

# A Comprehensive Analysis of Machine Learning Methods for Predicting Heart Disease

Manish G.S

Department of Computer Science, School of Computing Sciences  
Vels Institute of Science, Technology and Advanced  
Studies, Chennai, India.  
manishgurajala.s@gmail.com

Perumal S

Department of Computer Science, School of Computing  
Sciences Vels Institute of Science, Technology and Advanced  
Studies, Chennai, India.  
perumal.scs@velsuniv.ac.in

**Abstract**— Heart disease is a global issue that significantly impacts populations, highlighting the importance of early detection for individual well-being. Machine Learning (ML) techniques, such as those utilizing the Cleveland Heart Disease dataset, offer cost-effective methods for accurately predicting heart disease. This review focused on sharing the best practices of several ML approaches in the context of patient prognosis, diagnostics, and treatment of heart diseases. It begins with a Min-Max Normalized Data Scaling (MMNDS) to preprocess the dataset by addressing noise and missing values. Next, a Genetic-Based Crow Search Algorithm (GCSA) identifies an optimal feature subset. Finally, we employ Support Vector Machine (SVM), Naive Bayes (NB), and Random Forest Classifier (RFC) to predict heart disease based on clinical parameters. Furthermore, the proposed classification methods provide excellent results in identifying early heart diseases with performance metrics such as error rate, precision, recall, F1-score, and accuracy.

**Keywords**—Heart disease, Cleveland Heart disease dataset, MMNDS, GCSA, ML, RFC, SVM and Classification.

## I. INTRODUCTION

One of the most common chronic illnesses that affects a large number of people worldwide is heart disease. Early and efficient identification of heart illness is essential in healthcare, particularly in cardiology. Breathlessness, muscular weakness, and swelling feet are frequent symptoms of HD [1]. However, current methods for diagnosing heart disease are not very good at detecting the disease early for a few reasons, such as accuracy and the time required. However, the reality is that living a healthy lifestyle may significantly decrease the risk of it. Angina and myocardial infarction are examples of cardiovascular disorders, as are coronary heart disease, wherein the coronary arteries get coated with a waxy substance known as plaque. Given that heart attacks continue to be one of the world's leading causes of death, they can cause significant health issues and even death if they are not treated quickly [2]. Scientists have shifted to more contemporary methods for disease prediction, techniques like data mining, ML, and Deep Learning (DL), because the manual identification of cardiac disease has many drawbacks.

Moreover, ML models make it uncomplicated to analyze genetic data in its entirety. In addition to transforming and thoroughly analyzing medical data to improve predictions, models can be trained to generate knowledge of pandemic predictions. Many studies have been performed on various ML predictions of heart disease [3]. Classification algorithms like KNN to perform automatic blood pressure classification and DT and NB for HD forecast been used in recent studies. Three different SVM classifier types were used in another study to predict coronary artery disease [4]. The SVM classification of heart sounds provided an automated diagnosis system for heart valve disorders. Neural network models have effectively predicted data and solved various classification issues in recent years.

However, the ECG and patient data are often out of balance, which makes it hard for traditional ML to work impartially [5]. Potential risk factors that raise the likelihood of heart-related problems: unhealthy diets, inactivity, alcohol misuse, and tobacco use. Nevertheless, since the majority of the current dataset on heart disease involves unbalanced data, it is also necessary to discuss the method ML uses to process such data.

This section provides accurate predictions using data from the Cleveland Heart Disease (CHD) dataset obtained from Kaggle. Additionally, the MMNDS technique can be used in a preprocessing step to identify noise and missing fractions. Then, the GCSA method can be implemented to identify feature subsets and select the best features. Finally, RFC algorithms based on ML technology can be proposed to provide accurate results for heart disease prediction. We determine the value of each ML technique in terms of its computational efficiency and predictive accuracy by examining datasets related to heart disease.

This assessment helps determine the best practices for managing cardiac disease and highlights each strategy's advantages and disadvantages. We determine the efficacy of each ML technique in terms of its computational efficiency and predictive accuracy by analyzing datasets related to heart disease.

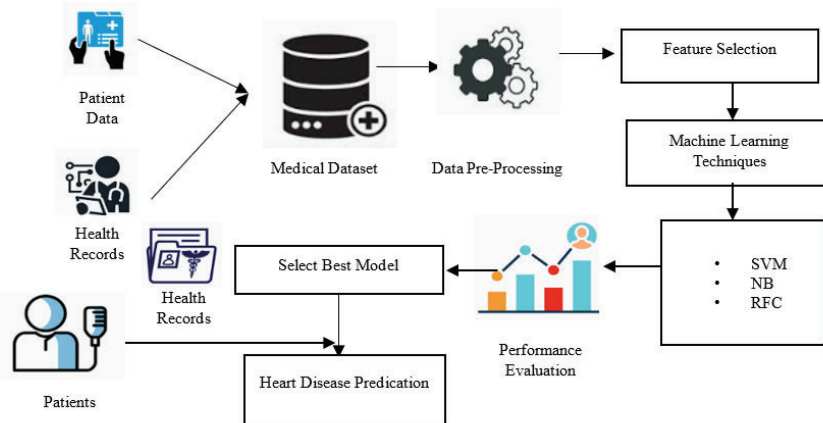


Fig 1. The Heart Disease Prediction based on Basic Architecture Diagram

As shown in Figure 1, Cardiac Disease Forecasting uses a structured architecture diagram that describes the data flows and processes involved in analyzing cardiovascular health. A graph typically consists of several layers of data input, pre-processing, feature selection, and model training. By using ML algorithms at the heart of this architecture, healthcare professionals can gain insights that aid in early detection.

## II. LITERATURE SURVEY

Specialists and medical professionals employ various ML algorithms to forecast heart disease [6]. In addition, optimization-based techniques were used to evaluate the performance analysis. However, medical data analytics must overcome several significant challenges. Likewise, the accuracy of heart disease predictions has been improved based on ML techniques that identify critical features [7]. Furthermore, various features and well-known classification approaches work together to create models that forecast. Cardiovascular disease prediction is made possible by providing healthcare organizations with useful ML technologies to address data analysis issues [8]. The proposed ML technique resolves healthcare data analytics issues. It is a logical synthesis of earlier techniques intended to increase efficiency and produce better results.

Furthermore, when applied to the CHD dataset, ML techniques showed exceptional accuracy in predicting heart disease. Due to the large number of features, they advise reducing the low-dimensional subspace of the CHD using the Jellyfish Optimization Algorithm (JOA) [9-10]. However, features in the dataset and optimization-based model can significantly affect the methodology execution during training. To determine what matters most appropriate parameters for assessing the suggested model's precision and functionality, GridSearchCV uses 5-fold cross-validation and a negative log loss metric for hyperparameter optimization [11].

In addition, a Keras-Based DL model is proposed, and the results are computed using a Dense Neural Network

(DNN) [12]. The offered technique is tested by implementing networks with 3 to 9 layers and different hidden layer structures. However, early detection can be difficult before heart disease develops. Specifically, it uses ML techniques to perform a thorough literature analysis and find practical solutions to the problems associated with unbalanced datasets [13]. Furthermore, it employs a Convolutional NN (CNN) technique to increase the accuracy of the prediction of cardiac disease.

The authors recommend a closed ensemble learning method using quantum support vector classifiers. In addition, Shapley Additive Explanations (SHAP) calculate the importance of each feature and provide a means by which the results of the proposed model can be explained using predictions [14]. To provide the optimal response, the Grey Wolf Horse Herd Optimization-based Shepard CNN (GWHHO-ShCNN) method employs a sophisticated DL technique [15]. However, the vast majority of applications require an enormous amount of data. Significant progress has been made in classification evaluation using Artificial Neural Network (ANN) models for cardiovascular disease prediction [16]. However, monitoring cardiovascular biomarkers to identify patients at risk for functional impairment is essential.

Predict cardiac disease utilizing the ML approach's Hyper-Parameter Optimization Technique (HPOT). Similarly [17], early and accurate heart disease predicts essential for effective analysis. This experiment used DL techniques instead of traditional methods using the UCI cardiology dataset [18]. Additionally, to increase the accuracy of traditional methods, a Cluster-Based Bidirectional Long-Short-Term Memory (C-BiLSTM) has been developed. To extract and forecast decision rules and find correlations between input and output variables, supervised techniques employ the Classification and Regression Tree (CART) methodology [19]. Compared to previous studies and comparisons in heart disease prediction, a novel [20] proposed Bi-directional Recurrent Network and Neuro-Fuzzy (BRN-NF) CART model significantly improves the accuracy prediction index.

## III. PROPOSED METHODOLOGY

With the CHDD from Kaggle serving as a primary data source, this section outlines suggested methods to employ sophisticated data analysis techniques to increase the accuracy of heart disease predictions. The pre-processing approach begins with the implementation of the MMNDS technique. Pre-processing is done by applying the MMNDS technique to account for noise and missing fractions in the dataset. Similarly, the GCSA method can be used to correctly identify and select the features that provide the most suitable features for predictive modeling. Specifically, we implement a set of ML classifiers such as SVM, NB, and RFC that can predict the probability of heart disease using selected patient clinical parameters. Finally, we propose RFC as a choice classification method that integrates multiple decision trees to generate more accurate predictions, providing an evidence-based approach to predicting heart disease. The proposed method can be used to improve the classification results of ML techniques for data pre-processing, feature selection, and achieving optimal results.

As medical informatics research advances, ML algorithms have become a viable method for forecasting cardiac issues. As illustrated in Figure 2, this dataset provides a strong foundation for predicting heart disease through ML algorithms, including preprocessing, feature selection, and classification techniques. Effective cleaning and transformation of the raw data will improve its quality and relevance and improve predictive accuracy. The most crucial factors influencing the prediction of heart disease can be found through a later feature selection stage. Finally, reliable predictive models can be obtained by using classification algorithms such as NB, RF, and SVM. Each ML algorithm provides different advantages for data processing and improving accuracy.

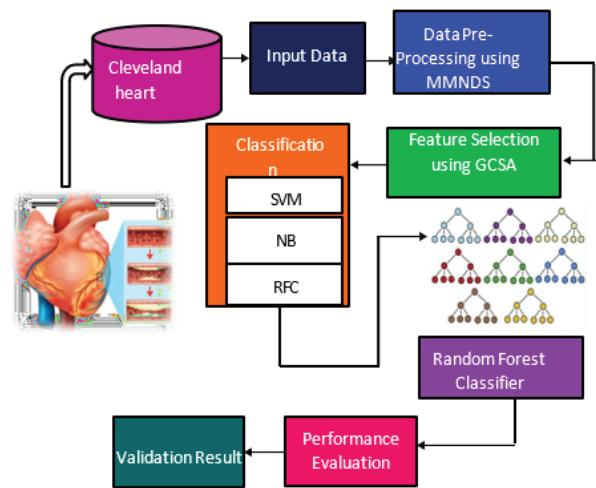


Fig. 2. The Presented Method Architecture Diagram

## A. Cleveland Heart Disease Dataset (CHDD)

The CHDD predicts cardiovascular disorders through supervised ML. The CHDD is sourced from the Kaggle ML repository. In 1988, the Cleveland Clinic Foundation began using the Cleveland dataset for health research. There were 303 participants with 76 distinct characteristics in the original dataset. The target class characteristic is one of just 14 features that most researchers employ. These characteristics include blood pressure, cholesterol, blood sugar, age, gender, and other health indicators. There are five class labels in the original Cleveland dataset.

A key component of cardiovascular analysis, the Cleveland Heart Disease Dataset serves as a fundamental resource for comprehending heart disease risk factors, as illustrated in Table 1. This dataset, which comes from the medical records of patients assessed for heart issues, has been widely used in statistical analysis and ML to enhance predictive models for detecting heart disease.

TABLE I: DATA SET FEATURE COLLECTION

age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
52	1	0	125	212	0	1	168	0	1	2	2	3	0
53	1	0	140	203	1	0	155	1	3.1	0	0	3	0
70	1	0	145	174	0	1	125	1	2.6	0	0	3	0
61	1	0	148	203	0	1	161	0	0	2	1	3	0
62	0	0	138	294	1	1	106	0	1.9	1	3	2	0
58	0	0	100	248	0	0	122	0	1	1	0	2	1
58	1	0	114	318	0	2	140	0	4.4	0	3	1	0
55	1	0	160	289	0	0	145	1	0.8	1	1	3	0
46	1	0	120	249	0	0	144	0	0.8	2	0	3	0
54	1	0	122	286	0	0	116	1	3.2	1	2	2	0
71	0	0	112	149	0	1	125	0	1.6	1	0	2	1

### B. Min-Max Normalized Data Scaling (MMNDS)

This section preprocesses the data using the MMNDS method to identify noise and missing scores in the heart disease dataset. The heart disease data can be preprocessed through missing data filtering and normalization using the MMNDS algorithm to analyze data for heart disease prediction effectively. Additionally, the MMNDS method can be used to clean data by removing noise and duplicate records. The technique can also be used to replace all missing values with the mean and median of the existing data to increase the variance and eliminate useless features. The MMNDS algorithm is also used to reduce numerical complexity during the computational process of heart disease predictions. The heart disease dataset is normalized between 0 and 1. Thus, all attributes can be evaluated to ensure consistency within a data set, allowing for accurate comparison and analysis of heart data. The dataset contains many missing values and meaningless or redundant data. Calculate the missing values using the mean value, as illustrated in Equation 1. Let's assume  $N$ -total number of datasets,  $X$ -input data,  $x_i$  –data row,  $M$  –Mean value.

$$M(X) = \frac{\sum_{i=0}^N x_i}{N} \quad (1)$$

Data normalization used in preprocessing can transform data values to a specific range from 0 to 1. Estimate the data range using zero mean or Z-score normalization, as shown in Equation 2. Let's assume  $X'_i$  –normalized data,  $SD(X)$  –standard deviation.

$$X'_i = \frac{x_i - M(X)}{SD(X)} \quad (2)$$

Calculate the average value of the standard deviation as shown in Equation 3. Let's assume  $\sigma$  –standard deviation.

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (X_i - M(X))^2} \quad (3)$$

To ensure uniformity and consistency in the heart disease dataset, normalized data can be scaled using the minimum and maximum values, as shown in Equation 4. Let's assume  $X_{Max} - X_{Min}$  is represented the minimum and maximum value,  $X^i$  –normalized data scaling.

$$X^i = \frac{X - X_{Min}}{X_{Max} - X_{Min}} \quad (4)$$

The noise and unwanted data in the heart disease dataset are adjusted using the MMNDS algorithm, which evaluates the minimum and maximum values during the data preprocessing stage.

### C. Genetic-Based Crow Search Algorithm (GCSA)

This section employed the GCSA algorithm to identify the optimal subset of features and select the best performers. Additionally, the total number of features optimized using GCSA features determines the length of

the dataset and feature selection for crow searches. In a binary string, 1 indicates a selected feature, and 0 indicates an unselected feature. Furthermore, the sum of the observer and crow searches equals the number of features in the dataset. Employing the Crow search algorithm, we can discuss the initial process of population genetics and identify optimal features in a heart disease dataset. The GCSA algorithm can optimize the contrast solution by simultaneously monitoring solutions from meta-heuristic optimization models. The optimal crow searches fitness function can be calculated using the basic functions. The fitness level of the crow's new posture is constrained to be the highest among the consecutive postures. Crow frequently updates its storage with new locations. When multiple iterations are executed, the optimal storage location associated with the target is referred to as the filtered optimal feature set solution. When multiple iterations are executed, the optimal storage location associated with the target is referred to as the filtered optimal feature set solution as following the below algorithm 1.

#### Algorithm

Input: Normalized data scale  $X^i$

Output: Optimal feature solution

Start

1. Feature sets are estimated using the genetic algorithm in the initial population.
2. To obtain the optimal features, evaluate fitness functionality
3. For each  $O_F$  do
4. Select the optimal feature fitness function
5. Calculate the crossover of randomly generated feature sets
6. Compute the new population created based on a genetic algorithm
7. Assess the opposition process for acquiring the initial population.
8. Compute the fitness function for the optimal solution according to the genetic algorithm procedure
9. Get the best solution
10. Compute the objective function to obtain the optimal solution
11. Create new position
12. If  $N_p = F$
13. Update the storage and location
14. Else
15. Compute the current optimum population
16. If obtained  $M_I$  then
17. Return optimal feature solution
18. End

Let's assume  $F$  –fesible, new position,  $O_F$  – optimal feature,  $M_I$  –maximum iteration. When multiple iterations are executed, the optimal storage location associated with the target is referred to as the filtered optimal feature set solution, as shown in algorithm 1.

### D. Classification Algorithms

#### Support Vector Machine (SVM):

Support vector machines (SVMs) are used to linearly separate classes in a variety of linear and non-linear patterns. Compared to the Naive Bayes classifier and SVM classifier, SVM is a supervised ML technique. The SVM model is based on the idea that a large number of trees will produce correct results. This SVM algorithm employs classification hyperplanes, to segregate more than one class. SVM has another application in regression analysis. This categorizes both linear and non-linear data. The focus of the SVM classifier is to find hyperplanes in the n-dimensional space. It also ensures that the maximum margin is achieved among positive and negative samples near the hyperplane when classifying multiple classes. SVM is highly affected by nonlinear and non-separable data. However, this problem can be solved using one of the Kernel functions, described in the next section, to transform the data from the input space to a new, higher-dimensional space. Through the confusion matrix, the performance of the model and the accuracy metrics were determined.

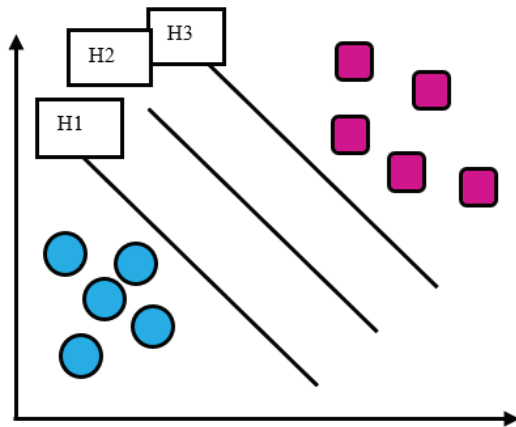


Fig. 3. Architecture Diagram for SVM.

As shown in Figure 3, all samples of one class can be analyzed on one side of the hyperplane and samples of another class on the other side to create a hyperplane. Furthermore, the optimal hyperplane can be designed to maximize the distance between the two classes. The support vector is the data point of the class that is closest to the hyperplane.

#### Naive Bayes (NB):

A naive Bayes classifier is a supervised learning methodology that can be used with the help of the Bayes theorem. The NB Classifiers create a prototype of a particular class based on the presence or absence of another class and its features. It is frequently used to determine the posterior probability of a specific observation and make inferences based on the most likely outcome. Heart disease feature reliance is appropriate for the medical industry and the identification of heart disease because it is simple to produce and generally functions effectively.

#### Random Forest Classifier (RFC):

Supervised learning algorithms, such as Random Forest, apply the development of ML classifiers that enhance the performance of decision trees. This method utilizes a random vector and randomly combines multiple tree predictors. Instead of splitting nodes based singly on specific variables, Random Forests select a subset of the best predictions randomly from the nodes. Using randomly chosen data samples, it builds decision trees, generates predictions from each tree, and then votes to choose the best response. Furthermore, this algorithm provides a more reliable indication of feature importance. To increase a dataset's predicting accuracy, a classifier random forest applies and averages several decision trees to various subsets of particular data. Instead of depending entirely on the trees, a random forest uses each tree's forecasts. These forecasts, which are based on the votes of the majority of participants, predict the outcome.

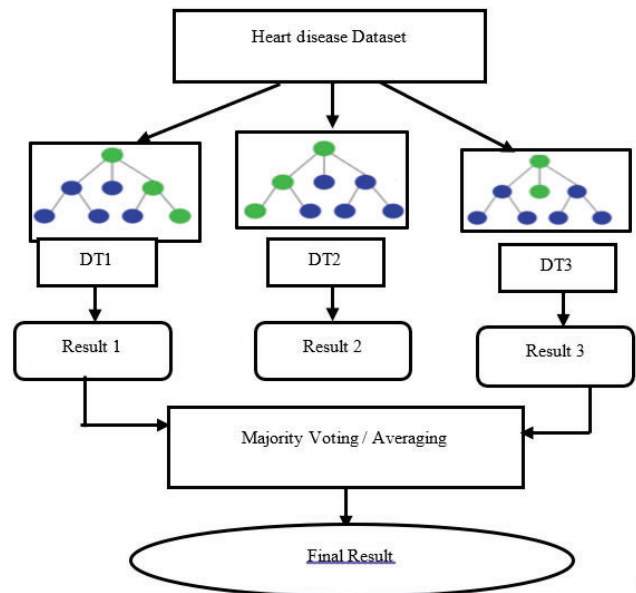


Fig. 4. The Architecture Diagram for Random Forest.

Figure 4 shows an architecture diagram of a random forest classifier that illustrates the significance of early identification of heart disease. This classifier improves accuracy by effectively combining predictions from several decision trees using an effective ensemble learning methodology.

## IV. RESULTS AND DISCUSSION

We used a review approach to predicting heart disease to assess the main effects. Furthermore, by presenting the KNN, SVM, NB, RF, DT, and other review methods in the previous CNN method, the performance evaluation has been improved significantly. Therefore, we evaluate these results by using criteria like F1 score, mistake rate, recall, precision, and precision to improve in predicting HD. Furthermore, illustrated in Table 2, an optimal estimate is obtained by training and evaluating the dataset's features using Python

and the tool Jupyter based on the simulation settings to forecast heart disease.

TABLE II: SIMULATION PARAMETERS

Simulation	Vales
Name of the Dataset	CHDD
No of Dataset	305
Used Language	Python
Used Tool	Jupyter
Data of Training	177
Data of Testing	128

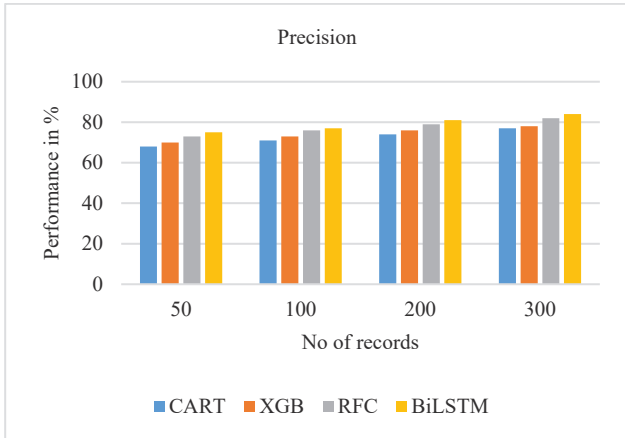


Fig. 5. Analysis of Precision

Figure 5 illustrate that the deployed precision analysis method is able to anticipate heart disease. Nonetheless, the suggested approach attains a precision analysis of 77%, 78%, and 82%, in difference to earlier methods like BiLSTM, CART, and XGB. The proposed RFC method's comparison rate with the previous CNN method is defined as 84%.

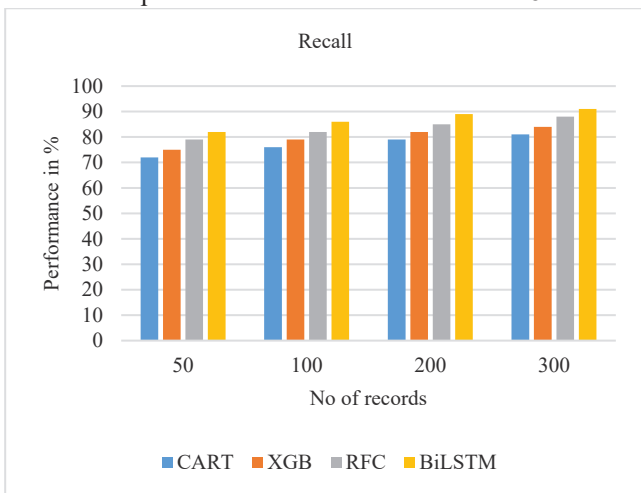


Fig. 6. Analysis of Recall

The accuracy of the suggested memory analysis approach in forecasting HD is illustrate in Figure 6. The suggested approach obtains recall rates of 81%, 84%, and 88% respectively, in comparison to previous techniques like BiLSTM, CART, and XGB. Additionally, the suggested RFC

method has a 91 % recall rate when compared to the earlier method.

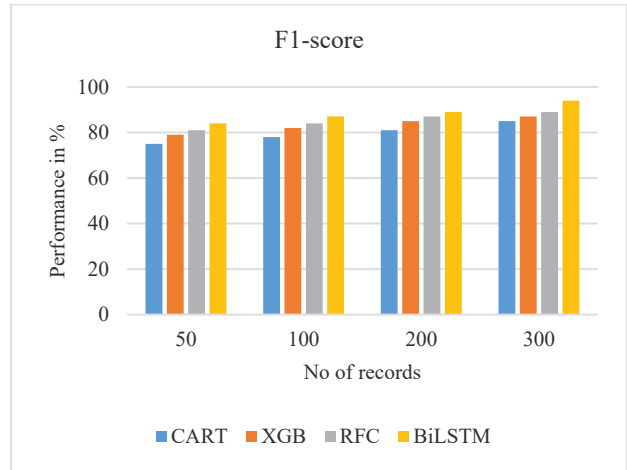


Fig. 7. Analysis of F1-Score

The suggested F1-score analysis method can accurately forecast heart disease, as shown in Figure 7. The proposed approach achieves F1-Score rates of 85%, 87%, and 89%, and 94%, respectively, compared to previous techniques like BiLSTM, CART, and XGB. Additionally, the suggested RFC method has a 94% F1- Score rate compared to the earlier methods.

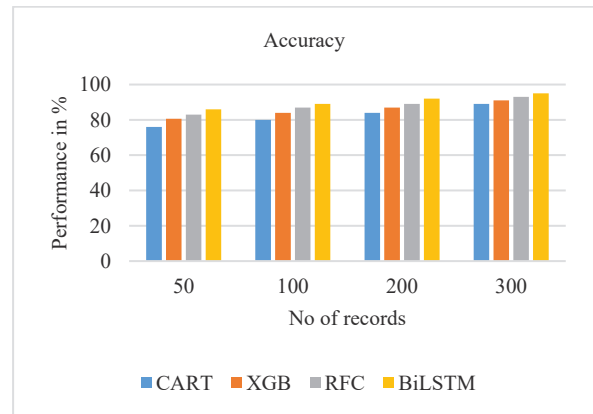


Fig. 8. Analysis of Accuracy

The proposed accuracy analysis method significantly predicts heart disease, as shown in Figure 8. This approach achieves accuracy rates of 89%, 91%, and 93%, respectively, outperforming previous techniques C-BiLSTM, CART, and XGB. Additionally, the proposed RFC method boasts a 95% accuracy rate, surpassing the earlier CNN method.

TABLE III: OVERALL PARAMETERS

Comparison methods/ metrics	Accuracy (%)	F1-score (%)	Precision (%)	Recall (%)
CART	89	85	77	81
XGB	91	87	78	84
RFC	93	89	82	88
BiLSTM	95	94	84	91

Table 3 shows the overall performance of quantitative metrics results.

## V. CONCLUSION

The need of carefully assessing feature selection and model combination while developing ML techniques to predict cardiac disease is demonstrated in the study's conclusion. The encouraging findings support more research and development of these methods to enhance health outcomes through improved prediction skills and develop an ML model that uses the CHDD to accurately forecast HD. Our novel illustrates the significance of feature selection and dataset selection to the effectiveness of ML systems. We have examined more than a few ML approaches, for instance, SVM, KNN, NB, DT, and RF via an evaluative lens to demonstrate their efficacy in forecasting HD. Methodology robustness is crucially indicated by the suggested examination of metrics, as well as F1 score, accuracy, recall, and precision. The conclusions from the study demonstrate that using a combination of classifier methodologies results in the most substantial improvement in performance measures for HD forecast. Furthermore, the deployed strategy improves the prior approach with an accuracy rate of 95%.

## REFERENCES

- [1] J. P. Li, A. U. Haq, S. U. Din, J. Khan, A. Khan and A. Saboor, "Heart Disease Identification Method Using Machine Learning Classification in E-Healthcare," in *IEEE Access*, vol. 8, pp. 107562-107582, 2020.
- [2] M. Swathy, K. Saruladha., A comparative study of classification and prediction of Cardio-Vascular Diseases (CVD) using Machine Learning and Deep Learning techniques, Volume 8, Issue 1, March 2022, Pages 109-116. <https://doi.org/10.1016/j.ict.2021.08.021>.
- [3] Bharti, Rohit, Khamparia, Aditya, Shabaz, Mohammad, Dhiman, Gaurav, Pande, Sagar, Singh, Parneet, Prediction of Heart Disease Using a Combination of Machine Learning and Deep Learning, *Computational Intelligence and Neuroscience*, 2021, 8387680, 11 pages, 2021. <https://doi.org/10.1155/2021/8387680>
- [4] Baccouche, Asma, Begonya Garcia-Zapirain, Cristian Castillo Olea, and Adel Elmaghraby. 2020. "Ensemble Deep Learning Models for Heart Disease Classification: A Case Study from Mexico" *Information* 11, no. 4: 207. <https://doi.org/10.3390/info11040207>
- [5] Md Manjurul Ahsan, Zahed Siddique, Machine learning-based heart disease diagnosis: A systematic literature review, *Artificial Intelligence in Medicine*, Volume 128, June 2022, 102289. <https://doi.org/10.1016/j.artmed.2022.102289> Intro
- [6] Bhavekar, Girish Shrikrushnarao, et al. "heart disease prediction using machine learning, deep Learning and optimization techniques-A semantic review." *Multimedia Tools and Applications* (2024): 1-28.
- [7] Srinivasan, S., Gunasekaran, S., Mathivanan, S.K. et al. An active learning machine technique-based prediction of cardiovascular heart disease from UCI-repository database. *SciRep* 13, 13588 (2023). <https://doi.org/10.1038/s41598-023-40717-1>
- [8] Ahmed, Rehan, Maria Bibi, and Sibtain Syed. "Improving heart disease prediction accuracy using a hybrid machine learning approach: a comparative study of SVM and KNN algorithms." *International Journal of Computations, Information and Manufacturing (IJCIM)* 3.1 (2023): 49-54.
- [9] Ahmad AA, Polat H. Prediction of Heart Disease Based on Machine Learning Using Jellyfish Optimization Algorithm. *Diagnostics (Basel)*. 2023 Jul 17;13(14):2392. doi: 10.3390/diagnostics13142392. PMID: 37510136; PMCID: PMC10378171.
- [10] Shah, Devansh, Samir Patel, and Santosh Kumar Bharti. "Heart disease prediction using machine learning techniques." *SN Computer Science* 1.6 (2020): 345.
- [11] Chandrasekhar, Nadikatla, and Samineni Peddakrishna. "Enhancing heart disease prediction accuracy through machine learning techniques and optimization." *Processes* 11.4 (2023): 1210.
- [12] Almazroi, Abdulwahab Ali, et al. "A clinical decision support system for heart disease prediction using deep learning." *IEEE Access* 11 (2023): 61646-61659.
- [13] Bizimana, Pierre Claver, et al. "[Retracted] an Effective Machine Learning-Based Model for an Early Heart Disease Prediction." *BioMed Research International* 2023.1 (2023): 3531420.
- [14] Abdulsalam, Ghada, Souham Meshoul, and Hadil Shaiba. "Explainable heart disease prediction using ensemble- quantum machine learning approach." *Intell. Autom. Soft Comput* 36.1 (2023): 761-779.
- [15] Kanchanamala, Pendela, A. Suja Alphonse, and PV Bhaskar Reddy. "Heart disease prediction using hybrid optimization enabled deep learning network with spark architecture." *Biomedical signal processing and control* 84 (2023): 104707.3
- [16] Sarra, R. Rone, A. Musa Dinar, and M. Abed Mohammed. "Enhanced accuracy for heart disease prediction using artificial neural network." *Indonesian Journal of Electrical Engineering and Computer Science* 29.1 (2023): 375-383.
- [17] Asif, Daniyal, et al. "Enhancing heart disease prediction through ensemble learning techniques with hyperparameter optimization." *Algorithms* 16.6 (2023): 308.
- [18] Dileep, P., et al. "An automatic heart disease prediction using cluster-based bi-directional LSTM (C-BiLSTM) algorithm." *Neural Computing and Applications* 35.10 (2023): 7253-7266.
- [19] Ozcan, Mert, and Serhat Peker. "A classification and regression tree algorithm for heart disease modeling and prediction." *Healthcare Analytics* 3 (2023): 100130.
- [20] Meenal, M. Revathy, and S. Mary Vennila. "Bidirectional recurrent network and neuro-fuzzy frequent pattern mining for heart disease prediction." *SN Computer Science* 4.4 (2023): 379.