

Research Article

Supervised Machine Learning-Based Cardiovascular Disease Analysis and Prediction

M. D. Amzad Hossen ¹, **Tahia Tazin** ¹, **Sumiaya Khan** ¹, **Evan Alam** ¹,
Hossain Ahmed Sojib ¹, **Mohammad Monirujjaman Khan** ¹, and **Abdulmajeed Alsufyani** ²

¹Department of Electrical and Computer Engineering, North South University, Bashundhara, Dhaka 1229, Bangladesh

²Department of Computer Science, College of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia

Correspondence should be addressed to Mohammad Monirujjaman Khan; monirujjaman.khan@northsouth.edu

Received 10 November 2021; Revised 23 November 2021; Accepted 26 November 2021; Published 10 December 2021

Academic Editor: Ewa Rak

Copyright © 2021 M. D. Amzad Hossen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Cardiovascular illness, often commonly known as heart disease, encompasses a variety of diseases that affect the heart and has been the leading cause of mortality globally in recent decades. It is associated with numerous risks for heart disease and a requirement of the moment to get accurate, trustworthy, and reasonable methods to establish an early diagnosis in order to accomplish early disease treatment. In the healthcare sector, data analysis is a widely utilized method for processing massive amounts of data. Researchers use a variety of statistical and machine learning methods to evaluate massive amounts of complicated medical data, assisting healthcare practitioners in predicting cardiac disease. This study covers many aspects of cardiac illness, as well as a model based on supervised learning techniques such as Random Forest (RF), Decision Tree (DT), and Logistic Regression (LR). It makes use of an existing dataset from the UCI Cleveland database of heart disease patients. There are 303 occurrences and 76 characteristics in the collection. Only 14 of these 76 characteristics are evaluated for testing, which is necessary to validate the performance of various methods. The purpose of this study is to forecast the likelihood of individuals getting heart disease. The findings show that logistic regression achieves the best accuracy score (92.10%).

1. Introduction

It is difficult to diagnose cardiac disease due to the presence of many health problems such as diabetes, high blood pressure, excessive cholesterol, and an irregular pulse rate. Numerous data analysis and neural network methods have been used to determine the severity of cardiac disease in people. The severity of illness is categorized using a variety of techniques, including the K-Nearest Neighbor (KNN) algorithm, DT, Genetic Algorithm (GA), and the Naive Bayes (NB) algorithm [1, 2]. Due to the complexity of cardiac disease, it must be treated with caution. Failure to do so may have a detrimental effect on the heart or result in early death. Medical science and statistical perspectives are utilized to identify different types of metabolic disorders. Data analysis with categorization is essential for heart disease prediction

and data research. Additionally, we have seen decision trees used to predict the accuracy of heart disease-related events [3]. Numerous approaches to knowledge abstraction have been employed in conjunction with well-established data mining techniques for heart disease diagnosis. Numerous analyses have been conducted in this study to develop a prediction model, not only utilizing different methods but also connecting two or more techniques. Data mining is the process of extracting needed information from massive databases in a variety of areas, including medicine, business, and education. Machine learning (ML) is one of the fields of artificial intelligence (AI) that is advancing at a breakneck pace. These algorithms are capable of analyzing massive amounts of data from a variety of areas, one of which being the medical field. It is a replacement for the conventional prediction modeling method that uses a computer to acquire

knowledge of complicated and nonlinear interactions between many variables by minimizing the difference between anticipated and actual results [4]. Data mining is the process of sifting through massive datasets in order to extract critical decision-making information from a collection of historical records for future study. The medical profession is replete with patient data. This data must be analyzed using a variety of machine learning techniques. Healthcare experts analyze these data in order to make appropriate diagnostic decisions. Through analysis, medical data mining using classification algorithms offers therapeutic assistance. It evaluates methods for classifying patients' risk of developing heart disease [5].

Several studies have been performed, and numerous machine learning models have been deployed, all with the goal of classifying and forecasting heart disease diagnoses. ANNs were developed to achieve the greatest prediction accuracy possible in the medical sector [6]. ANNs are used to forecast cardiac disease via back propagation multilayer perceptron (MLP). The resulting findings are compared to those of previously published models in the same area and found to be significantly improved [7]. The UCI laboratory's data on heart disease patients are utilized to identify patterns using NN, DT, Support Vector Machines (SVMs), and Naive Bayes. The performance and accuracy of various algorithms are compared. The proposed hybrid approach achieves an F-measure accuracy of 86.8%, which is comparable to other available methods [8]. The classification of Convolutional Neural Networks (CNNs) without segmentation is presented. This technique considers cardiac cycles with a variety of start locations derived from Electrocardiogram (ECG) data during the training phase. CNN is capable of generating features with varying locations throughout the patient's testing phase [9, 10]. Previously, a significant quantity of data produced by the medical sector was not used properly. The novel methods described here reduce the cost and enhance the accuracy of heart disease prediction in a simple and efficient manner. The numerous research approaches examined in this study for the prediction and classification of heart disease utilizing ML and deep learning (DL) techniques are very accurate in proving these methods' effectiveness [11, 12].

Golande et al. investigated a variety of machine learning methods that may be used to classify cardiac disease. Research was conducted to evaluate the accuracy of DT, KNN, and K-Means algorithms that may be utilized for classification [13]. This study indicates that DT achieves the greatest accuracy and that they may be made more efficient via the use of a mix of various methods and parameter tweaking. Nagamani et al. [14] developed a system that combined data mining methods with the MapReduce algorithm. For the 45 instances in the testing set, the accuracy achieved in this study was higher than the accuracy obtained using a typical fuzzy artificial neural network. Due to the usage of dynamic schema and linear scaling, the accuracy of the method was increased in this case. Alotaibi developed a machine learning model that compares five distinct methods [15]. A rapid miner was employed, which provided a better level of accuracy compared to MATLAB

and Weka. In this study, the classification algorithms DT, LR, NB, and SVM were compared for accuracy. The decision tree algorithm was the most precise. Repaka et al. developed a system [16] that combines NB (Naive Bayes) methods for dataset categorization and AES (Advanced Encryption Standard) for secure data transmission for illness prediction. Thomas and Princy conducted a study comparing several categorization algorithms used for heart disease prediction. The classification methods utilized were Naive Bayes, KNN, DT, and Neural Network, and the accuracy of the classifiers was evaluated over a range of attribute counts [17]. Lutimath et al. used Naive Bayes classification and SVM to predict cardiac disease. The performance metrics utilized in the study are the Mean Absolute Error, the Sum of Squared Error, and the Root Mean Square Error. It has been shown that SVM outperforms Naive Bayes in terms of accuracy [18]. The authors proposed an RNN-based prediction of the risk of depression based on ECG [19].

Cardiac disease may be cured if diagnosed early, but this is not always the case. We must learn more about a few heart illness markers if we want to avert major harm. The analysis of data from these indices and the use of three machine learning classification algorithms to predict cardiac disease are the primary goals of this project. The strategy with the greatest accuracy rate will be chosen.

After analyzing the aforementioned studies, the main objective of the proposed system was to develop a computer-aided diagnostic system using the inputs listed in Table 1. We compared the accuracy, precision, recall, and F1-scores of three classification algorithms. DT, RF, and LR are found to be the best classification methods for heart disease prediction. Because we employed multiple methods and reached 92% accuracy, which is greater than the prior publications, this research has a novel feature.

The remaining paper is divided into three parts. The methodology and methods are presented in Section 2, the findings and analysis are presented in Section 3, and the conclusion and future scope are presented in Section 4.

2. Methodology and Methods

This section includes details on the methods and materials utilized, as well as a dataset description, a schematic diagram, machine learning algorithms, and evaluation matrices.

2.1. Dataset. The Heart Disease dataset was utilized, which is a compilation of four distinct databases, but only the UCI Cleveland dataset has been used [20]. This database has 76 characteristics in total, but all published studies use just a subset of just 14 features [21]. As a result, for our study, we utilized the previously processed UCI Cleveland dataset accessible on the "Kaggle" website. Table 1 gives a detailed explanation of the 14 characteristics utilized in the proposed study.

There are a total of 165 cardiac disease and 138 non-cardiac disease datasets available in the target column. Figure 1 shows the visualization of the target column.

TABLE 1: Dataset feature’s information.

Feature	Explanation
Age	Age of the patient in completed years
Sex	Gender of the patient
Cp	Chest pain is classified into four types: (1) conventional angina, (2) unusual angina, (3) nonanginal pain, and (4) asymptomatic
Trestbps	Blood pressure in the resting state
Chol	Cholesterol in the blood
FBS	Fasting blood sugar levels
Resting	The results of an ECG taken while at rest are represented by three separate values: normal condition is represented by value 0, abnormality in the ST-T wave is represented by value 1 (which may include T-wave inversions and/or depression or an elevation of ST of >0.05 mV), and any possibility or certainty of LV hypertrophy per Estes’ criteria is indicated by value 2
Thali	The achievement of the maximal heart rate
Old peak	In compared to a resting state, exercise causes ST depression
Exang	Exercise-induced angina
Slope	The ST segment is depicted in three values based on the slope during peak exercise: (1) level, (2) flat, and (3) downsloping
Ca	Fluorescence imaging colored main vessels numbered from 0 to 3
Thal	The condition of the heart is represented by three distinct numerical values. Normal defects are numbered 3, fixed defects are numbered 6, and reversible defects are numbered 7
Target	It is the dataset’s last column. Column is a class or label. It denotes the number of classes in the dataset. This dataset has a binary categorization, which means it has two classifications (0, 1). In the class, “0” indicates that there is a low risk of heart illness, but “1” indicates that there is a high risk of heart disease. The value “0” or “1” is determined by the other 13 attributes

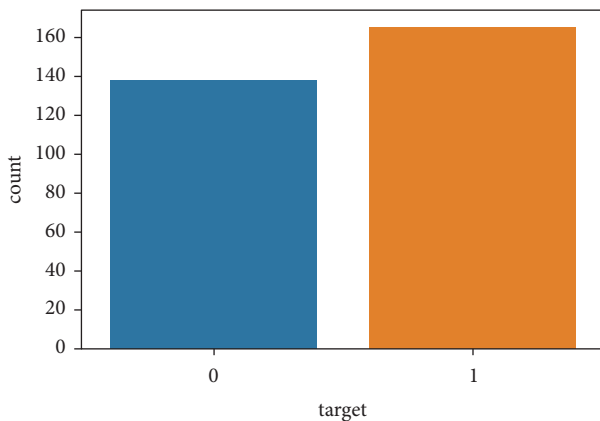


FIGURE 1: Visualization of the target column.

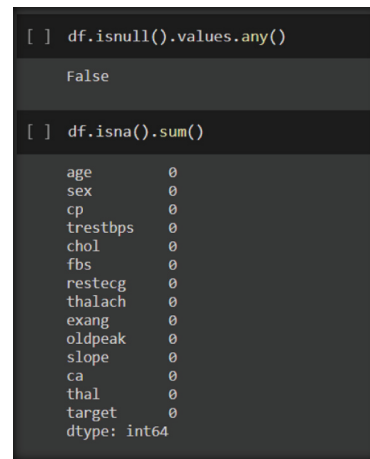


FIGURE 2: Zero missing data.

If similar and null data are not verified and handled, the model’s generality suffers. There is a chance that duplicates will appear in both the test and training datasets if duplicates are not handled effectively. As a result, during the pre-processing phase, all duplicate data were eliminated from the dataset. This dataset has no missing data, as shown in Figure 2.

Because the dataset contains no missing data, Figure 2 displays 0 values in all of the dataset’s attributes.

2.2. Schematic Diagram of the System. The proposed study indicated heart disease by examining the three classification methods listed above and carrying out performance analysis. The goal of this research is to accurately predict whether or not a patient has heart disease. The input values from the patient’s health report are entered by the health professional. The data are incorporated into a model that forecasts the chance of developing heart disease. Figure 3 depicts the system’s schematic diagram.

The properties listed in Table 1 are used as inputs for classification methods including Random Forest, Decision Tree, and Logistic Regression. The input dataset is divided into 80% of the training dataset and 20% of the test dataset. A training dataset is a collection of data that are being used to train a model. The testing dataset is also used to evaluate the trained model’s performance. The performance of each method is generated and analyzed using a variety of measures, including accuracy, precision, recall, and F1-scores, as discussed below.

2.3. Machine Learning Algorithms. Classification and regression techniques based on Random Forest are utilized. It constructs a tree for the data and then makes predictions using that tree. The RF technique is capable of processing enormous datasets and producing the same result even when substantial portions of the record values are missing. The decision tree’s produced samples may be stored and used on

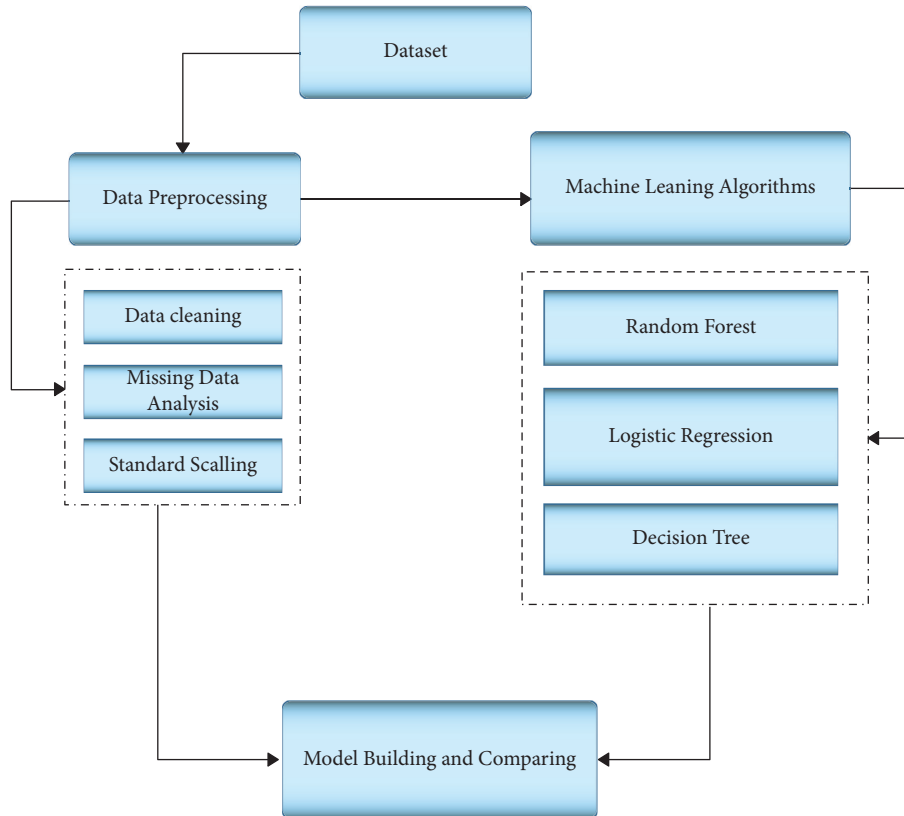


FIGURE 3: System's schematic diagram.

additional data. There are two steps in generating a random forest: first, generating a random forest and, second, using the Random Forest classifier built in the first stage, making a prediction. Figure 4 shows the schematic diagram of the Random Forest algorithm.

Also, RF is a decision tree-based method. After combining numerous separate decision trees, it is generally more accurate and reliable than a single tree. The random selection of samples and features, as well as the integration procedures, gives a Random Forest an edge over a DT. While the former resists overfitting better, the latter is more accurate. Random Forest uses the DT as the bagging model.

The DT method is represented visually as a flowchart, with the central node representing the dataset's properties and the outside branches representing the result. Decision trees are selected because they are quick, dependable, and simple to read and require little data preparation. In a DT, the prediction of the class label begins at the tree's root. The root attribute's value is compared to the record's attribute. The matching branch is explored for that value, and a move is made to another node based on the outcome of the comparison. Figure 5 shows the schematic diagram of the decision tree algorithm.

Strategic splits have a big influence on a decision tree's accuracy. They use different decision criteria. The development of subnodes increases their homogeneity. Because the target variable grows, the node's purity rises. A DT is simple to grasp and can handle both numerical and categorical data.

LR is a statistical approach that is often used to solve issues involving binary classification. Rather than fitting a straight line or hyperplane, logistic regression employs the logistic function to constrain the output of a linear equation to the range of 0 to 1. Due to the presence of 13 independent variables, logistic regression is well suited for categorization. Figure 6 shows the schematic diagram of the logistic regression algorithm.

2.4. Block Diagram of the Confusion Matrix. A confusion matrix is a technique for describing the performance of a classification system. The number of correct and incorrect predictions is summed and denoted by count values. This is the key to the misunderstanding matrix. The block diagram of the confusion matrix is shown in Figure 7.

It elucidates not only the errors made by the classifier but also the kind of faults committed. The expected row and predicted column for a class include the total number of correct predictions. Similarly, the expected row and projected column for a class value include the total number of incorrect guesses.

3. Result and Data Analysis

This section discusses the capabilities of the models, model predictions, inquiry, and final outcomes.

3.1. Data Visualization. A histogram displays the distribution of recurrences with infinite classes. It is a region outline composed of shapes with bases at class border spans and regions

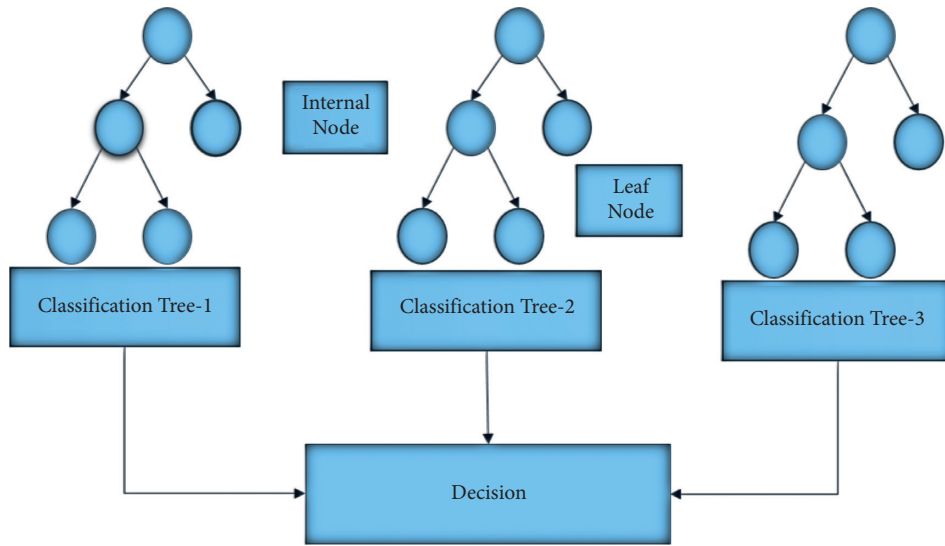


FIGURE 4: Schematic diagram of Random Forest.

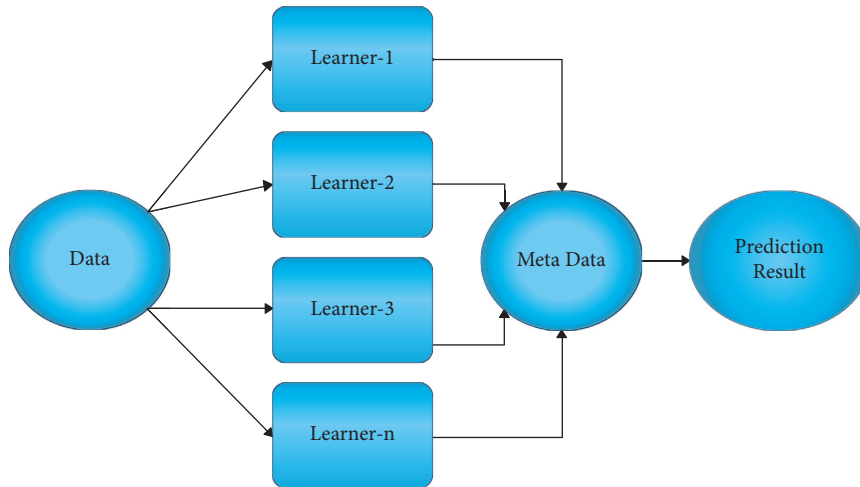


FIGURE 5: Schematic diagram of decision tree.

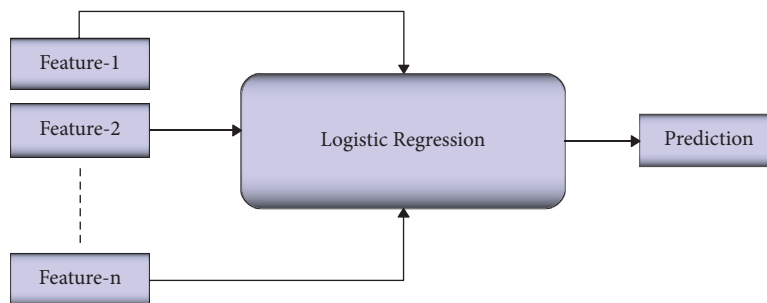


FIGURE 6: Schematic diagram of logistic regression.

proportional to the frequencies of the comparing classes. The square forms are all related because the base fills in the gaps between class boundaries. The square-form statures are proportional to the comparative class frequencies and recurrence

densities for different classes. Figure 8 depicts the distribution of age, blood pressure, cholesterol, heart rate, and old peak.

Figure 9 depicts the cardiac state of people of various ages.

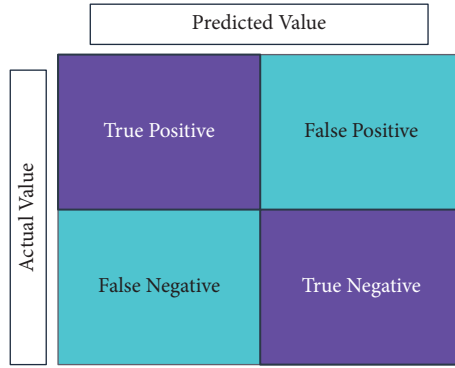


FIGURE 7: Block diagram of the confusion matrix.

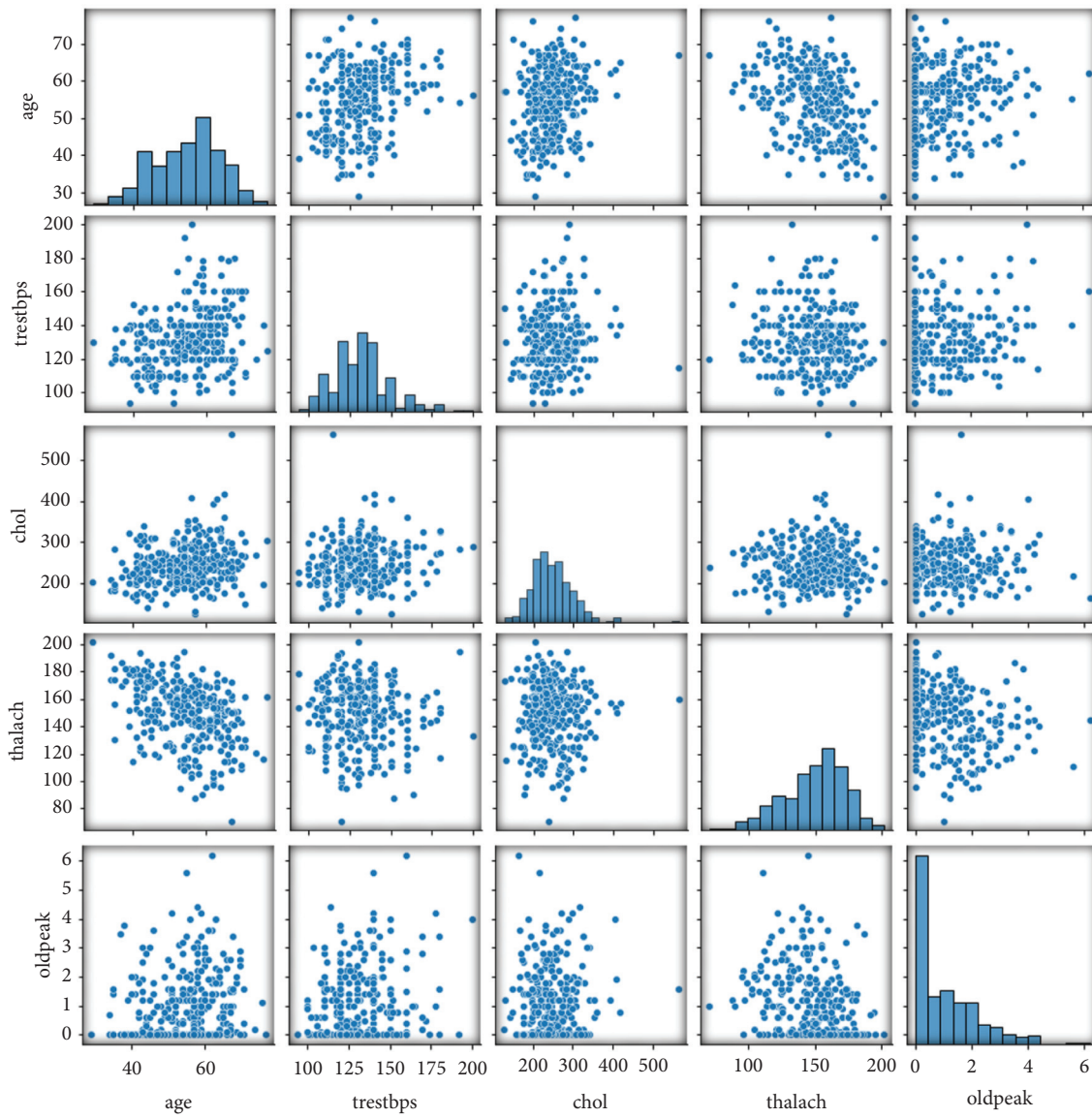


FIGURE 8: Distribution of the main features of the dataset.

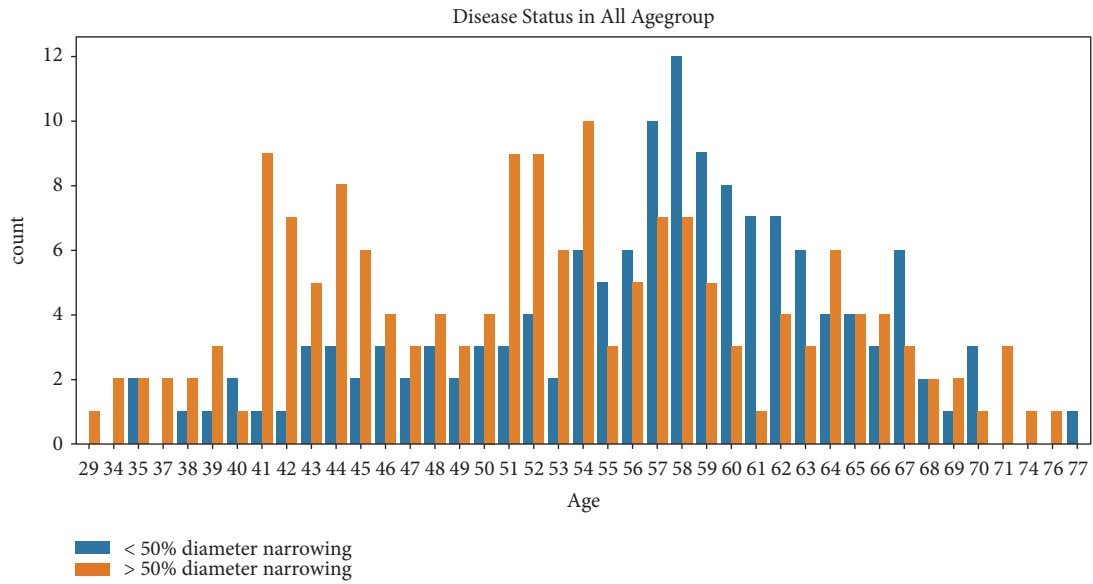


FIGURE 9: Cardiac state of people of various ages.

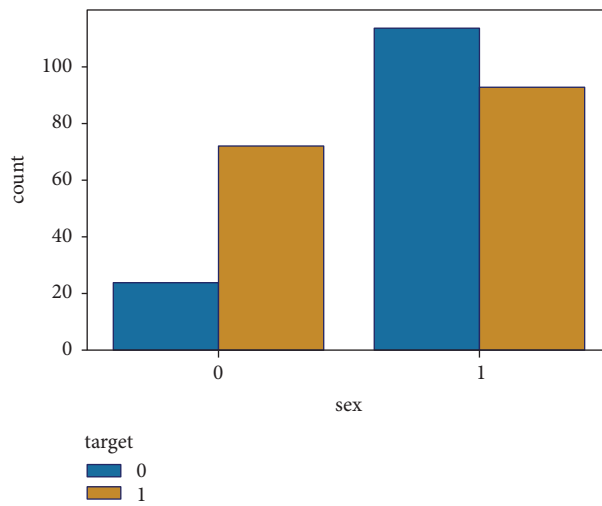


FIGURE 10: Disease status by gender.

Figure 9 reveals that an individual under the age of 35 does not have cardiovascular disease. The likelihood of developing cardiovascular disease rises with age. Target 0 indicates that the individual is healthy, whereas target 1 indicates that the individual has cardiac disease. Figure 10 depicts the illness status by gender.

The graph illustrates that a men are more likely than women to get cardiovascular disease. The probability distribution of four distinct types of characteristics is seen in Figure 11.

Figure 11 shows that the patterns of cholesterol levels, blood pressure levels, age, and maximal heart rate are not uniformly distributed. These will need to be addressed in order to prevent overfitting or underfitting of the data. In addition, cholesterol is an essential factor in the study of heart disease.

3.2. *Model Accuracy.* Table 2 shows the three different models' classification results.

According to Table 2, LR outperformed the other algorithms in terms of accuracy. The RF also performed well in terms of accuracy. The performance of the DT, on the other hand, is really low. The precision, recall, F1-score, and accuracy of the Random Forest algorithm are 77%, 87%, 82%, and 80%, respectively. Also, the precision, recall, F1-score, and accuracy of the logistic regression algorithm are 92%, 92%, 92%, and 92%, respectively.

3.3. *Confusion Matrix.* Figure 12 depicts the RF classifier's confusion matrix. This is the classifier that attained an accuracy rate of 80%.

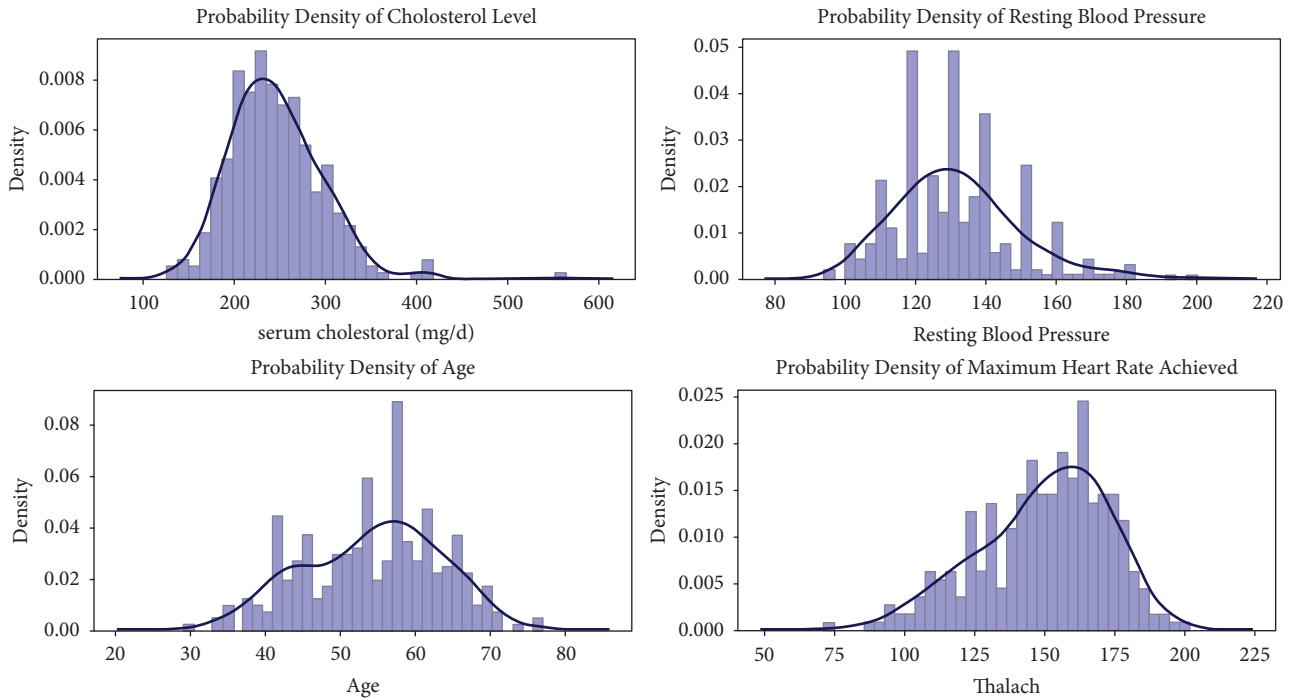


FIGURE 11: Probability density of features.

TABLE 2: Classification result of the three different models.

Model	Precision (%)	Recall (%)	F1-score (%)	Accuracy (%)
Random forest	77	87	82	80
Decision tree	71	74	72	72
Logistic regression	92	92	92	92

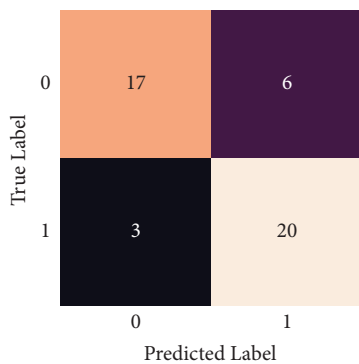


FIGURE 12: Confusion matrix of Random Forest Classifier.

Figure 12 illustrates that the FR classifier properly predicts 37 data points and wrongly predicts 9 data points.

Figure 13 depicts the prediction's ROC (receiver operating characteristic) curve. A Random Forest classifier has an AUC (accuracy under the curve) of 88%. The confusion matrix of the decision tree algorithm is shown in Figure 14.

Figure 14 shows the DT classifier accurately predicting 33 data points and wrongly predicting 13 data points.

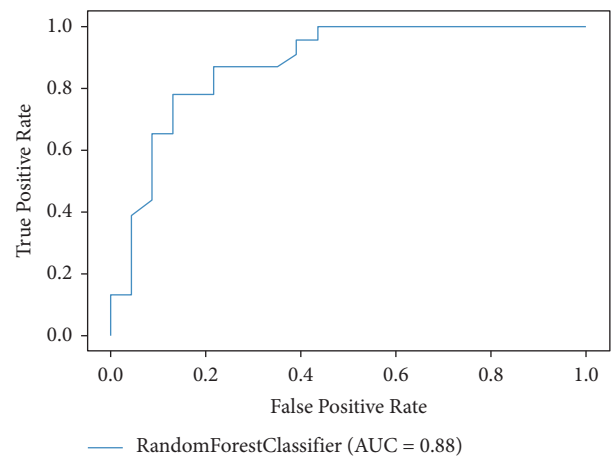


FIGURE 13: AUC for Random Forest.

Figure 15 depicts the prediction's AUC. The DT classifier has an accuracy under the curve of 72%. Figure 16 depicts the LR algorithm's confusion matrix.

Figure 16 demonstrates that the logistic regression classifier properly predicts 70 data points and wrongly predicts 6 data points.

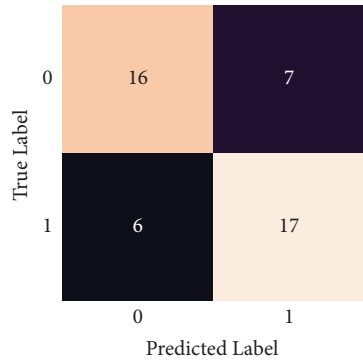


FIGURE 14: Confusion matrix of decision tree classifier.

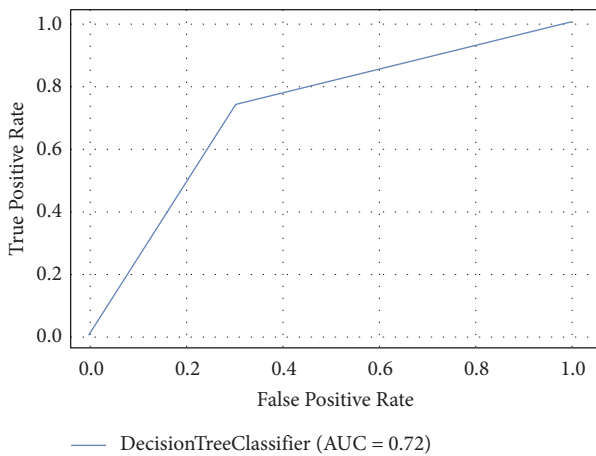


FIGURE 15: AUC of decision tree.

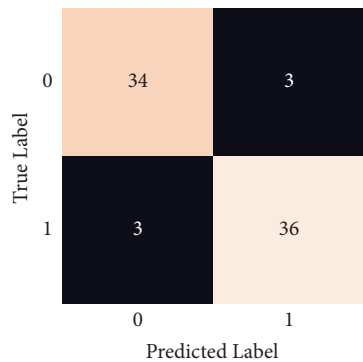


FIGURE 16: Confusion matrix of logistic regression.

Figure 17 depicts the prediction’s AUC. For the LR classifier, the accuracy under the curve is 95%. Table 3 compares the models to those in previous research articles. It clearly shows that logistic regression is the best model among the framework’s various models. It has a higher accuracy rate.

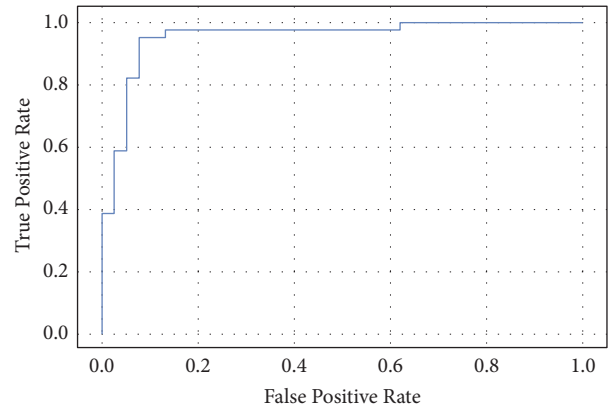


FIGURE 17: ROC curve for logistic regression.

TABLE 3: Result comparison.

Reference	Model name	Accuracy (%)	Accuracy in this study (%)
[20]	Logistic regression	75	92
[22]	Decision tree	82.5	72
[23]	Random Forest	80.3	80
[24]	Random Forest	87.6	80

4. Conclusions

Three machine learning techniques are provided in this work, and their comparative assessment is described. The goal of the article was to determine which machine learning classifier would be the most effective in predicting heart disease based on the dataset utilized. Three classifiers were built, and their results were compared. Some of the comparison approaches used include the confusion matrix, accuracy, specificity, and sensitivity. For the 14 variables in the sample, the LR classifier performed admirably in the ML approach. The logistic regression technique outperformed the other two classifiers employed, with an accuracy of 92%. The RF classifier had an accuracy of 80%, whereas the DT classifier had an accuracy of 72%. This idea has the potential to be a game changer in the medical field. Patients at risk of heart disease might be recognized quickly with this method, which could help to lower the rising death rate. The properties in the dataset that the prediction model is built on are not prohibitively costly to record. As a result, this kind of diagnostics may be made accessible to patients at a reasonable cost, allowing it to reach a considerably larger number of people. This kind of diagnosis will become more common in the future as machine learning algorithms improve as a result of continuous research. If additional patient information is utilized, the model may be refined and adjusted. A bigger dataset ensures more precise and accurate findings. This is critical since medical diagnosis is a very delicate problem that requires high degrees of accuracy and

precision. A web application that integrates these methods and uses a larger dataset than the one used in this study might be developed in the future. As a result, healthcare providers will be better able to predict and treat cardiac abnormalities with more precision and efficiency. This will improve the framework's reliability as well as its presentation.

Data Availability

The data utilized to support this research findings are accessible online at <https://www.kaggle.com/ronitf/heart-disease-uci>.

Conflicts of Interest

The authors declare no conflicts of interest regarding the present study.

Acknowledgments

The authors are thankful for the support from the Taif University Researchers Supporting Project (TURSP-2020/115), Taif University, Taif, Saudi Arabia.

References

- [1] M. Durairaj and V. Revathi, "Prediction of Heart Disease Using Back Propagation MLP Algorithm," *Information Communication and Embedded Systems*, vol. 4, no. 08, pp. 235–239, 2015.
- [2] M. Gandhi, "Predictions in heart disease using techniques of data mining," in *Proceedings of the International Conference on Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE)*, pp. 520–525, Noida, India, February 2015.
- [3] A. S. Abdullah, "A Data mining Model for predicting the Coronary Heart Disease using Random Forest Classifier," *International Journal of Computer Application*, vol. 22, 2012.
- [4] S. F. Weng, J. Reys, J. Kai, J. M. Garibaldi, and N. Qureshi, "Can machine learning improve cardiovascular risk prediction using routine clinical data?" *PLoS One*, vol. 12, no. 4, Article ID e0174944, 2017.
- [5] V. V. Ramalingam, A. Dandapath, and M. K. Raja, "Heart disease prediction using machine learning techniques: a survey," *International Journal of Engineering & Technology*, vol. 7, no. 2.8, pp. 684–687, 2018.
- [6] L. Baccour, "Amende d fuse d TOPSIS-VIKOR for classification (ATOVIC) applied to some UCI data sets R," *Expert Systems With Applications*, vol. 99, pp. 115–125, 2018.
- [7] R. Das, I. Turkoglu, and A. Sengur, "Expert Systems with Applications Effective diagnosis of heart disease through neural networks ensembles," *Expert Systems with Applications*, vol. 36, no. 4, pp. 7675–7680, 2009.
- [8] C. Cheng and H. Chiu, "An artificial neural network model for the evaluation of carotid artery stenting prognosis using a nationwide database," in *Proceedings of the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 2566–2569, Jeju Island, Republic of Korea, July 2017.
- [9] J. Nahar, T. Imam, and K. S. Tickle, "Expert Systems with Applications Association rule mining to detect factors which contribute to heart disease in males and females," *Expert Systems with Applications*, vol. 40, no. 4, pp. 1086–1093, 2013.
- [10] S. Zaman and R. Toufiq, "Codon based back propagation neural network approach to classify hypertension gene sequences," in *Proceedings of the 2017 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, pp. 443–446, Cox's Bazar, Bangladesh, February 2017.
- [11] D. K. Ravish, K. J. Shanthi, N. R. Shenoy, and S. Nisargh, "Heart function monitoring, prediction and prevention of heart attacks: using artificial neural networks," in *Proceedings of the 2014 International Conference on Contemporary Computing and Informatics (IC3I)*, pp. 1–6, Mysore, India, November 2014.
- [12] W. Zhang and J. Han, "Towards heart sound classification without segmentation using convolutional neural network," in *Proceedings of the 2017 Computing in Cardiology (CinC)*, pp. 1–4, Rennes, France, September 2017.
- [13] A. Golande and T. P. Kumar, "Heart disease prediction using effective machine learning techniques," *International Journal of Recent Technology and Engineering*, vol. 8, pp. 944–950, 2019.
- [14] T. Nagamani, S. Logeswari, and B. Gomathy, "Heart disease prediction using data mining with mapreduce algorithm," *International Journal Of Innovative Technology And Exploring Engineering*, vol. 8, no. 3, 2019.
- [15] F. S. Alotaibi, "Implementation of machine learning model to predict heart failure disease," (IJACSA), *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 6, 2019.
- [16] A. N. Repaka, S. D. Ravikanti, and R. G. Franklin, "Design and implementing heart disease prediction using naives bayesian," in *Proceedings of the 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, pp. 292–297, Tirunelveli, India, April 2019.
- [17] J. Thomas and R. T. Princy, "Human heart disease prediction system using data mining techniques," in *Proceedings of the 2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT)*, pp. 1–5, Nagercoil, India, March 2016.
- [18] M. N. Lutimath, C. C. Basavaraj, and S. Pol, "Prediction of heart disease using machine learning," *International Journal of Recent Technology and Engineering*, vol. 8, pp. 474–477, 2019.
- [19] S. T. Noor, S. T. Asad, and M. M. Khan, "Predicting the risk of depression based on ECG using RNN," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 1299870, 12 pages, 2021.
- [20] "Heart disease UCI," <https://www.kaggle.com/ronitf/heart-disease-uci>.
- [21] C. B. Rjeily, G. Badr, and E. Hassani, "Medical data mining for heart diseases and the future of sequential mining in medical field," in *Machine Learning Paradigms*, pp. 71–99, Springer, Cham, Switzerland, 2019.
- [22] K. Srinivas, G. Raghavendra Rao, and A. Govardhan, "Analysis of coronary heart disease and prediction of heart attack in coal mining regions using data mining techniques," in *Proceedings of the 2010 5th International Conference On Computer Science & Education*, pp. 1344–1349, IEEE, Hefei, China, August 2010.
- [23] R. Bharti, A. Khamparia, M. Shabaz, G. Dhiman, S. Pande, and P. Singh, "Prediction of heart disease using a combination of machine learning and deep learning," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 8387680, 11 pages, 2021.
- [24] G. Guidi, M. C. Pettenati, P. Melillo, and E. Iadanza, "A machine learning system to improve heart failure patient assistance," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 6, pp. 1750–1756, 2014.