International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V71I8P222

Volume 71 Issue 8, 253-262, August 2023

Original Article

Comparative Analysis and Evaluation of Machine Learning Algorithms for Early Detection of Cardiac Disease

Ritika[1], Rajender Singh Chhillar[2], Sandeep Dalal[3]

[1,2,3]*Department of Computer Science and Applications, MDU Rohtak, Haryana, India*

[1]*Corresponding Author : ritikarathee@gmail.com*

Received: 21 April 2023 Revised: 03 July 2023 Accepted: 22 July 2023 Published: 15 August 2023

Abstract *- The diagnosis of Cardiac disease at the heart attack or stroke stage is very costly with lesser chances of survival. The online and automated healthcare system can be adapted to diagnose cardiac disease at an earlier stage. Various machine learning models are methods integrated within the health care systems to optimize the performance of cardiac disease prediction. In this paper, six supervised learning algorithms are implemented on an authenticated cardiac disease dataset. Bayesian network, decision tree, random forest, SVM, Neural network and Radial Basis Function classifiers are applied in this work on a filtered and preprocessed dataset. The analytical results are obtained in terms of accuracy, precision, recall, TP rate, FP rate and F-score parameters. The analysis results identified that the SVM and Naive Bayes classifiers achieved the most effective results. SVM with Gaussian kernel achieved the maximum accuracy rate, TP rate and least FP rate.*

Keywords *- Machine Learning, Supervised learning, Cardiac disease, Feature selection, Classifiers.*

1. Introduction

In recent years, Cardiovascular Disease (CVD) has become one of the most critical diseases, with millions of deaths per year. Based on a study, 31% of total deaths are because of heart disease. Any kind of abnormal condition or behaviour of blood vessels can cause cardiac disease or attack. This abnormal blood pressure and blood flow can build up plague that narrows up the veins and arteries. It disturbs the blood flow from the heart, which increases the risk of heart stroke or failure. Various reasons or symptoms of heart disease include high blood pressure, high cholesterol levels, genetic mutations, physical inactivity, poor nutrition and obesity. Once a heart attack or heart failure situation arises, it is very costly and risky for a patient. The chances of survival are also very lesser in such situations. But if the heart disease symptoms are detected at an earlier stage, then precautionary treatment can be taken to increase the survival chances [1].

Various machine learning, mathematical and deep learning-based automated models are investigated and implemented in numerous applications to predict heart disease. The effectiveness of these models depends on the number of available or collected features and the data collection methods. Data collection methods are also improved in recent times. The real-time and IoT-based methods are integrated to capture the real situation of a patient. Various health organizations are collecting this kind of information to take decisive results accurately. The machine learning methods use these features or data to extract the pattern to identify healthy and heart patients. The researchers used the supervised, semi-supervised and unsupervised learning methods in different models to correlate the features and generate the patterns. The size and reliability of these patterns are reflected in the accuracy and performance of the disease prediction and classification models [1]. A prototype model for IoT and ML[22][24][31] based on heart disease prediction is provided in Figure 1. The figure shows that the functional server is responsible for processing the adequate feature collected from patients. These general features, such as age, gender, pain type, etc., can be collected directly from the patient. The actual heart-associated features, such as heart rate, glucose level, and blood pressure, can be captured through various IoT devices. An adequate model can be applied to these composite features, and the predictive result can be stored in the available dataset[3][25][26][27].

In this paper, an analytical and comparative study is provided on various machine learning methods. The supervised learning methods are implemented on the cardiac disease dataset. In this section, the criticality and features of the cardiac disease system are described.

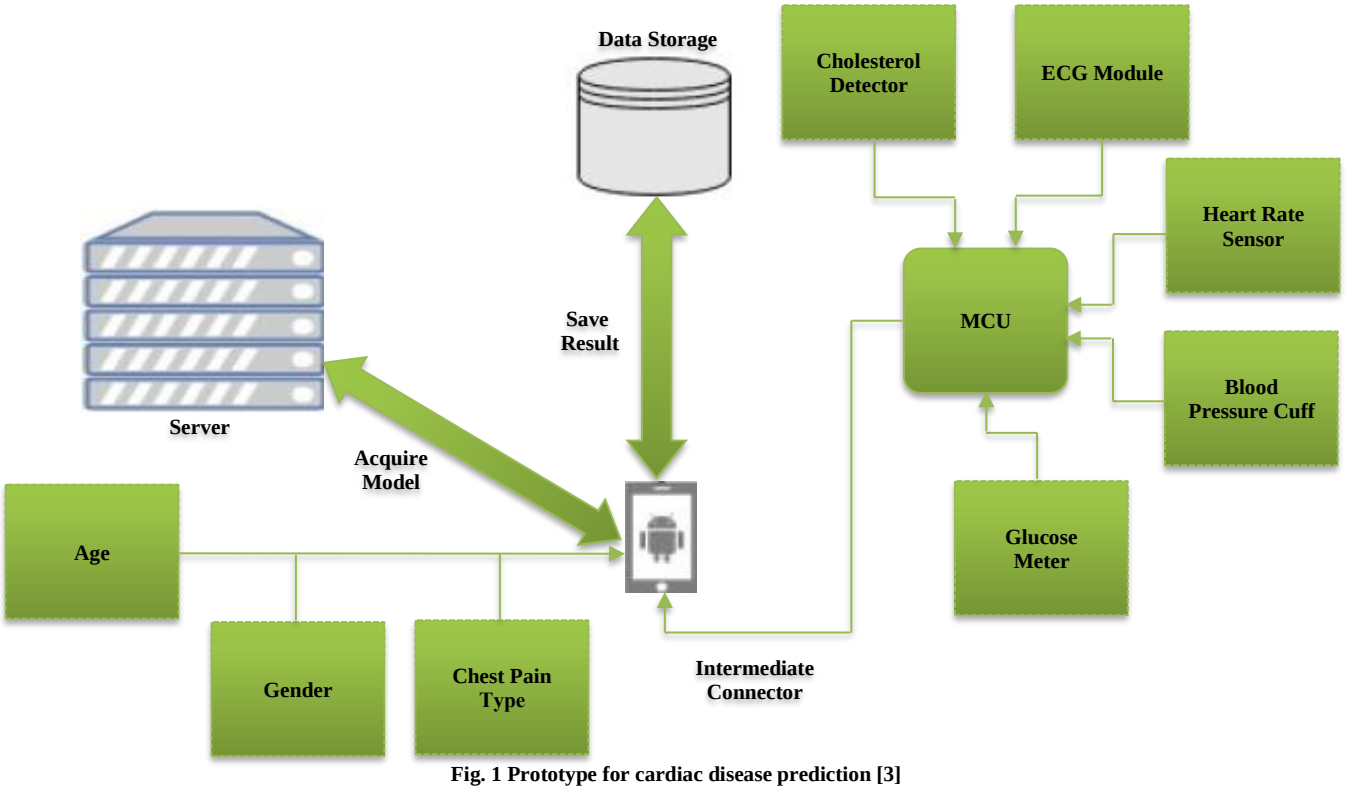

Fig. 1 Prototype for cardiac disease prediction [3]

The standard prototype used that captures the data from different sources and processes the model is also defined. Section 2 discusses various machine learning methods and models defined by earlier researchers with their associated features and capabilities. Section 3 discusses functional behaviour and features of various supervised machine learning methods. The adaptation of the classification algorithm is described for the cardiac health care system. Section 4 provides the description of the dataset and the results obtained for different classifiers. Section 5 includes the conclusion and future scope of the work.

2. Related Work

The severity of cardiac disease attracts researchers to predict the occurrence of heart disease. The researchers proposed various heart disease prediction, classification methods, and models to predict cardiac disease symptoms accurately. The author used various hybrid methods, optimization methods and feature selection methods to improve the capabilities of machine learning models. Some of the work provided by the researchers for machine learning methods and models is provided in this section.

Li et al.[5] proposed a conditional mutual information-based feature selection method to optimize the heart disease classification. The author improved the existing Support Vector Machine (SVM) for optimizing the performance of cardiac disease. The fuzzy system-based ensemble learning model was proposed for accurately diagnosing heart disease. Fitriyani et al.[6] proposed a heart disease-based clinical decision support system to identify cardiac disease at the earlier stage. The author integrated the density-based spatial clustering of application with noise (DBSCAN) clustering method to avoid the outlier. The filtered dataset was processed under a hybrid synthetic minority over-sampling technique with the edited nearest neighbour (SMOTE-ENN) method with XGBoost classifier for heart disease prediction. The model was applied on Statlog and Cleveland datasets and achieved 95,90% and 98.40% accuracy, respectively. A fuzzy and genetic algorithm based hybrid model [7] was proposed to optimize the cardiac disease prediction. The author defined the rough sets for effective feature selection and fuzzy rule within the classification stage. An adaptive genetic algorithm was applied to a filtered feature set for predicting heart disease. The method was applied to the publicly available UCI heart disease dataset and achieved higher accuracy than existing methods.

Amin et al.[8] provided an analytical study on various feature selection and prediction methods to identify the most effective classification technique. The author applied the model to KNN, Naive Bayes, Decision Tree, SVM, Logistic Regression and Neural network methods. The analysis results identified that the voting-based ensemble learning method achieved 87.4% accuracy method. Haq et al.[11] proposed a hybrid machine learning model to predict cardiac disease accurately. The author implemented the seven machine learning algorithms with a feature selection stage to optimize

the predictive results. The author implemented the pre-processing, feature selection, machine learning and validation algorithms to evaluate the performance results. The health care system is defined to optimize the diagnosis of cardiac disease. Latha et al.[12] combined weak predictive algorithms and proposed an ensemble learning algorithm to optimize the performance of cardiac disease prediction—the author combined and compared the multiple classifiers with existing classifiers and ensemble methods.

The proposed ensemble classifier improved the accuracy to 7% over the weak classifiers. Gupta et al.[13] proposed an intelligent machine learning framework by processing the factor analysis of mixed data (FAMD). This method identified the effective features of the dataset and applied different predictive models. The proposed model improved the accuracy and performance over baseline methods. Singh et al.[16] provided a study and analytical work for predicting the performance of various machine learning algorithms. The author analyzed the decision tree, KNN, linear regression, and support vector machine methods for predicting their performances. The author identified the KNN as the most effective method with maximum accuracy.

Rani et al.[30] defined a hybrid feature selection method to minimize the dataset features. The author used the single scalar-based multiple classifiers to optimize the work. This hybrid model achieved 86.6% accuracy. Garate- Escamila et al.[20] used the chi-square and principal component analysis-based hybrid filtration method and combined it with a random forest classifier. The proposed method achieved 98.7% accuracy for Cleveland and 99% for Hungarian datasets. The method achieved the performance in comparison with existing methods. Muhammad et al.[21] proposed a computational predictive system for the prediction of cardiac disease. The study was defined on various feature filtration and classification method. The author identified that the proposed model achieved effective results for Chi-square and P-value filtration methods.

3. ML Models/Approaches

Cardiovascular disease can lead to death. It has the highest death rate overall diseases. The early prediction of heart disease can save the life of an individual. The researchers investigated various predictive methods and classification models. In this section, some of the most effective machine learning models are discussed. Researchers used these models for the prediction of various medical diseases [18]. Some researchers have used these machine-learning models to predict the chances of cardiac disease. In this paper, we have defined a cardiac disease prediction model further integrated with multiple machine-learning models. Figure 2 shows the model used in this work for implementing and analyzing different machine-learning methods. In this section, a detailed description of the model and functional description of each of the machine learning methods is provided.

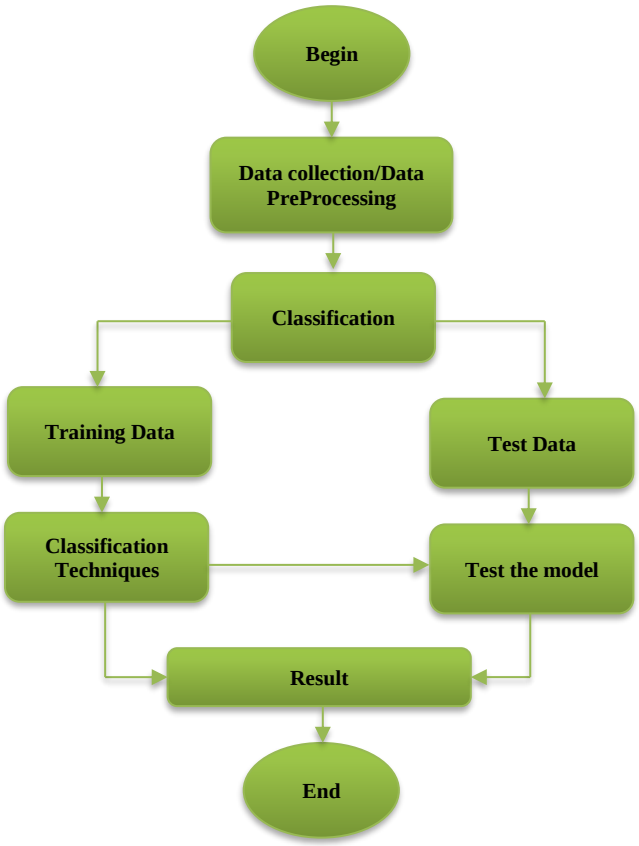

Fig. 2 Prototype for ML model on cardiac disease

Figure 2 shows the cardiac disease prediction model that is used in this research. All the machine learning algorithms defined in this paper are implemented in the classification stage of this model. In the first step of this model, the cardiac disease dataset is collected. In this work, the dataset is collected from authenticated external sources, and its description is provided in section 4. Once the dataset is obtained and the classification model is applied to it. According to this model, the dataset is divided into two subsets: training and testing set. While performing the implementation and experimentation, 90% of the data is taken as a training set and 10% is considered as a test set. The classification algorithms applied in this work are Decision Tree, Neural network, Bayesian network, SVM, and RBF networks. The model is first applied to the training set and generates the classification. Later, the test set is applied to this model to predict cardiac disease. The analysis results are obtained on the parameters described in section 4.2. In this section, the functional description of each of the classification methods used in this paper is provided.

3.1. Decision Tree

The decision tree is an effective tree-based classifier that performs a level-specific analysis of individual features to identify the instance class. It is a recursive process that analyzes each feature to take the decision over the information

space. The node-level split is performed over the dataset with the specification of certain functions. The attribute value-based decision is performed on each node, and this process is repeated on each feature until the leaf node does not occur. A particular class represents each of the leaf nodes. There are multiple forms of decision trees with specific integrated methods [15][29]. In this work, we have applied the decision tree with three integrated steps.

1. Compute the Gini Index for each individual feature
2. Apply the node-based decision to split the dataset. Identify the optimal value for this split ion.
3. Apply this process recursively to generate the complete decision tree. The process is pleated till the terminal node does not occur.
4. Represent each of the leaf nodes with a specific class.

3.2. Random Forest

Random forest is an integrated ensemble learning method that has provided good results for many classification applications. It combines multiple decision trees and generates a forecast to perform effective prediction. It uses a likelihood estimation-based analysis method to identify the prediction class. It uses a voting method on each decision tree to identify the respective instance class[15][29]. The functional process of the random forest algorithm is shown in Figure 3[17]. The steps followed by the random forest algorithm are given below:

1. Take a sub-sample set from the training set.
2. Create a decision tree on this sub-sample set.
3. Identify each sub-tree's interior node and the best split for that part.
4. Evaluate the values for each tree separately and define the expected classes for each leaf.

Apply the majority voting method while combining the trees and take the final decision of the relativeinstance class.

3.3. Naïve Bayes

Naive Bayes is a probabilistic learning model that uses the Bayes rule on independent features. The subsequent probability computation is performed on different features, and the class is allocated to the highest probability value. The Gaussian function is implied within the method with prior probability for computing the data class. The conditional probability is computed for dependent features[4].

The Gaussian function-based priority probability P(Xf) i used for feature class computation is provided in equations (1) and (2). It is one of the simplest and faster methods that can predict the best results when the assumption and independence of features are known[9][10]. Algorithm- 1 defines the algorithmic process followed by the Naïve bayes classifier.

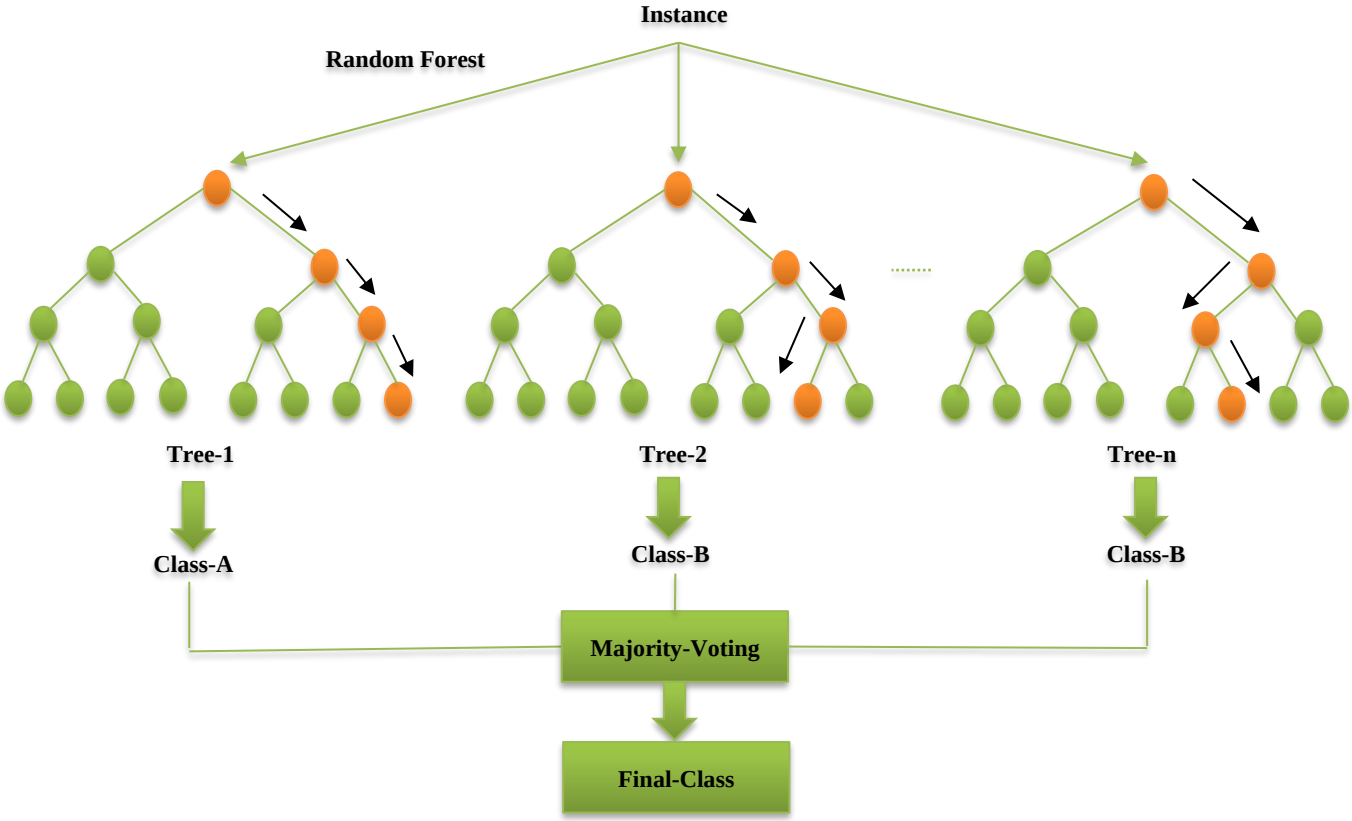

Fig. 3 Functional process of random forest

$$P(X_{f1}, X_{f2}, \ldots X_{fn}|c) = \prod_{i=1}^{n} P(X_{fi}|c) \quad (1)$$

$$P(X_{fi}|c) = \frac{P(c_i|X_f)\,P(X_f)}{P(c_i)} c \in \{cardic\ patient, healthy\} \quad (2)$$

Algorithm-1

1. Define the dataset D with n features called X(x1,x2....xn) and m is number of classes C(c1,c2....cm)
2. Define the Bayes theorem for mapping the features to the classes
$$P(X|C_i) = \frac{P(X|C_i)*P(C_i)}{P(X)}$$
3. Compute probabilistic estimation for each feature P(X) and respective to each class $P(X|C_i)$ such that P(Ci) is maximized
4. Compute the class driven conditional independence $P(X|C_i) = P(x1|C_i) * P(x2|C_2)\ldots\ldots\ldots P(xm|C_i)$
5. $P(X|C_i)$ $P(C_i)$ is evaluated for obtaining the class of X and providing it as the final identified class.

3.4. Support Vector Machine (SVM)

Support Vector Machine (SVM) is an effective classifier that establishes the best relation between the features and the relative class. The method is effective in minimizing the classification error. Sequential Minimal Optimization is the algorithm used to identify the optimum class for a particular feature set. The effectiveness of this algorithm depends on the kernel used in the algorithm. Two of the most common kernel functions we have compared in this work are Linear and Gaussian Kernels. Figure 4 shows the graphs for Gaussian and Linear Kernels. As shown in the figure, a Linear Kernel provides a straight-line division to separate the classes. Whereas the Gaussian kernel can perform a Gaussian kernel-based interpretation to separate the classes. It uses a curved or non-planer separation, as shown in Figure 4. Equation (3) shows the linear kernel, and equation (4) shows the Gaussian kernel specification. The linear kernel is basically used when features are more than available samples, and if the samples are lesser, then the Gaussian kernel is used. The dimension of data affects the selection of the kernel[2].

$$Ker_{linear}(x^i, x^j) = x^i * x^j \quad (3)$$

$$Ker_{gaussian}(x^i, x^j) = \exp\left(-\frac{||x^i - x^j||^2}{2\sigma^2}\right) x^i * x^j \quad (4)$$

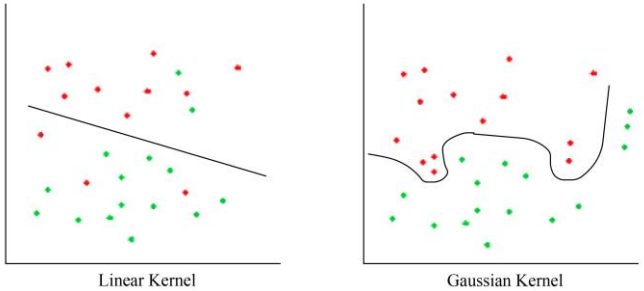

Fig. 4 SVM kernels

3.5. Neural Network

A neural network uses the brain's neuron cell structure to solve a problem. It is an intelligent machine learning algorithm that uses functionality similar to a human brain. It is a layered architecture that is provided in Figure 5. This architecture consists of three layers called Input, Hidden and output layers. The neural network is a non-linear and statistical architecture that can solve complex problems. In the first layer of this layer, the featured input is provided. In the hidden layer, the weights are applied to the input to map the input to a specific class. In this stage, the feature-driven weights are applied to improve the performance and accuracy of the model. Once the data is processed over the model, the particular disease class is obtained as the final output.

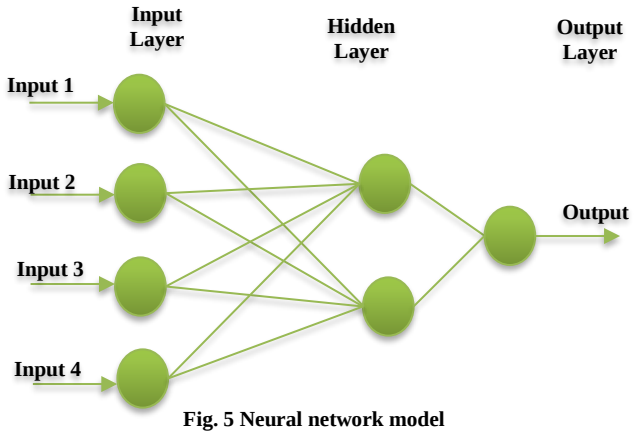

Fig. 5 Neural network model

The model's effectiveness depends on the neurons' behaviour, and the modulation of these neurons can be improved within the hidden layer. The hidden layer is the actual processing layer that can identify the pattern from the input and perform the computation over it. A neural network can have multiple hidden layers. This work processes the hidden layers with backpropagation called Multi-layer perceptron. After processing the pattern and performing the computations, the particular disease class is obtained by the model and presented as output through the output layer. Another effective element of this classifier is the activation function. Linear, sigmoid, logistic, etc., are different activation functions that control this classifier's flow. In this work, we have used the sigmoid activation function to optimize the result. The implementation and comparative results of this classifier are provided in the next section.

3.6. Radial Basis Function

The radial basis function is a neural network that performs the interpolation on data points within the multidimensional space. The functional link in the network form is provided in Figure 6. It defines a regulation with a basis function. The associated stabilizer is defined for achieving radial symmetry. A multivariate continuous function is defined with domain-specific modeling to achieve higher accuracy. The regularization function is defined to

minimize the error in the system. It is a three-layer network in which the radial basis function is defined in the hidden layer. The nonlinear activation function is defined to map the input to the output with better accuracy.

It is an optimized function in which the activation function is applied at the hidden layer. The activation function of the classifier is defined as $\emptyset_0 = 1$. In this optimization function, the main objective is to minimize the mean square error (MSE). The weights are adjusted in the hidden layer to generate optimized results. The radial function is here defined with the Gaussian function. The output function is given as

$$y_i(\vec{x}) = \sum_{k=1}^{J2} \omega_{ki} \emptyset(||\vec{x} - \overrightarrow{c_k}||), \text{ i=1,2,….J3} \quad (5)$$

Where, $y_i(\vec{x})$ represents ith output,

ω_{ki} represents the connection weight for the ith unit

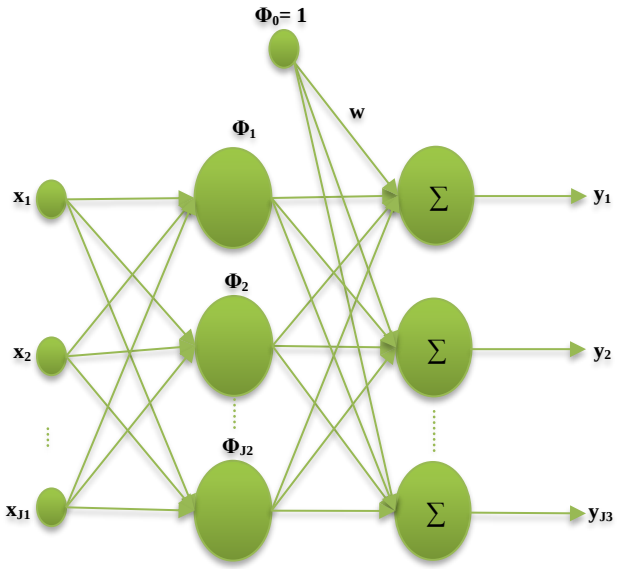

Fig. 6 Radial Basis Network Architecture

4. Results

In this paper, the defined standard model for cardiac disease prediction is defined and experimented with. The experimental results are provided and discussed in this section.

4.1. Dataset

The experiments and implements of the proposed model and classification methods are implemented to an authenticated dataset called Statlog Heart disease dataset[28]. This dataset is taken from the UCI repository. The dataset contains 14 features, including six numeric and eight categorial features. The dataset has two main classes, called healthy and cardiac disease. The analysis measures, experimentation details, and results are provided in this section.

4.2. Performance Measures

The performance measures are used to identify the effectiveness and reliability of each classification method. The key measures used for identification of the prediction class are True Positive (TP), defined correctly identified, and False Positive (FP), defined correctly rejected. True Negative (TN), defined as Incorrectly Identified and False Negative (FN), defined as incorrectly rejected. Other derived measures to analyze the performance of the classifiers are given below:

4.2.1. Accuracy

Accuracy defines the ratio of correctly classified instances over the available instances. The accuracy measure is provided in equation (6).

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

4.2.2. Precision

Precision is the ratio of True positive and total positive instances identified over the sample space. Equation (7) shows the computation of precision measure.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (7)$$

4.2.3. Recall

Recall rate defines the ratio of actual positives identified for the test set. Equation (8) shows the computation of the recall measure.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (8)$$

4.2.4. F-Score

This measure is defined as the harmonic mean of precision and recall values. F-score can be computed using equation (9).

$$\text{F-Score} = 2\ x\ \frac{Precision\ x\ Recall}{Precision+Recall} \quad (9)$$

4.3. Comparative Results

A comparative evaluation of various machine learning methods for cardiac attack prediction is provided in this section. The analytical analysis parameters are already described in section 4.2. The experimentation is conducted by using 10 fold method. Figure 7 shows the accuracy-based analysis results. The figure shows that the accuracy obtained for decision tree and random forest methods is ineffective, with a lesser accuracy rate of 77.78% and 76.54%. Naïve Bayes and SVM methods achieved the most effective results with 86.42% accuracy. The analysis results show that the SVM and Bayes network can predict cardiac disease more accurately. The neural and RBF neural network-based methods also provided significant results with higher than 80% accuracy.

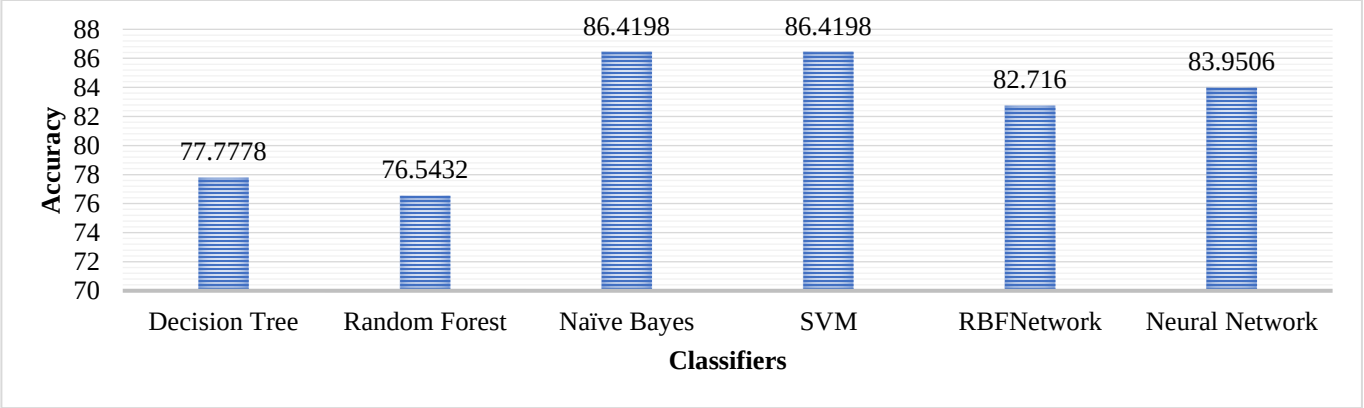

Fig. 7 Accuracy-based analysis

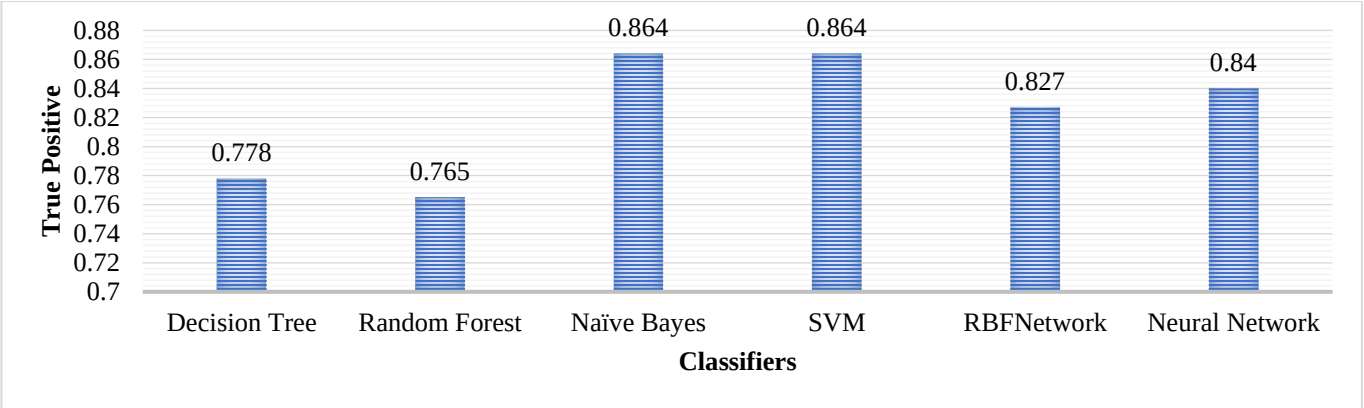

Fig. 8 True positive analysis

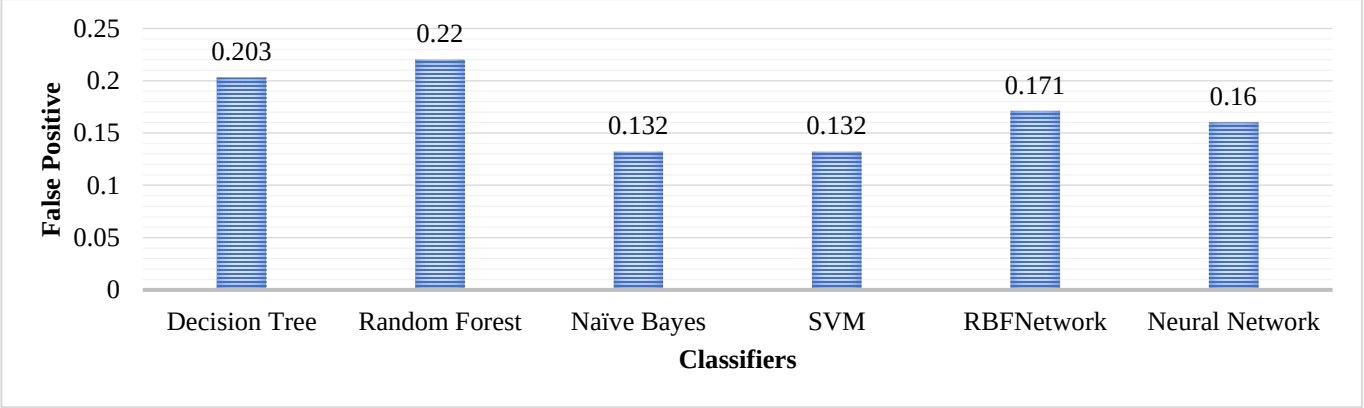

Fig. 9 False positive analysis

Another performance measure considered in this research is true positive analysis. The analysis results on true positives are provided using Figure 8. The figure shows that the true analysis results for decision tree and random forest methods are ineffective with .778 and .765. Naïve bayes and SVM methods achieved the most effective results with a 0.864 true positive rate. The analysis results show that the SVM and Bayes network can predict cardiac disease with the most adaptive and effective true positive rate. The neural and RBF neural network-based methods also provided significant results with a true positive rate higher than 0.8.

False positive identifies the error in the predictive system. The analysis results on false positives are provided using Figure 9. The figure shows that the true analysis results for decision tree and random forest methods are ineffective, with higher false positive rates over 0.2. Naïve bayes and SVM methods achieved the most effective results with the least lease false positive rate of 0.132. The analysis results show that the SVM and Bayes network can predict cardiac disease with the most adaptive and effective false positive rate. The neural and RBF neural network-based methods also provided significant results with a false positive rate higher than 0.15.

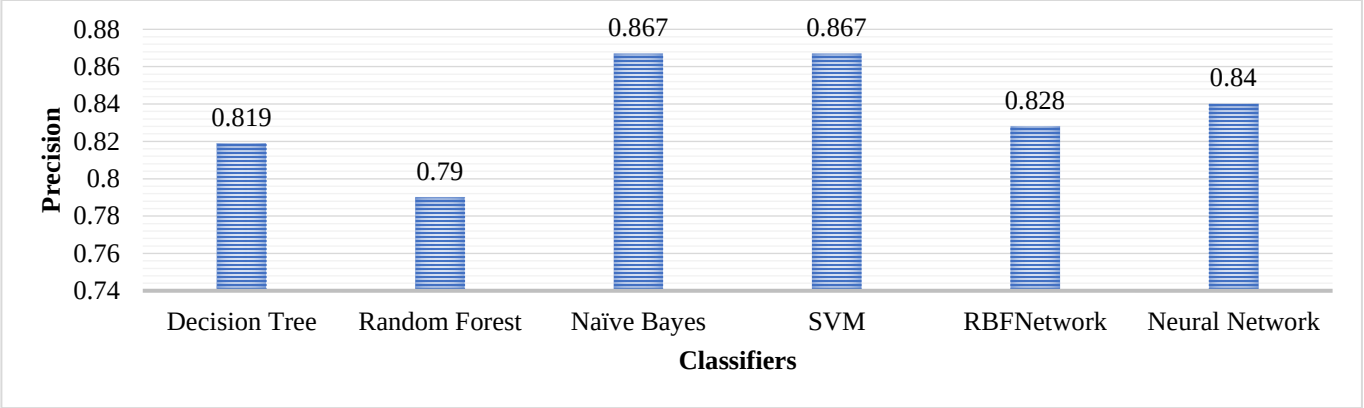

Fig. 10 Precision analysis

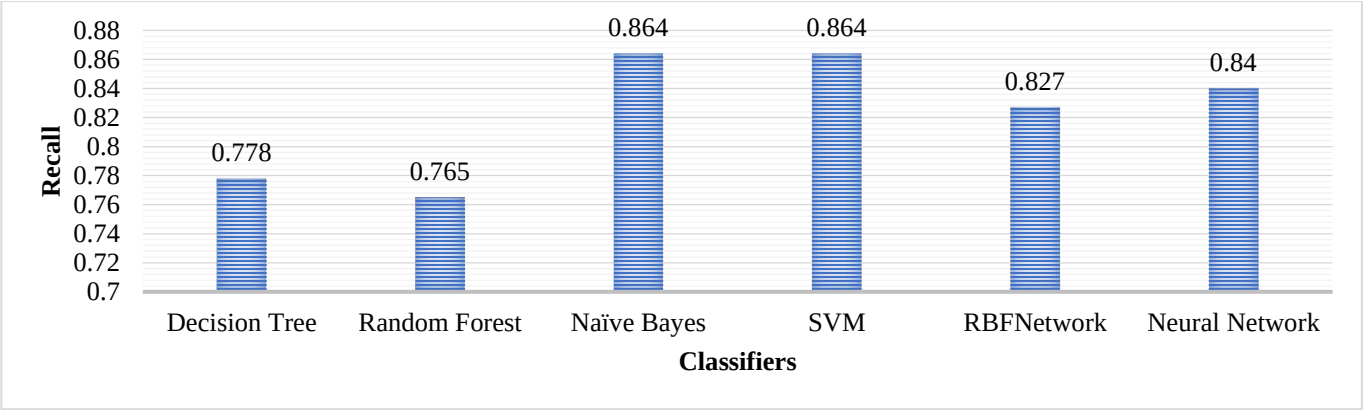

Fig. 11 Recall analysis

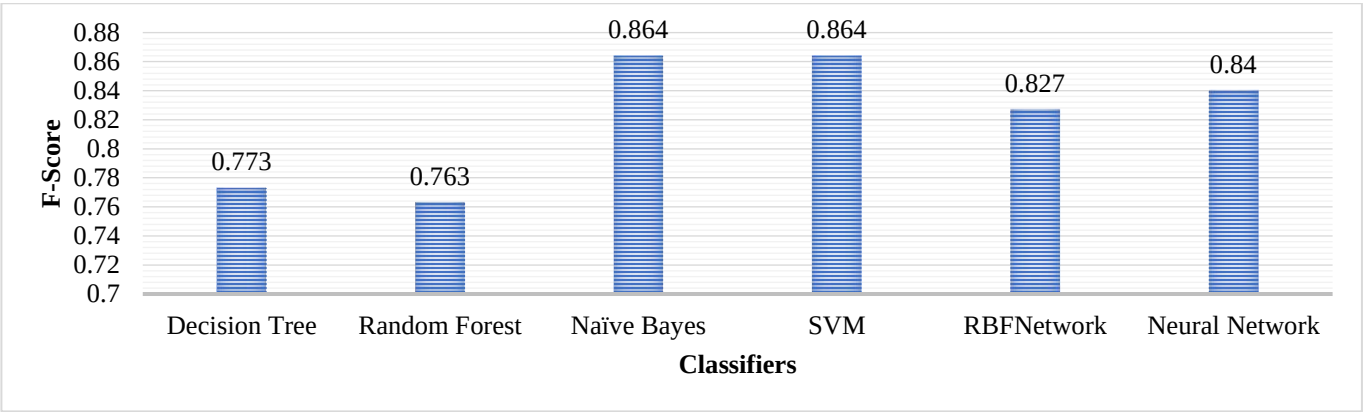

Fig. 12 F-Score measure analysis

Precision is another accuracy-based prediction measure used in this research. The analysis results on precision ratio are provided using Figure 10. The figure shows that the precision rate analysis results for decision tree and random forest methods are ineffective, with lesser precision rates of 0.819 and 0.79. Naïve bayes and SVM methods achieved the most effective results with the least lease precision rate of 0.867. The analysis results show that the SVM and Bayes network can predict cardiac disease with the most adaptive and effective precision rate. The neural and RBF neural network based methods also provided significant results.

The recall parameter-based analysis results are provided in Figure 11. The figure shows that the recall rate analysis results for decision tree and random forest methods are ineffective, with lesser recall rates of 0.779 and 0.765. Naïve bayes and SVM methods achieved the most effective results with the least lease recall rate of 0.864. The analysis results show that the SVM and Bayes network can predict cardiac disease with the most adaptive and effective recall rate. The neural and RBF neural network based methods also provided significant results.

Another effective accuracy-based analysis measure is F-Score. The analysis results based on the F-score parameter are provided in Figure 12. The figure shows that the F-score analysis results for decision tree and random forest methods are the least effective, with 0.773 and 0.763. Naïve bayes and SVM methods achieved the most effective results with the least lease F-score of 0.864.

The analysis results show that the SVM and Bayes network can predict cardiac disease with the most adaptive and effective F-Score rate. The neural and RBF neural network based methods also provided the effective Fscore rate.

5. Conclusion

The diagnosis of cardiac disease at an early stage can reduce the chances of death. Various machine learning methods and models are available to predict cardiac disease. In this paper, a standard model of cardiac disease prediction is defined. Various classifiers are integrated into this model to predict cardiac disease accurately. The classifiers used in this research work are Bayesian network, decision tree, neural network, radial basis network and SVM. The analysis results show that the SVM and Bayesian network classification-based machine learning model achieved the maximum accuracy rate of 86.42. The method also achieved effective results under true positive, true negative, precision and recall measures.

References

[1] Princy R. Jane Preetha et al., "Prediction of Cardiac Disease Using Supervised Machine Learning Algorithms," *4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, pp. 570-575, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[2] Samir S. Yadav et al., "Automated Cardiac Disease Diagnosis Using Support Vector Machine," *3rd International Conference on Communication System, Computing and IT Applications (CSCITA)*, pp. 56-61, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[3] Shaira Tabassum et al., "The Cardiac Disease Predictor: IoT and ML-Driven Healthcare System," *4th International Conference on Electrical Information and Communication Technology (EICT)*, pp. 1-6, 2019. [Google Scholar] [Publisher Link]

[4] Senthilkumar Mohan, Chandrasegar Thirumalai, and Gautam Srivastava, "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques," *IEEE Access*, vol. 7, pp. 81542-81554, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[5] Jian Ping Li et al., "Heart Disease Identification Method Using Machine Learning Classification in E-Healthcare," *IEEE Access*, vol. 8, pp. 107562-107582, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[6] Norma Latif Fitriyani et al., "HDPM: An Effective Heart Disease Prediction Model for a Clinical Decision Support System," *IEEE Access*, vol. 8, pp. 133034-133050, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[7] G. Thippa Reddy et al., "Hybrid Genetic Algorithm and a Fuzzy Logic Classifier for Heart Disease Diagnosis," *Evolutionary Intelligence*, vol. 13, pp. 185-196, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[8] Mohammad Shafenoor Amin, Yin Kia Chiam, and Kasturi Dewi Varathan, "Identification of Significant Features and Data Mining Techniques in Predicting Heart Disease," *Telematics and Informatics*, vol. 36, pp. 82-93, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[9] Apurb Rajdhan et al., "Heart Disease Prediction Using Machine Learning," *International Journal of Research and Technology*, vol. 9, no. 4, pp. 659-662, 2020. [Google Scholar]

[10] Anjan Nikhil Repaka, Sai Deepak Ravikanti, and Ramya G Franklin, "Design and Implementing Heart Disease Prediction using Naives Bayesian," *3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, pp. 292-297, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[11] Amin Ul Haq et al., "A Hybrid Intelligent System Framework for the Prediction of Heart Disease Using Machine Learning Algorithms," *Mobile Information Systems*, vol. 2018, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[12] C. Beulah Christalin Latha, and S. Carolin Jeeva, "Improving the Accuracy of Prediction of Heart Disease Risk Based on Ensemble Classification Techniques," *Informatics in Medicine Unlocked*, vol. 16, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[13] Ankur Gupta et al., "MIFH: A Machine Intelligence Framework for Heart Disease Diagnosis," *IEEE Access*, vol. 8, pp. 14659-14674, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[14] Vaibhav Gupta, and Pallavi Murghai Goel, "Heart Disease Prediction Using ML," *SSRG International Journal of Computer Science and Engineering*, vol. 7, no. 6, pp. 17-19, 2020. [CrossRef] [Publisher Link]

[15] Safial Islam Ayon, Milon Islam, and Rahat Hossain, "Coronary Artery Heart Disease Prediction: A Comparative Study of Computational Intelligence Techniques," *IETE Journal of Research*, vol. 68, no. 4, no. 2488-2507, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[16] Archana Singh, and Rakesh Kumar, "Heart Disease Prediction Using Machine Learning Algorithms," *International Conference on Electrical and Electronics Engineering (ICE3)*, pp. 452-457, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[17] V.V. Ramalingam, Ayantan Dandapath, and M. Karthik Raja, "Heart Disease Prediction Using Machine Learning Techniques: A Survey," *International Journal of Engineering & Technology*, vol. 7, no. 2.8, pp. 684-687, 2018. [Google Scholar] [Publisher Link]

[18] Rahul Katarya, and Sunit Kumar Meena, "Machine Learning Techniques for Heart Disease Prediction: A Comparative Study and Analysis," *Health and Technology*, vol. 11, no. 1, pp. 87-97, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[19] Nandakumar Pandiyan, and Subhashini Narayan, "Prediction of Cardiac Disease using Kernel Extreme Learning Machine Model," *International Journal of Engineering Trends and Technology*, vol. 70, no. 11, pp. 364-377, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[20] Anna Karen Gárate-Escamila, Amir Hajjam El Hassani, and Emmanuel Andrès, "Classification Models for Heart Disease Prediction Using Feature Selection and PCA," *Informatics in Medicine Unlocked*, vol. 19, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[21] Yar Muhammad et al., "Early and Accurate Detection and Diagnosis of Heart Disease Using Intelligent Computational Model," *Scientific Reports*, vol. 10, 2020.[CrossRef] [Google Scholar] [Publisher Link]

[22] Ricardo Buettner, and Marc Schunter, "Efficient Machine Learning Based Detection of Heart Disease," *IEEE International Conference on E-Health Networking, Application & Services (HealthCom)*, pp. 1-6, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[23] S. Sivasubramaniam, and S. P. Balamurugan, "Nature Inspired Optimization with Hybrid Machine Learning Model for Cardiovascular Disease Detection and Classification," *International Journal of Engineering Trends and Technology*, vol. 70, no. 12, pp. 127-137, 2022. [CrossRef] [Publisher Link]

[24] Santhana Krishnan, and S. Geetha, "Prediction of Heart Disease Using Machine Learning Algorithms," *1st International Conference on Innovations in Information and Communication Technology (ICIICT)*, pp. 1-5, 2019. [CrossRef] [Publisher Link]

[25] Hager Ahmed et al., "Heart Disease Identification from Patients' Social Posts, Machine Learning Solution on Spark," *Future Generation Computer Systems*, vol. 111, pp. 714-722, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[26] Aditi Gavhane et al., "Prediction of Heart Disease Using Machine Learning," *Second International Conference on Electronics, Communication And Aerospace Technology (ICECA)*, pp. 1275-1278, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[27] Tougui Ilias, Abdelilah Jilbab, and Jamal El Mhamdi, "Heart Disease Classification Using Data Mining Tools and Machine Learning Techniques," *Health and Technology*, vol. 10, no. 5, pp. 1137-1144, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[28] Online Available : https://archive.ics.uci.edu/dataset/145/statlog+heart

[29] Michael J. Pencina et al., "Quantifying Importance of Major Risk Factors for Coronary Heart Disease," *Circulation*, vol. 139, no. 13, pp. 1603-1611, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[30] Pooja Rani et al., "A Decision Support System for Heart Disease Prediction Based Upon Machine Learning," *Journal of Reliable Intelligent Environments*, vol. 7, no. 3, pp. 263-275, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[31] Harshit Jindal et al., "Heart Disease Prediction Using Machine Learning Algorithms," *IOP Conference Series: Materials Science and Engineering*, vol. 1022, 2021. [CrossRef] [Google Scholar] [Publisher Link]