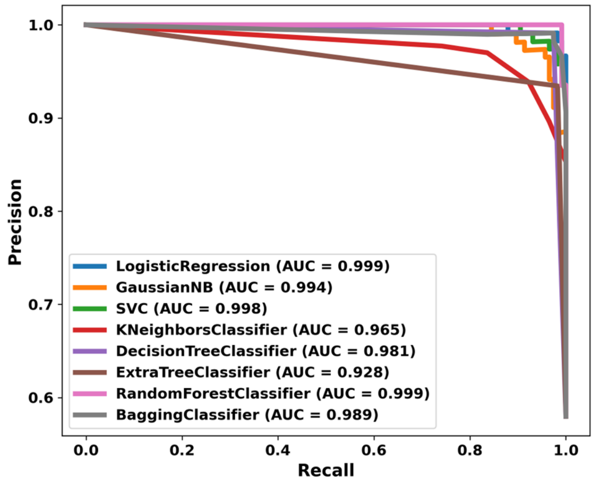


Fig. 10. Comparison of ML classifiers for heart disease prediction based on the Precision-Recall curve.

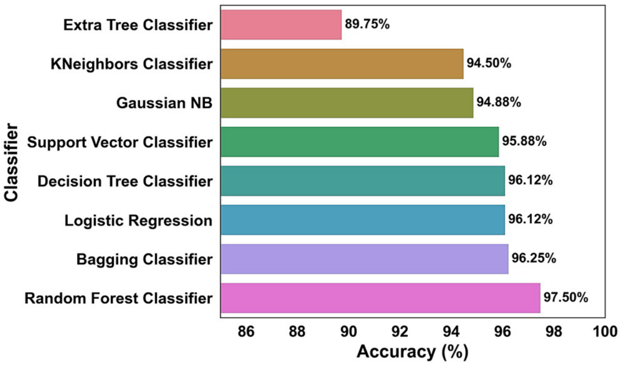


Fig. 11. Accuracy comparison.

Figure 12 shows the precision comparison of the ML models. The best performance was obtained with RF (97.51%), indicating its ability to correctly diagnose heart disease cases with few false alarms. Bagging also showed high precision (96.30%), followed by LR (96.15%). SVC and DT reported competitive precision (95.96%, 96.15%), while the GNB followed (94.92%). On the other hand, relatively low precision scores were achieved by ET and KNN.

Figure 13 compares the Recall performance of the ML classifiers for heart disease prediction. RF achieved the best recall (97.50%), which suggests that it is very good at correctly predicting almost all cases with heart disease. Bagging yielded the second-highest recall (96.25%), closely matched by LR (96.12%). SVC, DT, and GNB had recalls of 95.88%, 96.12%, and 94.88%, indicating good precision. KNN and ET demonstrated lower Recall values.

Figure 14 shows the F1-scores achieved by the ML models. RF performed the best with an F1-score of 97.50%, representing a strong trade-off between identifying heart disease cases and minimizing FPs. Bagging was also impressive at 96.25%, while LR had achieved 96.12%. Both SVC and DT classifiers achieved consistent performance (95.87%, 96.12%), while moderate performance was obtained by GNB and KNN, and ET obtained the least F1-score. These results reveal that RF yields more balanced and robust prediction results in heart disease prediction.

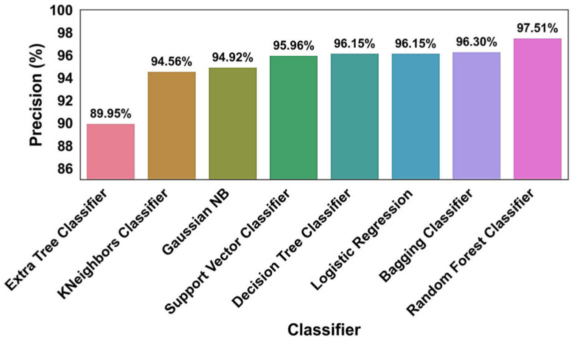


Fig. 12. Comparison of ML models on Precision.

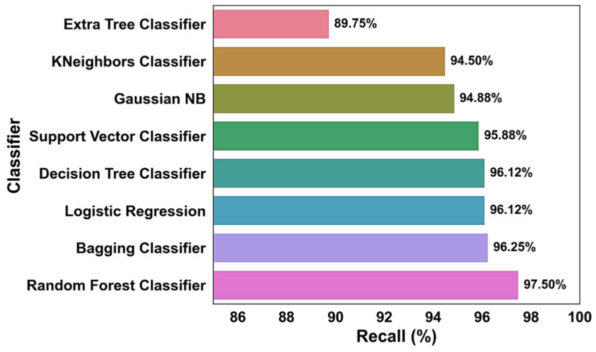


Fig. 13. Comparison of ML models on Recall.

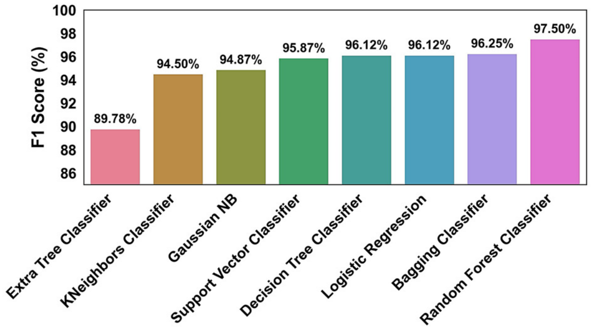


Fig. 14. Comparison of ML models on F1-score.

Figure 15 shows the feature importance values obtained from the RF classifier, representing the relative significance of each clinical attribute in the prediction of heart disease. The feature with the highest importance score is slope (S), which suggests that the value S has during the ST-segment induced by exercise dictates the presence of heart disease. Chest pain and resting blood pressure ranked second and third, respectively, in prediction importance. Moderate importance can be seen for features such as the number of major vessels, serum cholesterol, and resting electrocardiographic results, which indicate that they play a supportive role in the model. Oldpeak and maximum heart rate are lower but still meaningful contributors. In contrast, a lower importance score was obtained for other features such as gender, age, fasting blood sugar, and exercise-induced angina, indicating a lesser contribution to the final prediction. This figure provides evidence that RF can extract clinically relevant features from the data, enhancing the interpretability and reliability of heart disease prediction.

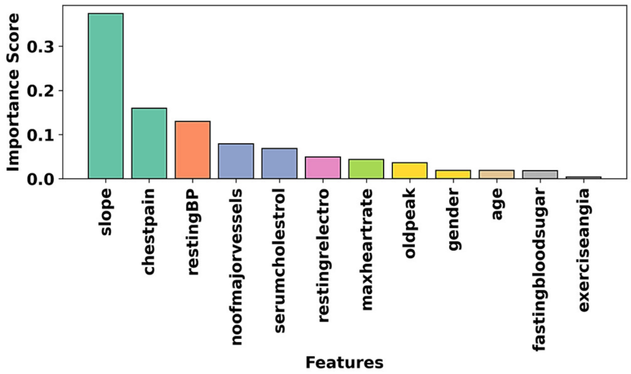


Fig. 15. Examining feature importance with RF for heart disease prediction.

IV. CONCLUSION

This work presented the design of an ML-based data-driven model that accurately and reliably predicts heart disease based on clinical attributes. The most accurate model for early detection of heart disease was chosen for classification. The experiments showed that ensemble-based methods surpassed single classifiers in both prediction performance and stability. The RF classifier performed best compared to the other models assessed, achieving the best classification performance and strong generalization with Accuracy, Precision, Recall, and F1-score of 97.50%, 97.51%, 97.50%, and 97.50%. RF successfully captured complex nonlinear relationships between clinical features and minimized misclassification errors. The negligible difference between training and validation results demonstrated the model's robustness and resistance to overfitting.

The findings indicate that the proposed RF-based system shows a good potential for decision support in early diagnosis of heart disease and determination of clinical risk. Early and accurate prediction can help doctors diagnose early, take preventive measures, and develop a better treatment plan to ultimately reduce heart disease mortality. In general, the proposed model has promising prospects and practical value in real-world healthcare applications. In the future, the model can be improved by integrating more complete and diverse clinical datasets for better predictive power and generalization. Furthermore, the implementation of transparent AI and deep learning techniques could be beneficial for model interpretability and more sophisticated clinical decision support.

ACKNOWLEDGEMENT

This research project was funded by the University of Technology and Applied Sciences through the Internal Research Funding Program-2025 (grant no. UTAS-Shinas-cy02-2025-007).

ETHICAL STATEMENT

The patient data were obtained from a publicly available and fully de-identified dataset [24]. Therefore, ethical approval and informed consent were not required.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest that could have influenced the results of this work.

ACKNOWLEDGMENT

Not applicable to this work.

DATA AVAILABILITY STATEMENT

The results were generated using the Cardiovascular Disease Dataset, available at [24].

REFERENCES

[1] R. Katarya and S. K. Meena, "Machine Learning Techniques for Heart Disease Prediction: A Comparative Study and Analysis," *Health and Technology*, vol. 11, no. 1, pp. 87–97, Jan. 2021, https://doi.org/10.1007/s12553-020-00505-7.

[2] D. Shah, S. Patel, and S. K. Bharti, "Heart Disease Prediction using Machine Learning Techniques," *SN Computer Science*, vol. 1, no. 6, Nov. 2020, Art. no. 345, https://doi.org/10.1007/s42979-020-00365-y.

[3] A. Srivastava and A. K. Singh, "Heart Disease Prediction using Machine Learning," in *2022 2nd International Conference on Advanced Computing and Innovative Technologies in Engineering (ICACITE)*, Apr. 2022, pp. 2633–2635, https://doi.org/10.1109/ICACITE53722.2022.9823584.

[4] C. M. Bhatt, P. Patel, T. Ghetia, and P. L. Mazzeo, "Effective Heart Disease Prediction Using Machine Learning Techniques," *Algorithms*, vol. 16, no. 2, Feb. 2023, Art. no. 88, https://doi.org/10.3390/a16020088.

[5] H. Jindal, S. Agrawal, R. Khera, R. Jain, and P. Nagrath, "Heart disease prediction using machine learning algorithms," *IOP Conference Series: Materials Science and Engineering*, vol. 1022, no. 1, Jan. 2021, Art. no. 012072, https://doi.org/10.1088/1757-899X/1022/1/012072.

[6] M. Kavitha, G. Gnaneswar, R. Dinesh, Y. R. Sai, and R. S. Suraj, "Heart Disease Prediction using Hybrid Machine Learning Model," in *2021 6th International Conference on Inventive Computation Technologies (ICICT)*, Jan. 2021, pp. 1329–1333, https://doi.org/10.1109/ICICT50816.2021.9358597.

[7] A. Garg, B. Sharma, and R. Khan, "Heart disease prediction using machine learning techniques," *IOP Conference Series: Materials Science and Engineering*, vol. 1022, no. 1, Jan. 2021, Art. no. 012046, https://doi.org/10.1088/1757-899X/1022/1/012046.

[8] A. K. Shukla and V. S. Kumar, "Cloud Computing with Artificial Intelligence Techniques for Effective Disease Detection," *International Journal of Data Informatics and Intelligent Computing*, vol. 2, no. 1, pp. 32–41, Mar. 2023, https://doi.org/10.59461/ijdiic.v2i1.45.

[9] Kumar Mohan, Yousuf Fadhil Salim AL-Mamari, and Mohammed Ahmed Mohammed AL-Najadi, "Techniques of Machine Learning for Detecting Heart Failure," *International Journal of Data Informatics and Intelligent Computing*, vol. 2, no. 2, pp. 47–54, June 2023, https://doi.org/10.59461/ijdiic.v2i2.62.

[10] J. Hu *et al.*, "Explainable AI models for predicting drop coalescence in microfluidics device," *Chemical Engineering Journal*, vol. 481, Feb. 2024, Art. no. 148465, https://doi.org/10.1016/j.cej.2023.148465.

[11] L. Riyaz, M. A. Butt, M. Zaman, and O. Ayob, "Heart Disease Prediction Using Machine Learning Techniques: A Quantitative Review," in *International Conference on Innovative Computing and Communications*, vol. 1394, A. Khanna, D. Gupta, S. Bhattacharyya, A. E. Hassanien, S. Anand, and A. Jaiswal, Eds. Springer Singapore, 2022, pp. 81–94.

[12] D. E. Salhi, A. Tari, and M. T. Kechadi, "Using Machine Learning for Heart Disease Prediction," in *Advances in Computing Systems and Applications*, vol. 199, M. R. Senouci, M. E. Y. Boudaren, F. Sebbak, and M. Mataoui, Eds. Springer International Publishing, 2021, pp. 70–81.

[13] N. Chandrasekhar and S. Peddakrishna, "Enhancing Heart Disease Prediction Accuracy through Machine Learning Techniques and Optimization," *Processes*, vol. 11, no. 4, Apr. 2023, Art. no. 1210, https://doi.org/10.3390/pr11041210.

[14] S. R. Reddy and G. V. Murthy, "Cardiovascular Disease Prediction Using Particle Swarm Optimization and Neural Network Based an Integrated Framework," *SN Computer Science*, vol. 6, no. 2, Feb. 2025, Art. no. 186, https://doi.org/10.1007/s42979-025-03723-w.

[15] P. Shao *et al.*, "Explainability of AI predictive models for drop coalescence in microfluidics channels: Experimental validation," *Chemical Engineering Science*, vol. 320, Jan. 2026, Art. no. 122559, https://doi.org/10.1016/j.ces.2025.122559.

[16] K. Arumugam, M. Naved, P. P. Shinde, O. Leiva-Chauca, A. Huaman-Osorio, and T. Gonzales-Yanac, "Multiple disease prediction using Machine learning algorithms," *Materials Today: Proceedings*, vol. 80, pp. 3682–3685, 2023, https://doi.org/10.1016/j.matpr.2021.07.361.

[17] M. Diwakar, A. Tripathi, K. Joshi, M. Memoria, P. Singh, and N. Kumar, "Latest trends on heart disease prediction using machine learning and image fusion," *Materials Today: Proceedings*, vol. 37, pp. 3213–3218, 2021, https://doi.org/10.1016/j.matpr.2020.09.078.

[18] A. Saboor, M. Usman, S. Ali, A. Samad, M. F. Abrar, and N. Ullah, "A Method for Improving Prediction of Human Heart Disease Using Machine Learning Algorithms," *Mobile Information Systems*, vol. 2022, pp. 1–9, Mar. 2022, https://doi.org/10.1155/2022/1410169.

[19] J. Patel, A. A. Khaked, J. Patel, and J. Patel, "Heart Disease Prediction Using Machine Learning," in *Proceedings of Second International Conference on Computing, Communications, and Cyber-Security*, vol. 203, P. K. Singh, S. T. Wierzchoń, S. Tanwar, M. Ganzha, and J. J. P. C. Rodrigues, Eds. 2021, pp. 653–665.

[20] D. AbdElminaam, M. Radwan, N. Mohamed Abdelrahman, H. Wael Kamal, A. Khaled Abdelmonem Elewa, and A. Moataz Mohamed, "MLHeartDisPrediction: Heart Disease Prediction using Machine Learning," *Journal of Computing and Communication*, vol. 2, no. 1, pp. 50–65, Jan. 2023, https://doi.org/10.21608/jocc.2023.282098.

[21] M. Rahardi, B. P. Asaddulloh, A. Aminuddin, F. F. Abdulloh, I. Saifudin, and F. P. Kusumawijaya, "Optimizing Machine Learning Models for Class Imbalance in Heart Disease Prediction," *Engineering, Technology & Applied Science Research*, vol. 15, no. 3, pp. 23599–23604, June 2025, https://doi.org/10.48084/etasr.10407.

[22] M. K. Hossen, "A Study of Large-eddy Simulation using Statistical and Machine Learning Techniques," *Journal of Engineering Advancements*, pp. 96–103, Sept. 2022, https://doi.org/10.38032/jea.2022.03.004.

[23] T. Roopa and G. D. Ramanjinappa, "Heart Disease Predictive Modeling with XGBoost and SMOTE-Driven Class Imbalance Mitigation," *Engineering, Technology & Applied Science Research*, vol. 15, no. 6, pp. 29914–29918, Dec. 2025, https://doi.org/10.48084/etasr.14301.

[24] B. P. Doppala, "Cardiovascular_Disease_Dataset." Mendeley, Apr. 16, 2021, https://doi.org/10.17632/DZZ48MVJHT.1.