Early Prediction of Cardiovascular Disease Using Machine Learning Algorithms

Charu Khandelwal, Simran Agarwal, Jyotsna, Deepti Sahu, and Sudeshna Chakraborty

Abstract Cardiovascular disease (CVD) is regarded as one of the world's leading causes of death. Individuals who are dealing with various risk factors such as high blood cholesterol, obesity (overweight), hypertension, and diabetes are more susceptible to CVD and thus need early detection. Advancements of technologies are assembling terabytes of data every day from the healthcare industry to keep records. However, this data is not mined well to anticipate the likelihood of a patient getting a cardiopulmonary arrest. Therefore, with the assistance of disparate machine learning and data mining techniques, it is feasible to extract useful insights and discover hidden patterns from the datasets to get a more accurate diagnosis and decision-making. The paper aims to review different research papers with comparative results that have been done on the prognostication of CVD to get an integrated, synthesized overview of machine learning techniques, their performance measures in several datasets and to also make vital conclusions. From the study, we observed that various techniques such as decision trees and artificial neural network (ANN) give the highest CVD prediction system accuracy in different scenarios. This procedure could possibly be useful for cardiologists to forecast the occurrence of cardiovascular disease beforehand and come up with proper medical treatment.

Keywords Data mining · Machine learning · Machine learning techniques · Cardiovascular disease · Decision tree

1 Introduction

The human body consists of internal organs like the brain, heart, liver, lungs, kidney, etc., that are very liable to diseases but among these organs, the heart is considered as the most vital and more prone to disease. Despite leading a healthy lifestyle, the propensity of having CVD due to plaque buildup in arteries causing narrowed blood

---

C. Khandelwal · S. Agarwal · Jyotsna (✉) · D. Sahu (✉) · S. Chakraborty
Department of Computer Science, Sharda University, Noida, India

S. Chakraborty
e-mail: sudeshna.chakroborty@sharda.ac.in

vessels resulting in stroke, hypertension, chest pain, arrhythmia, and other symptoms is frequent [1]. According to WHO, millions of people die every year due to cardiac disease. Electrocardiogram (ECG) plays a crucial role in quickly monitoring the heart's health and detecting the electrical signals produced by it. Heart rates differ widely from person to person depending upon their lifestyle. It's considered to be lower when at rest and higher during exercise. The standard heart rate is typically 60–100 bps (beats per minute). Lower resting heart rate is an indication of good health and wellness. Many doctors suggest a proper nutritious diet and regular exercise to keep the heart healthy. Figure 1 depicts the overview of the 3 phases with the help of ECG signals which shows that heartbeat is moderate.

A lot of factors such as cholesterol level, age, smoking habits, diabetes, genetic mutations, and pulse rate contribute to heart disease. Identifying people at risk of CVD is a cornerstone of preventive cardiology. With the constant boom of data in the healthcare industry, their collection techniques are getting ameliorated on a daily basis through the application of wearable technology and the Internet of Things (IoT). It is impossible for a human being to amalgamate zillions of facts and data and infer the specific patient's malady. Nevertheless, machine learning could be applied as a predictive mechanism to find the insights and patterns in the data [5]. The datasets are largely collected via Kaggle and the UCI machine learning repository.

The robust strategies related to deep learning and machine learning techniques help to foretell the people who are in the similitude of getting the cardiac disease an accurate time, provide affordable services, and save precious lives. These algorithms and techniques can be directly applied to a dataset using various machine learning
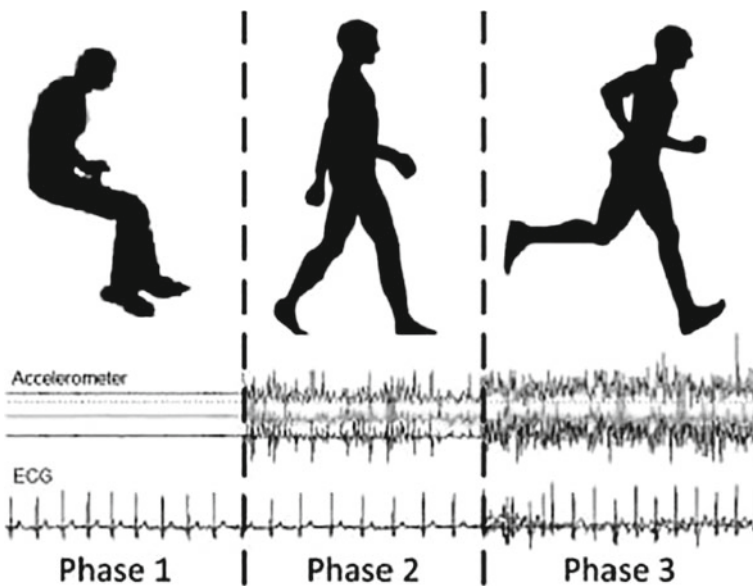

Fig. 1 Phase1: standing; Phase 2: walking; and Phase 3: running [6]

frameworks to draw analytical conclusions. Even though different types of CVD usually have a lot of different symptoms, many have identical warning signs such as respiratory infection, irregular heartbeat, dizziness, loss of appetite, and restlessness. It is a fatal disease which gives rise to more adversities. Table 1 summarizes the different types of CVD.

This article structure is as follows—Sect. 2 presents the various machine learning algorithms used, Sect. 3 of the paper gives a comprehensive literature review, Sect. 4 presents the summary, and Sect. 5 provides a conclusion.

Table 1 Different types of cardiovascular disease, its symptoms, and risk factor

| Cardiovascular diseases	Brief description	Warning symptoms	Risk factors of disease
Rheumatic heart disease	Inflammatory disease, heart valves are permanently damages, tonsils in children	Nodules, chest pain, fever, dyspnea, complications during pregnancy	Untreated or under-treated strep infections; over-crowding; age group of 5–15 years are more prone to this disease
Coronary artery disease (atherosclerosis)	when the coronary arteries become too hard and narrow, restricting flow of blood to the heart	Nausea, angina, chest pain, discomfort in the back, jaw, shortness of breath	Smoking, poor nutritious diet, laziness, diabetes, hypertension
Peripheral vascular disease (claudication)	Flow of blood decreases causing blood vessels to block outside of the brain and heart, hardening of arteries	Leg, thigh, buttocks, calf pain or cramping, loss of hair on legs, muscles feel numb	Diabetes, overweight, history of stroke, aging high blood pressure, kidney disease, sedentary lifestyle, drug use, high cholesterol
Deep vein thrombophlebitis (DVT)	Clots of blood in a deep vein, generally occurs in the legs which can dislodge and move to the heart and lungs	Pain, swelling, and tenderness in one leg; warm skin around the areas of the clot; skin turning to a reddish color over the affected area	Long bed rest, injury to veins, obesity, bowel disease, smoking
Stroke (brain attack)	Blood supply gets reduced to part of our brain	Weakness of face and arm, sudden headache, unstable walking	High blood pressure, alcohol intake in excess amount, family history, atrial fibrillation, vision loss
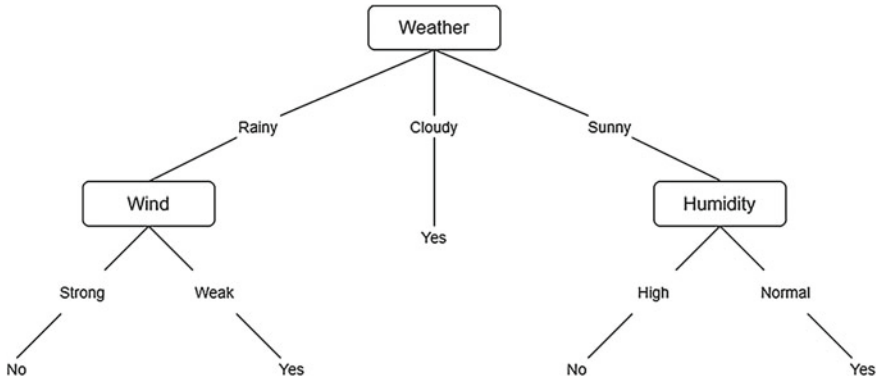
Fig. 2 Decision tree, tree-like structure

2 Exploratory Study of Various Machine Learning Algorithms

2.1 Decision Trees

Decision trees are a type of construction that distributes a large number of records into smaller groups using a set of simple decision rules. It handles both continuous and categorical variables and they are mostly used for classification problems. In decision trees, each tree comprises nodes and branches where every node serves as an attribute in a group for classification, and every branch will act as a value that the node can take. Decision trees aren't very robust on their own but are used in other methods that leverage their simplicity and create some very powerful machine learning algorithms. Recently, it is reborn with new upgrades and those upgrades are advanced methods such as gradient boosting and random forest that build on top of decision trees. The main advantage of decision trees is that without requiring much computation, it can provide a clear indication of which fields are more important for prediction. Figure 2 shows the tree-like structure of a decision tree.

2.2 Support Vector Machine

Support vector machine (SVM) comprises learning models that are supervised in nature. It's mostly used for classification and works on the concept of margin calculation. It finds the best line (decision boundary) that helps to separate space into classes. This line is searched through the maximum margin which is equidistant from the points on both sides. The sum of distance of these two points from the line has to be maximized to get the best result. This line is called maximum margin hyperplane or classifier, and the two points are called support vectors which support
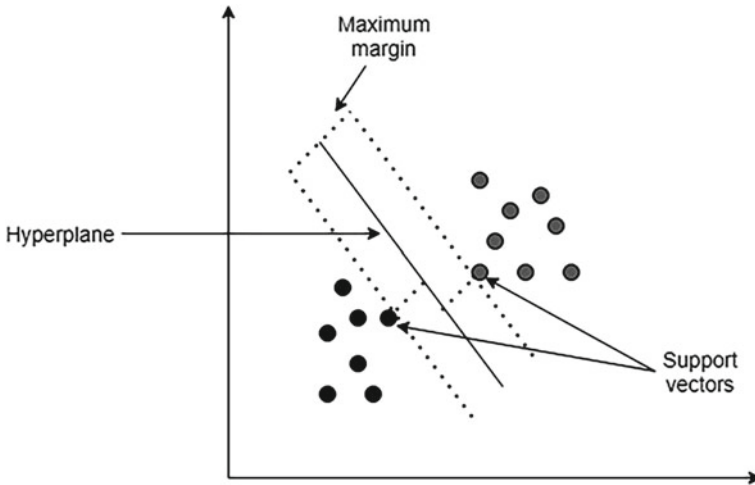
Fig. 3 Support vector machine

the whole algorithm. SVM is better to be thought of as a more rebellious and risky type of algorithm because it looks at extreme points that are very close to the decision boundary, and uses it to construct analysis. That in itself makes the SVM algorithm very special and different from other machine learning algorithms. Figure 3 shows the example of SVM.

2.3 Random Forest

Random forest (random decision forest) is a technique that constructs and integrates various decision trees to get a more stable and accurate prognosis. The random forest works using an algorithm as discussed.

- Initially, random K data points will be taken out of the train set.
- The decision tree will be created using K data points.
- Select the number N-tree of trees and follow the above two steps.
- To select a new data point, create N-tree trees for each data point to predict the class to which it belongs and allocate it to the class that wins the majority vote.

The random forest starts with one tree and builds it up to n-number, which is randomly selected from the data. Though each one of these trees might not be perfect, overall on average it can perform very well, and therefore it is a major advantage of this algorithm. Figure 4 shows an example of a random forest.
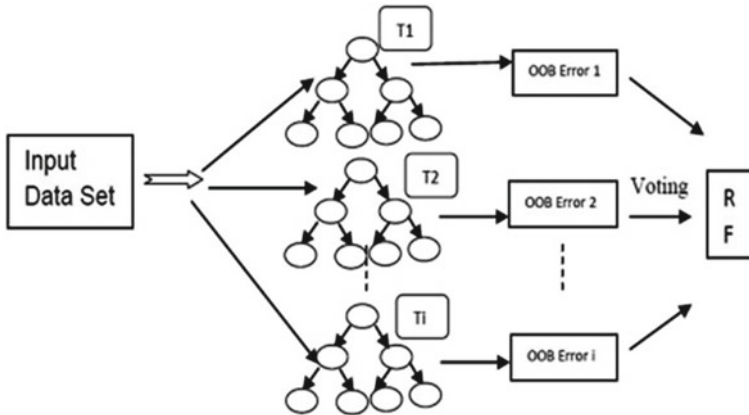
Fig. 4 Random forest [7]

2.4 Naive Bayes

In statistics, Bayes' theorem (Bayes' law) follows the maxim of the conditional and marginal probabilities of two sets of random outcomes from an experiment. It is a family of algorithms which sticks to the standard theory, that each set of features which is classified is free from each other. It is often used to calculate posterior probabilities. Naive Bayes handles both discrete and continuous data. The Bayes model in combination with decision rules contributes to probability independently like hypotheses and studies without considering correlations. The probability model of naive Bayes can be efficiently trained using supervised learning algorithms. It is insensitive to irrelevant features and doesn't require as much training data. Naive Bayes can solve diagnostic problems as it aids to specify if a patient is at high risk of certain diseases.

2.5 Artificial Neural Network

Artificial neural networks (ANN) are vaguely prompted by the biological neural circuit. It is considered to be the most useful and powerful machine learning algorithm. An input layer, a concealed layer, and a final output layer make up the three layers. The input layer takes the input which is assessed by the hidden layer. Finally, the output layer sends the calculated output. A multilayer perceptron (MLP), a form of feedforward ANN, is the most common type of neural network. It uses a supervised learning technique called backpropagation (backward propagation of errors) for training. ANN can find complicated patterns in data and thus improve its performance. It has a feature where the failure of one or more cells does not prevent it from generating results, thus making the networks fault-tolerant. It can perform more than

Fig. 5 Naive Bayes
classifier
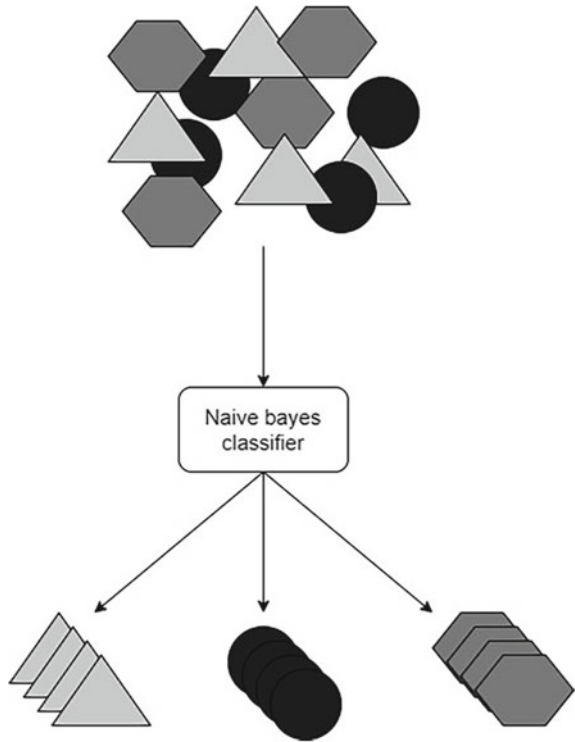
one job at the same time, making it a widely used algorithm for solving complex problems. Figure 5 shows the actual structure of ANN.

2.6 K Nearest Neighbor

K-Nearest Neighbor (KNN) is a kind of lazy learning or non-parametric learning that uses a pliable number of parameters where a function is estimated locally and all enumeration is withheld until classification. It is a smooth algorithm. To get a better understanding of KNN, let's walk through the steps:

- Pick the number K of neighbors. The most common default for K is 5.
- Select the K closest neighbors of the new data point using Euclidean distance. Distances such as Manhattan distance can also be applied.
- Compute the number of data points that fall into each category.
- Assign the newly created data point to the category with the most neighbors.

KNN algorithm fares across all parameters of deliberation. It is frequently used for its low computation time and trouble-free implementation as only the value of
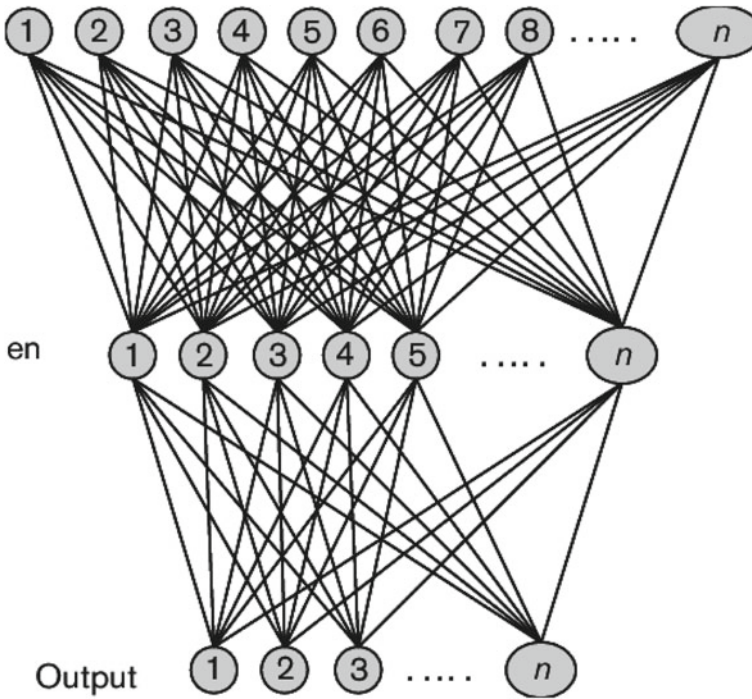
Fig. 6 Artificial neural network of multiple layers and outputs [9]

k and the distance function are required to execute. Figure 6 shows an example of KNN.

2.7 Logistic Regression

A logistic regression model is used to perform predictive analysis (modeling) to estimate the probability of a given output based on input variables, in contrast with a binary classifier. It is incredibly simple to execute and very efficient to train. It has a good baseline that one can use, to compute the performance of other advanced or complex algorithms. It is a valuable model to be selected when different sources of data are combined into a binary classification task. Since a linear model does not extend to classification problems with multiple classes, logistic regression is considered as a solution for it. Figure 7 shows the logistic function graph. The algorithm compresses the outcome of a linear (algebraic) mathematical equation between 0 and 1 using the logistic function.
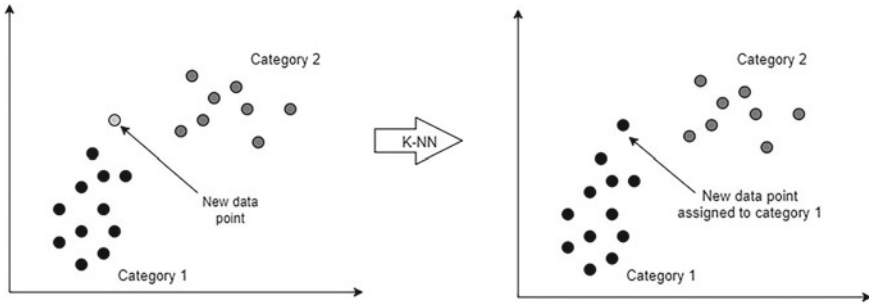
Fig. 7 K nearest neighbor

2.8 *Gradient Boosting*

Gradient boosting is considered as one of the most robust techniques for constructing predictive models due to its high speed and accuracy. It follows a greedy approach and produces a highly robust solution for both classification and regression problems. It requires three elements to function—loss function, weak learner, and an additive model. A loss function is optimized. To generate predictions, a weak learner is used, and an additive model. In this, the additive model is used to add on the weak learner to minimize the loss function, and finally lower the overall prediction error. It integrates the previous one with the best possible next model. In other words, it tries to develop a new sequential model. Gradient boosting is commonly used as it is generic enough to use any differentiable loss function. Figure 8 depicts the working of gradient boosting.
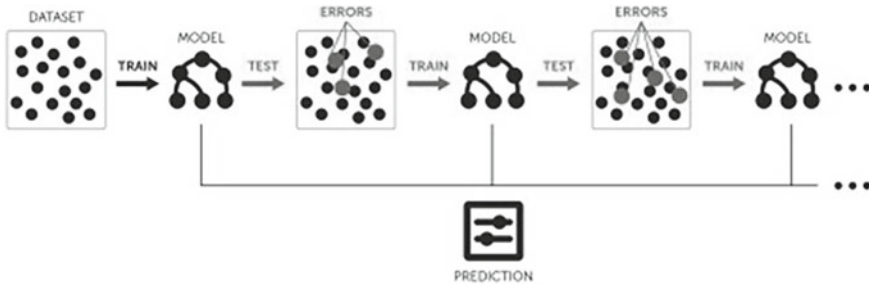
Fig. 8 The logistic function
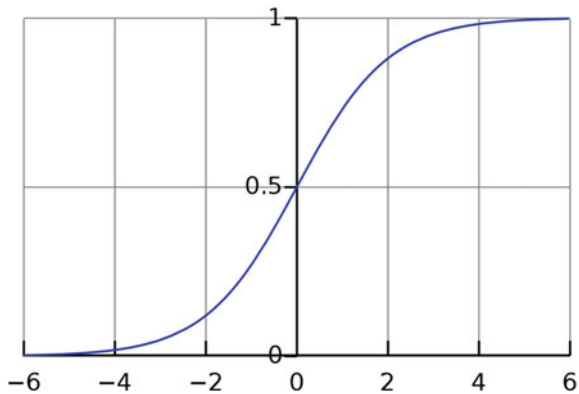
Fig. 9 Working of gradient boosting

2.9 Rough Set

Rough set theory has engaged the attention of many research workers and practitioner workers throughout the world. Methods build on a rough set have broad applications in many real-life projects. It can find a minimal set from data for dimension reduction in classification. They set interconnection with many other approaches such as statistical methods and fuzzy set theory. A rough set solves the problems such as finding the dependency between the most significant attributes, reducing the surplus one, and describing a set of objects based on attribute values. It is widely used for feature extraction, feature selection, decision rule generation, and also for discovering hidden patterns inside the data. Thus, it is valuable to mention that a rough set plays a crucial role in solving prediction problems. Figure 9 connotes the rough set theory concept where the lower and upper approximation sets are known to be crisp sets, and the same sets can also be called fuzzy sets these sets in other variations.

3 Literature Survey

In this paper, the dataset is subjected to a variety of machine learning methods to prognosticate the likelihood of a patient getting cardiac arrest based on various controlled and uncontrolled variables [1]. Parameters, namely age, blood pressure, alcohol intake, gender, chest pain, fasting blood sugar, cholesterol, etc., are considered for predictions of CVD. Initially, the dataset contains some missing records which are recognized and replaced with the most relevant values. The missing values are calculated using the mean method. After preprocessing the data, classification algorithms like SVM, decision tree, and ANN are applied to the dataset. Due to the broad range of relevancy of ANN and its capability to understand advanced or complex relationships along with modeling of non-linear processes, an ANN algorithm is considered as the best performing algorithm with accuracy 85.00%. The conclusion made during this study was that the accuracy of ANN could be more

Fig. 10 Rough set theory
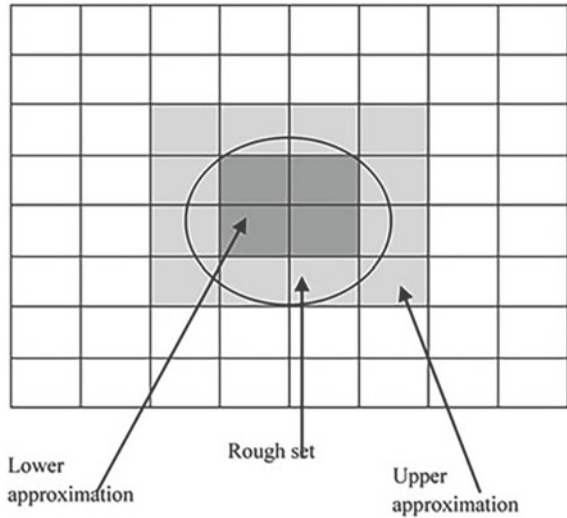
precise if a larger dataset is used. Figure 10 indicates the accuracy across various algorithms.

In this paper, support vector classifier, logistic regression, and decision tree are presented to forecast CVD using machine learning paradigms with 301 sample data, and 12 attributes [2]. The entire data is required to split into two parts, one set for training comprises 80% of total are split into two sets and other for testing with 20% of total data. Data visualization techniques are also applied to extract the hidden insights from the dataset which would help doctors to analyze the pattern effectively for further medical diagnosis. Performance assessment is carried through these four algorithms and their accuracy is deliberated. Moreover, while analogizing these classification algorithms, the outcome reveals that the performance of logistic regression is better than the other three algorithms. The precision, recall, F1-score, and support are also calculated for logistic regression. Later, a comparative study is also performed with the UCI dataset using the same algorithms where the support vector classifier provides better results with an accuracy of 86.1%. Figure 11 shows the performance of the algorithm on two different datasets (Fig. 12).

The prediction of cardiac disease using machine learning techniques has been proposed [3]. The dataset is taken from the UCI repository with 13 medical parameters such as blood pressure level, and electrocardiographic results as input. Python programming is used as a tool for data analysis and machine learning paradigm. Data preprocessing is applied to transform the unrefined data into a comprehensible format. The dataset is divided into two parts, 70% for training while 30% for testing. A scatter plot is applied to both the training and test sets to represent patients having heart disease or not. Two classification-based machine learning techniques, naive Bayes and decision tree, are used. Though naïve Bayes can handle enormous, tangled, non-linear dependent data, decision trees perform better with an accuracy
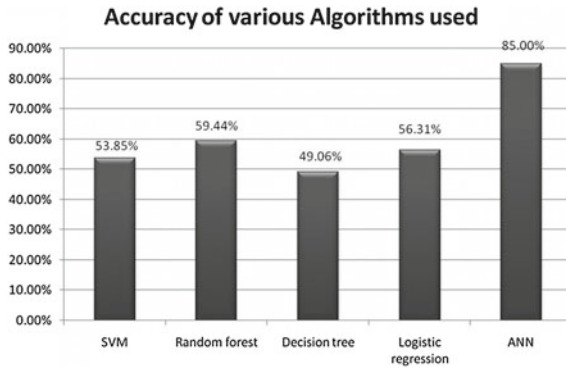
Accuracy of various Algorithms used

Fig. 11 Accuracy of different algorithms

Algorithm accuracy on two different datasets

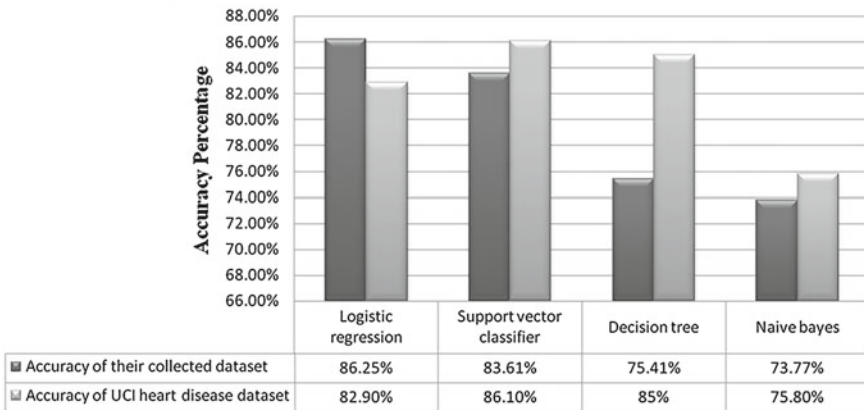|	Logistic regression	Support vector classifier	Decision tree	Naive bayes
Accuracy of their collected dataset	86.25%	83.61%	75.41%	73.77%
Accuracy of UCI heart disease dataset	82.90%	86.10%	85%	75.80%

Fig. 12 Algorithm accuracy on two different datasets

level of 91% as this model analyzes the dataset in the tree shape structured format because of which each attribute is completely analyzed.

CVD prediction using machine learning techniques is discussed. A dataset of cardiac disease has been taken from the UCI repository consisting of 14 attributes as input. R language is used as it has the best compatibility with UNIX and Windows and also proffers a better outcome compared to other languages. Data preprocessing is applied to make the mining process more efficient and to avoid fault prediction. The records are classified into a training and a testing dataset. The system also demonstrates powerful visualization using a box plot, scatter plot, and mosaic plot of interrelation and traits of all the attributes for the graphical representation of data. Then, for prediction, SVM, naive Bayes, random forest, logistic regression, and gradient boosting are used. The analogizing of classification algorithms is made
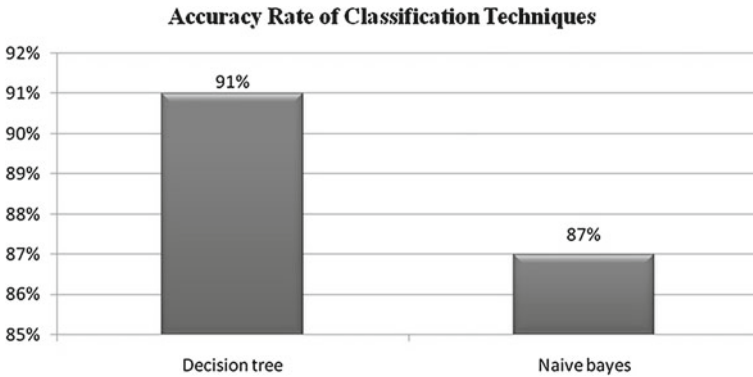
Accuracy Rate of Classification Techniques

Fig. 13 The accuracy rate of the two models

which signifies that the best performing algorithm is logistic regression. The user interface is designed where the parameters of the patients such as type of chest pain, height, age (in years), resting blood pressure, and cholesterol are recorded, and based on the algorithm, the interface of the system calculates the patient's risk of heart disease.. Figure 13 represents the accuracy of each algorithm tested.

The detection of CVD using a new ensemble classifier is proposed [4]. On a dataset acquired from the UCI laboratory, classification-based machine learning techniques such as decision trees, naive Bayes, multilayer perceptron neural network with hidden layers, and rough set are deployed. Information of 303 patients is collected having total features as 76 where the filtering method called Pearson's correlation coefficient is applied to select the most discriminative features. Hence, 14 attributes such as age, cholesterol, and fasting blood sugar are used in this dataset for prediction. Later applied data preparation, where the missing rate of each feature is calculated. The data is assessed using tenfold cross-validation. Performance metrics such as sensitivity, precision, F-measure, and accuracy are calculated with the usage of a confusion matrix. Here, F-measure combines sensitivity and precision into a single value. naïve Bayes, rough set, and neural network achieve the highest performance on the basis of F-measure. Later, fusion strategy is applied to combine these three best classifiers by weighted majority vote which further improves the accuracy of the model. The statistically revelatory difference in assessing the performance of the classifier is observed by the fusion of outputs which further enhance decision support. F-measure achieves 86.8% outperforming other individual metrics, in the domain of classification. Figure 14 shows the performance of the classifier.

In this research, supervised machine learning methods such as naive Bayes, decision tree, logistic regression, and random forest are used to predict the illness related to cardiac problems [5]. The dataset of cardiac disease patients is taken from Kaggle with 12 essential attributes such as systolic and diastolic blood pressure, chest pain, gender, cholesterol, smoker/drinker, and age to forecast the likelihood of patients developing heart illness. Furthermore, the dataset has been divided into two parts with 70% and 30% of total data used for training and testing, %respectively. The
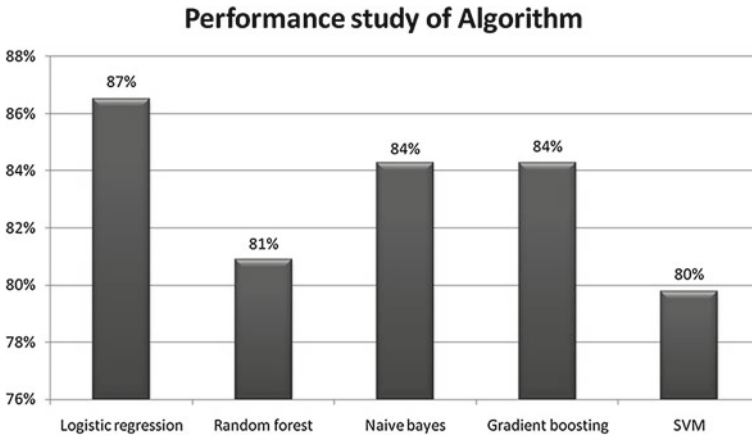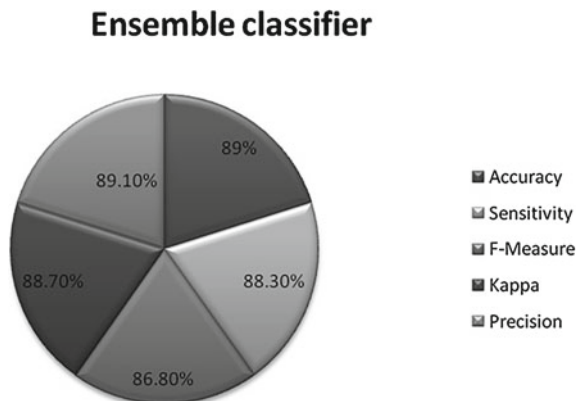
Fig. 14 Performance study of algorithm

confusion matrix (error matrix) is exerted which shows the correlation between all available features and with the help of it and the classification algorithm's precision, recall, F1-score, and accuracy are calculated. The performance-based model is estimated and their results are examined. The testing results show that the decision tree algorithm provides a superior forecast than the other algorithms, with an accuracy rate of 73%. The author later used the technique of dimensionality reduction, where the entities which are negatively correlated are skimmed from the dataset and then tested. As a result, the accuracy of the random forest and KNN algorithm changed either positively or negatively. However, precision values of the decision tree algorithm, before and after the dimensionality reduction, remain the same. It gives the highest accuracy of 73% in both the cases. Figure 15 compares the accuracy (pre- and post-dimensionality reduction) between the algorithms (Fig. 16).

Fig. 15 Evaluation of the proposed ensemble classifier's performance

Comparison of Accuracy

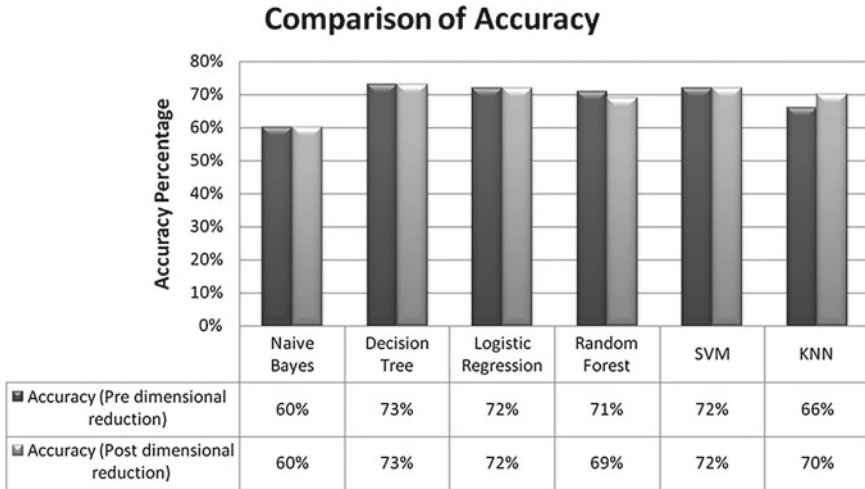|	Naive Bayes	Decision Tree	Logistic Regression	Random Forest	SVM	KNN
■ Accuracy (Pre dimensional reduction)	60%	73%	72%	71%	72%	66%
▣ Accuracy (Post dimensional reduction)	60%	73%	72%	69%	72%	70%

Fig. 16 Comparison of accuracy

4 Comparative Analysis of Machine Learning Technique

Table 2 shows different machine learning techniques used on cardiovascular disease predictions with accuracy.

5 Conclusion

The research examines a variety of machine learning techniques for estimating the total number of CVDs. Machine learning takes leverage of structured and unstructured data sources and therefore plays a crucial role in the healthcare industry. From this study, it shows that decision tree delivers better prediction by providing 91% accuracy consisting of 14 clinical parameters. ANN has also performed well with 85% accuracy. Therefore, we conclude that different methodologies used give different accuracies depending upon the type of dataset taken and tools used for implementation. It is also crucial to note that each domain is non-identical thereby, it is foremost to endeavor various data optimization techniques to escalate the accuracy of the model.

There are numerous upgrades that could be explored in order to increase the system's performance. Hereby, we recommend some of the following observations that need to be considered in future research work to get a more accurate diagnosis of CVD by using a robust prediction system.

- Real data of patients from medical organizations can be incorporated in a large quantity to increase the accuracy of the prediction model.

Table 2 Comparison of machine learning techniques

Author and year	Purpose	Techniques used	Accuracy/Result	
Esfahani et al. [4]	Cardiovascular disease prediction using a new ensemble classifier		Accuracy	F- Measure
		Neural network	86.9%	86.1%
		SVM	75.4%	80.3%
		Naive Bayes	85.4%	84.6%
		SVM (Poly Function)	75.4%	80.1%
		Decision tree	74.8%	77.8%
		Rough set	88.1%	85.7%
Kumar et al. [1]	Performance-based analysis of machine learning paradigms for prediction of cardiovascular disease	Logistic regression	87%	
		Random forest	81%	
		Naive Bayes	84%	
		Gradient boosting	84%	
		SVM	80%	
Krishnan [3]	Prediction of heart disease using classification-based data mining techniques	Decision tree	91%	
		Naive Bayes	87%	
Chauhan et al. [1]	Cardiopulmonary arrest prediction based on various variables using machine learning algorithms	SVM	53.85%	
		Decision tree	49.06%	
		Random forest	59.44%	
		Logistic regression	56.31%	
		ANN	85.00%	
Islam et al. [2]	Efficient forecasting of cardiovascular disease using machine learning paradigms and comparing the accuracy of their collected dataset with that of UCI heart disease dataset		Accuracy of collected dataset	Accuracy of UCI dataset
		Logistic regression	86.25%	82.9%
		Support vector classifier	83.61%	86.1%
		Decision tree	75.41%	85%
		Naive Bayes	73.77%	75.8%

(continued)

Table 2 (continued)

Author and year	Purpose	Techniques used	Accuracy/Result	
Princy et al. [10]	Prognosis of cardiac disease prediction using supervised machine learning algorithms		Accuracy (Pre-dimensional reduction)	Accuracy (Post-dimensional reduction)
		Naive Bayes	60%	60%
		Random forest	71%	69%
		Decision tree	73%	73%
		SVM	72%	72%
		Logistic Regression	72%	72%
		KNN	66%	70%
Akella et al. [9]	Build a predictive model for the detection of coronary artery disease (CAD)	Neural Network	93.03%	
		KNN	84.27%	
		Generalized linear model	87.64%	
		Decision tree	79.78%	
		SVM	86.52%	
		Random forest	87.64%	
Chowdhury et al. [4]	Heart diagnostics Diagnosis machine learning techniques	Decision tree	87.546	
		Logistic regression	86%	
		KNN	87%	
		Naïve Bayes	84%	
		SVM	91%	

- There is a lag between data collection and preprocessing which needs to be addressed.
- Consulting a highly experienced doctor in cardiology will help to prioritize the attributes and to add more vital parameters of cardiac disease for better prediction.
- There is a need to apply more feature extraction and feature selection methods to improve the accuracy performance of the algorithms.
- To lower the overall prediction error, more complex hybrid models should be designed by integrating diverse machine learning and data mining techniques.

- The genetic algorithm is one of the finest and simplest random-based evolutionary algorithms that can be used for optimization which makes the overall performance of intelligent prediction models better.
- To evaluate data in a clinical setting and for better comparison insights in the future study, new analytical frameworks and methodologies, such as regression, association rule, and clustering algorithm, are needed.

References

1. Chauhan U, Kumar V, Hauhan V, Tiwary S, Kumar A (2019) Cardiac arrest prediction using machine learning algorithms. In: 2019 2nd international conference on intelligent computing, instrumentation and control technologies (ICICICT) , pp 886–890. https://doi.org/10.1109/ICICICT46008.2019.8993296
2. Islam S, Jahan N, Khatun ME (2020) Cardiovascular Disease Forecast using Machine Learning Paradigms. Fourth International Conference on Computing Methodologies and Communication (ICCMC) 2020:487–490. https://doi.org/10.1109/ICCMC48092.2020.ICCMC-00091
3. Krishnan GS (2019) Prediction of heart disease using machine learning algorithms. In: 2019 1st international conference on innovations in information and communication technology (ICIICT), pp 1–5
4. Esfahani HA, Ghazanfari M (2017) Cardiovascular disease detection using a new ensemble classifier. In: 2017 IEEE 4th international conference on knowledge-based engineering and innovation (KBEI), pp 1011–1014. https://doi.org/10.1109/KBEI.2017.8324946
5. Princy RJ, Parthasarathy S, Hency PS, Raj Lakshminarayanan A, Jeganathan S (2020) Prediction of cardiac disease using supervised machine learning algorithms. In: 2020 4th international conference on intelligent computing and control systems (ICICCS), pp 570–575. https://doi.org/10.1109/ICICCS48265.2020.9121169
6. Heurtefeux K, Hamida E, Menouar H (2014) Design and implementation of a sustainable wireless BAN platform for remote monitoring of workers health care in harsh environments. In: 6th international conference on new technologies, mobility and security (NTMS), pp 1–5. https://doi.org/10.1109/NTMS.2014.6814014
7. Patel S, Jokhakar VN (2016) A random forest based machine learning approach for mild steel defect diagnosis. In: 2016 IEEE international conference on computational intelligence and computing research (ICCIC), pp 1–8. https://doi.org/10.1109/ICCIC.2016.7919549
8. Enzo G, Massimo B (2008) Introduction to artificial neural networks. Eur J Gastroenterol Hepatol 19: 1046–1054. https://doi.org/10.1097/MEG.0b013e3282f198a0
9. Aravind A, Akella S (2021) Machine learning algorithms for predicting coronary artery disease: efforts toward an open-source solution. Future Sci OA 7: FSO698. https://doi.org/10.2144/fsoa-2020-0206