

Comparison of result of machine learning algorithms in predicting heart disease

Sajad Yousefi^{1*}, Maryam Poornajaf²

¹Faculty Member, Department of Electrical Engineering, Technical and Vocational University (TVU), Tehran, Iran

²Faculty Member, Department of Computer Engineering, Technical and Vocational University (TVU), Tehran, Iran

Article Info

Article type:
Research

Article History:
Received: 2022-12-25
Accepted: 2023-04-16
Published: 2023-04-18

*** Corresponding author:**
Sajad Yousefi

Department of Electrical
Engineering, Technical and
Vocational University (TVU), Tehran,
Iran

Email: syosefi_1980@yahoo.com

Keywords:
F1-Score
Machine Learning
Heart Disease
Classification
Importance Score
Accuracy

ABSTRACT

Introduction: Heart disease is, for the most part, alluding to conditions that include limited or blocked veins that can prompt a heart attack, chest torment or stroke. Earlier identification of heart disease may reduce the death rate. The cost of medical diagnosis makes it perverse to cure it for the large amount of people early. Using machine learning models performed on dataset. This article aims to find the most efficient and accurate machine learning models for disease prediction.

Material and Methods: Several supervised machine learning algorithms were utilized to diagnosis and prediction of heart disease such as logistic regression, decision tree, random forest and KNN. The algorithms are applied to a dataset taken from the Kaggle site including 70000 samples. In algorithms, methods such as the importance of features, hold out validation, 10-fold cross-validation, stratified 10-fold cross-validation, leave one out cross-validation are the result of effective performance and increase accuracy. In addition, feature importance scores was estimated for each feature in some algorithms. These features were ranked based on feature importance score. All the work is done in the Anaconda environment based on python programming language and Scikit-learn library.

Results: The algorithms performance is compared to each other so that performance based on ROC curve and some criteria such as accuracy, precision, sensitivity and F1 score were evaluated for each model. As a result of evaluation, random forest algorithm with F1 score 92%, accuracy 92% and AUC ROC 95%, has better performance than other algorithms.

Conclusion: The area under the ROC curve and evaluating criteria related to a number of classifying algorithms of machine learning to evaluate heart disease and indeed, the diagnosis and prediction of heart disease is compared to determine the most appropriate classifier.

Cite this paper as:

Yousefi S, Poornajaf M. Comparison of result of machine learning algorithms in predicting heart disease. Front Health Inform. 2023; 12: 135. DOI: [10.30699/fhi.v12i0.402](https://doi.org/10.30699/fhi.v12i0.402)

INTRODUCTION

Heart disease is one of the most common diseases nowadays, and an early diagnosis of such a disease is a crucial task for many health care providers to prevent their patients and save lives. Tiny malfunction in the heart may cause numerous heart diseases, it can be damaged or defective in one of the heart valves, because of some infection, due to an unhealthy diet, lack of exercise, overweight, or smoking. On the other hand, sometimes few people born with defects in the heart. Some heart deceases more dependent on gender also and found more in males. Heart disease causes narrowing of the arteries

and arteries that carry blood due to the abnormal condition of the blood vessels in the heart caused by deposits. Impaired blood flow in the arteries causes a heart attack and stroke. Components such as high blood pressure, high blood cholesterol levels, high weight, and alcohol consumption are the causes of heart disease. Heart disease can be predicted by analyzing the datasets that have been updated and examining the components affecting heart disease. Various work is done on the different heart data sets. Using machine learning models, a large amount of data can be analyzed and patterns between data can be used to predict [1]. In machine learning, performance evaluation metrics are processed for

data sets. The researchers used various data mining methods such as classification and clustering to create a model for predicting heart disease. Researchers have applied different data mining methods such as association rules, classification and clustering to build a model for the prediction of heart disease.

With increasing research in the field of healthcare along with advanced machine learning, various experiments and researches were carried out in the last few years that provides significant information about the potential of modern day technologies in the healthcare sector [2].

In this paper, the classification of machine learning techniques and algorithms are used to increase the F1-score metric, accuracy metric and the area under the received operating curve (ROC). In machine learning (ML), classification algorithms are supervised learning approach in which the computer learns from the input data and learn from it. Here we investigated how such cheap and simple algorithms might be enough utility to use clinically, and point the way to improved services.

MATERIAL AND METHODS

Classification

Classification is a method of machine learning and is used to learn how to assign a class tag to an input instance. For example, classification can determine whether a person is ill. Class labels must be converted to numeric values. Two classes are considered: class zero (having heart disease) and class one (not having heart disease). Classification is actually a predictive issue that predicts class labels. The data set records under analysis are divided into two categories: training set and test set. The individual records that make up the test dataset are randomly sampled from the r set under analysis. Test data set records are independent of training records. Due to the presence of the target variable and two classes the diagnosis of heart disease has been classified as supervised learning, therefore, supervised learning algorithms such as K-nearest neighbor (KNN), logistic regression (LR), support vector machines (SVM), decision tree (DT), stochastic gradient descent, etc. have been investigated to predict heart disease. To achieve better results, ensemble learning methods such as random forest (RF), bagging, extra tree (ET) and XGBoost have been used. The introduction of some classification techniques is given below.

Logistic Regression (LR)

Logistic regression forecast a categorical variable from the set of predictor attributes. Along with a categorical dependent attribute, a discriminant task study is usually used where all of the predictors are

continuous and distributed. Logistic regression is also frequently used if predictor variables are a mix of continuous and categorical attributes also if they are not truly distributed (it makes no assumptions for the distribution of the predictor attribute). In logistic regression, the prediction of the dependent variable is a function of probability [3].

Decisions Tree (DT)

A decision tree is very popular among other classifiers [4]. It is a tree-type structure where the interior node represents a condition on an attribute, while every branch signifies the outcome of it, leaf node signifies the class label, and then the conclusion is taken after calculating all attributes. A path from a root to a leaf denotes classification rules [5]. In the health field, decision trees define the order of attributes. Primarily, it produces a set of resolved cases. After that, the set is split into training data and testing data. Training data is used for the training of a decision tree, and testing data is used to calculate the accuracy of an acquired solution [6].

Random Forest (RF)

Random forest algorithm is used for supervised and classification, but mostly it's used for classification problems. It generates decision trees based on data samples and then gets the prediction from each of them. After prediction, it selects the most suitable solution by means of voting. It is an aggregate method that is better than a single decision tree because it decreases the over-fitting by averaging the result [7].

K-Nearest Neighbor (KNN)

KNN algorithm is for classifying substances built on neighboring training samples in the feature space. It is a type of instance-based training, also called lazy learning, here the function is approximated till nearby only and all computation waits until classification. KNN is the basic and easiest classification method. There is less or no previous information for data distribution [8].

Extra Tree (ET)

Extra tree classifier [9] generates randomized multiple decision trees with different sub-samples without bootstrapping. It avoids the problem of over-fitting and results in better accuracy. Three important parameters for extra tree classifiers are: (i) M , represents the total number of trees to be generated, (ii) K , represents the number of attributes chosen for tree construction and (iii) n_{\min} denotes minimum required samples. When compared with random forest mechanism, extra tree classifier differs by choosing random K attribute and random split values for generating the tree. The attribute which shows minor bias-variance is identified as the best split attribute

Support Vector Machine (SVM)

SVM is an extremely popular supervised machine learning technique (having a pre-defined target variable) that can be used as classifier as well as a predictor. A SVM model represents the training data points as points in the feature space [10].

Multi-Layer perceptron (MLP)

MLP is a well-established neural network-based classification algorithm, which consists of three or more types of layers: an input layer, output layer and one or more hidden layers between input and output layers [11].

Every layer contains a number of 'neurons' connecting all the layers with each other. MLP is a universal multivariate non-linear mappings calculator that results from the capacity of training data to learn and generalize from training data using backpropagation learning methods [12].

The construction of MLP classifiers consists of adequate input variables and specification of the type of network, relevant data pre-processing and partitioning, the configuration of network infrastructure, specification of success parameters, specification of training algorithm (optimization of relation weights), and finally evaluation model [13].

Bagging

Bagging or bootstrap aggregating is applied where the accuracy and stability of a machine learning algorithm needs to be increased. It is applicable in classification and regression. Bagging also decreases variance and helps in handling overfitting.

Extreme Gradient Boosting (XGBoost)

XGBoost, which stands for extreme gradient boosting, is a scalable, distributed gradient-boosted decision tree (GBDT) machine learning library. It provides parallel tree boosting and is the leading machine learning library for regression, classification, and ranking problems. XGBoost is a robust machine-learning algorithm that can help understand data and make better decisions. XGBoost is designed for speed, ease of use, and performance on large datasets.

Stochastic gradient descent (SGD)

Stochastic gradient descent (often abbreviated SGD) is an iterative method for optimizing an objective function with suitable smoothness properties (e.g., differentiable or subdifferentiable). It can be regarded as a stochastic approximation of gradient descent optimization, since it replaces the actual gradient (calculated from the entire data set) by an estimate thereof (calculated from a randomly selected subset of the data). Especially in high-dimensional optimization problems this reduces the

very high computational burden, achieving faster iterations in exchange for a lower convergence rate. While the basic idea behind stochastic approximation can be traced back to the Robbins–Monro algorithm of the 1950s, stochastic gradient descent has become an important optimization method in machine learning [14].

Performance Measures

In the classification problems, the process of selecting the best metrics for evaluating the performance of a particular classifier for a given data set depends on the number of considerations including class balance and expected outcomes. One performance metric may evaluate a classifier from a single perspective while the others fail to measure it, and vice versa. Hence, there is no standardized (unified) metric for defining the generalized performance measurement of the classifier. In this paper, several metrics are chosen to measure how well models perform such as, such as accuracy, precision, recall, and F1 score.

These metrics are drawn from the following four categories:

True positives (TP): case where the true class of the instance was 1 (True) and the model prediction is also 1 (True).

False Positives (FP): a case where the true class of the instance was 0 (False) and the model prediction is 1 (True).

True Negatives (TN): a case where the true class of the instance was 0 (False) and the model prediction is also 0 (False).

False Negatives (FN): a case where the true class of the instance was 1 (True) and the model prediction is 0 (False) [15, 16].

Accuracy

Accuracy measure described as the average number of correct predictions. However, this is not quite robust for the unbalanced dataset.

$$Accuracy = (TP + TN) / (TP + FN + FP + TN)$$

Precision

Called positive predictive value measures the capability of a model to identify the correct instances for each class. This is a strong matrix for multi-class classification and unbalanced datasets.

$$precision = TP / (TP + FP)$$

Recall or Sensitivity

This metric measures a models performance in recognizing the true positive out of the total true positive cases.

$$\text{Sensitivity}(TPR) = TP / (TP + FN)$$

F1 score

Named as balanced F-score or F-measure, can be defined as a weighted average of precision and recall.

$$F1 = 2 \cdot \text{precision} \cdot \text{recall} / (\text{precision} + \text{recall})$$

ROC is a graphical way to show how good the performance of a classifier is. Basically, it is a plot of a true positive rate against a false positive rate.

That is, the number of correct predictions is divided by the number of actual positive results, and the rate of positive predictions is calculated. FPR, on the other hand, indicates the number of positive identifications among negative observations. This ratio is also used as a false positive rate in the ROC chart. Area under the curve (AUC) is used as a criterion for evaluating the performance of a classifier. Therefore, the closer the area under the graph to the number one, the better the classifier performance [17, 18].

Dataset

A heart disease dataset in Kaggle has been used for this paper. The Kaggle heart disease dataset contains total 70000 numbers of instances, each instance contains 12 attributes (independent variables) and the associated dependent Boolean variable target: diagnosis of heart disease which has the following two values: 0 or 1. Table 1 summarizes characteristics of attributes. The data set includes a series of characteristics such as age, gender, height, weight, Ap-hi, Ap-lo, cholesterol, Gluc, smoke, Alco, Alco and Cardio. In Table 1, the range of values is mentioned. For example, value 1 for women and value 2 for men are considered. Cardio feature as target, determines whether a person is sick or not. The value of the cardio feature is the same as the value 1 means being sick and the value zero means not being sick.

All the work is done in the Anaconda environment based on python programming language and Scikit-learn library.

Data Pre-Processing

Data pre-processing includes detecting outliers, making decisions what to do with the outliers, finding and filling appropriate missing values and searching for inconsistency in data. Data needs to be normalized or standardized and reduced before applying to machine learning algorithms. Normalization of data is done when the features of dataset have different measuring units. In order to build up a more accurate ML model, data preprocessing is required. Data pre-processing is the process of cleaning the data [19].

Table 1: Data set description

ID	Attribute name	Description	Domain
1	Age	Age	int (years)
2	Gender	Gender	categorical code
3	Height	Height	int (cm)
4	Weight	Weight	float (kg)
5	Ap-hi	Systolic blood pressure	int
6	Ap-lo	Diastolic blood pressure	int
7	Cholesterol	Cholesterol	1: normal, 2: above normal, 3: well above normal
8	Gluc	Glucose	1: normal, 2: above normal, 3: well above normal
9	Smoke	Smoking	binary
10	Alco	Alcoholic	binary
11	Alco	Physical activity	binary
12	Cardio	Heart disease (1 = no, 0 = yes)	binary

It will remove all the NAN values from our data. This process is also known as data wrangling. This includes the identification of missing data, noisy data and inconsistent data. The heart dataset is set up in comma separated document (CSV) format from Excel file. Different things required are the expulsion of right qualities for missing records, copy records evacuate pointless information field, standard information position, adjust information in a convenient way and so on. The considered heart data set do not have any missing data values for different attributes.

RESULTS

In this paper, using the classification algorithms in the test phase, the value of the target variable is predicted according to other features. After obtaining and predicting the variable target, this predicted value is compared with its actual value in the test set and degree of proximity of the prediction values to real values in the test set is calculated. Accuracy is calculated by several accuracy methods to hold out validation, 10-fold cross-validation, stratified 10-fold cross-validation, leave one out cross-validation, repeated random test-train splits, using Scikit-learn. In this study, k-fold cross validation was used to train and test the model. K-fold cross validation is used to prevent overfitting in forecasting models. In this paper, k is considered equal to 10. Feature importance score is also used in classification algorithms. Fig 1 shows the dependency values between all attributes in the dataset. Values below zero indicate a negative dependence and values

above zero indicate a positive dependence.

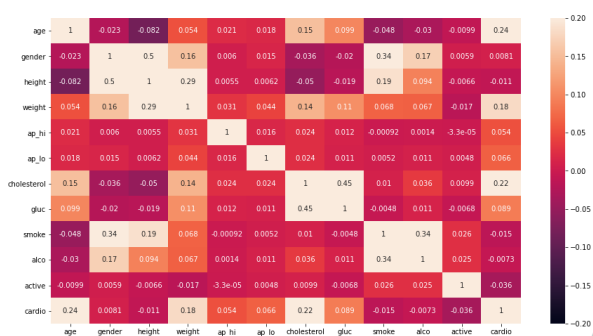


Fig 1: Correlation between all available features

The characteristics of age, cholesterol and weight have a high correlation with the characteristics of cardio. For example, the correlation coefficient of cardio and age is equal to 0.24, the correlation coefficient of cardio and weight is equal to 0.18, and the correlation coefficient of cardio and cholesterol is equal to 0.22.

The dataset consists of some features and some features are not important for the good-fit of learning models to improve the performance of machine learning models. This study deployed several feature selection techniques such as forward feature selection, backward feature elimination, bi-directional elimination, and machine learning feature selection. These techniques help to extract the important features from the dataset to train the machine learning models. In machine learning, feature selection is crucial to designing a good model and obtaining the best model performances. The redundant and undesired features may need to be removed from the original datasets to train the model faster, easily interpret the data, and avoid overfitting problems. This paper has considered several methods for feature selection, as determining the right set of features for heart disease classification is essential. The feature selection is based on the specific ML algorithm used to fit the dataset in the used method. A greedy selection method selects the combination of feature sets and evaluates the performance of the feature set combinations against the evaluation criteria. The evaluation criteria may include metrics such as accuracy, F1-score, etc., to assess the performance of feature set combinations.

Table 2 and Fig 2 to 5 show the performance outcome parameters of the classification algorithms employed, namely accuracy, precision, recall and F1 score. The random forest outperformed others by producing 92% F1score. The decision tree made it with 84%. The random forest algorithm outperformed others by producing 92% accuracy. Decision tree delivered 84%. The value of the precision criterion for the random forest algorithm is equal to 91%, which has the highest value compared to the rest. Also, the value of this parameter is equal

to 81% for decision tree algorithm. As a result of evaluation, random forest algorithm with F1 score value equal to 92% and accuracy equal to 92% have better performance than other algorithms.

Table 2: Models performance evaluation metrics

Classifier	Accuracy	Precision	Recall	F1-Score
Random Forest	0.9228	0.9102	0.9398	0.9247
Decision Tree	0.8428	0.8108	0.8823	0.8450
Extra Tree	0.7293	0.7385	0.7070	0.7224
XGBoost	0.7305	0.7459	0.6952	0.7197
MLP	0.6657	0.6193	0.8541	0.7180
LR	0.7037	0.7216	0.6607	0.6898
Bagging	0.6908	0.7056	0.6439	0.6733
SVM	0.5864	0.5570	0.8321	0.6730
KNN	0.6659	0.7192	0.5334	0.6125
SGD	0.5185	0.8241	0.0552	0.1034

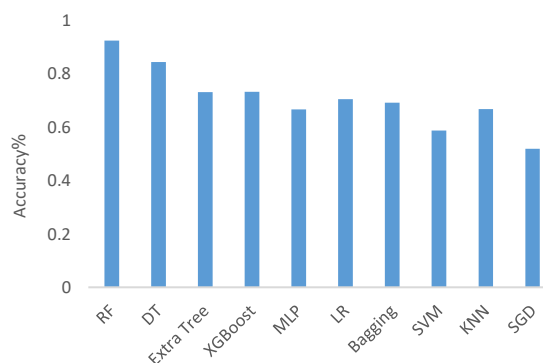


Fig 2: Accuracy of learning techniques

Fig 6 to 11 are for a better understanding of feature ranking and importance according to classification algorithms. These figures show the feature ranking based on feature importance and coefficient scores for all the applied classification algorithms except MLP and KNN. These figures also tend to represent the highly responsible attributes for heart disease.

Fig 6 to 11 show the important features according to feature importance and correlation value. According to the figures, it is found that Ap-hi is the significant feature or factor for identification and prediction. Besides age, cholesterol, weight and Ap-lo also significant factors predicting heart disease. Analysis of this result for doctors identified very important features for diagnosing and predicting heart disease.



Fig 6: Feature importance

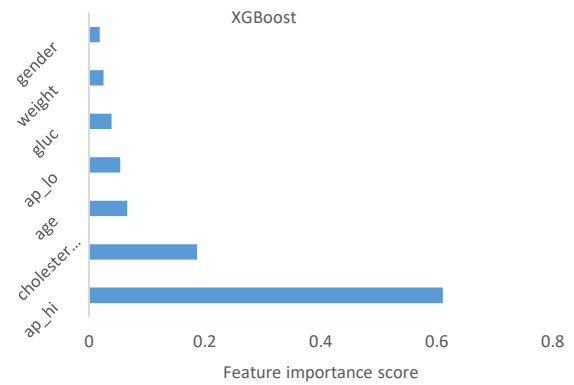


Fig 10: Feature importance

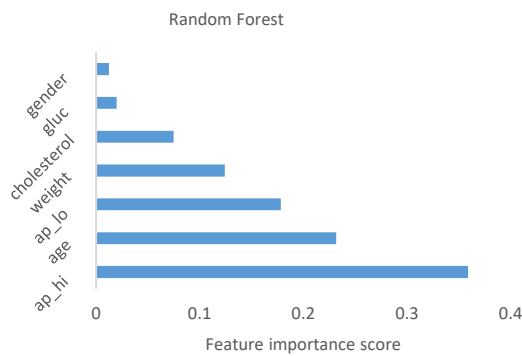


Fig 7: Feature importance

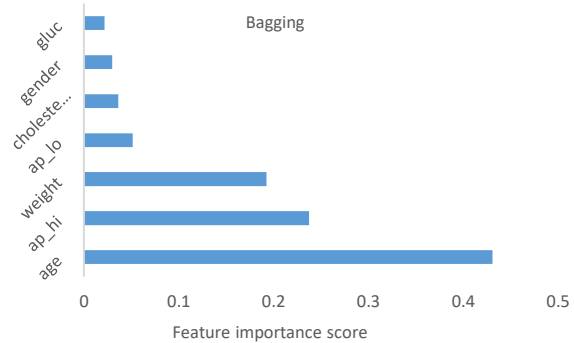


Fig 11: Feature importance

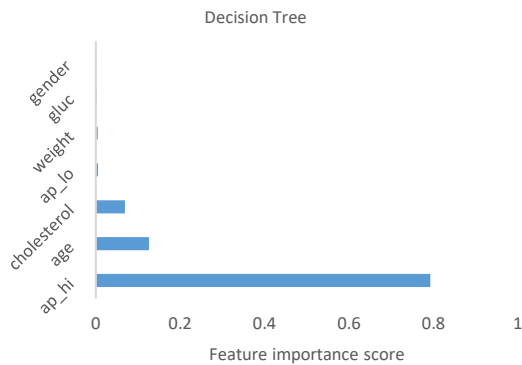


Fig 8: Feature importance

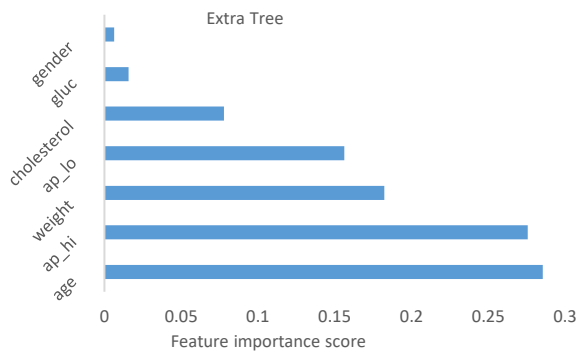


Fig 9: Feature importance

Table 3 shows the value of area under ROC for all the applied classification algorithms. Random Forest shows higher performance than them. On the other hand, decision tree provided the good result. Fig 12 represents the ROC, which is built by the value of the true positive rate and false positive rate. The area is bigger than others, the performance of the classifier is better. The random forest has the highest AUC score 0.95%. As a result of evaluation, random forest algorithm with F1 score value equal to 92% and AUC ROC equal to 95% have better performance than other algorithms.

Table 3: Area under ROC for classification algorithms

Algorithms	AUC ROC
Random Forest	0.951
Decision Tree	0.866
XGBOOST	0.800
Extra Tree	0.789
MLP	0.770
LR	0.763
Bagging	0.743
KNN	0.713
SVM	0.589
SGD	0.522

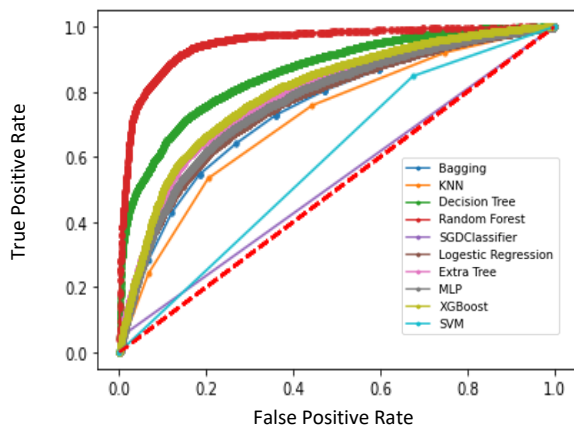


Fig 12: ROC curves of ten classifiers

DISCUSSION

Machine learning techniques in the field of medicine can be used to analyze a set of data related to a disease and to predict the heart disease. This study presents ML algorithms to identify high correlated features that are closely associated with heart disease identification. In this paper ML algorithms improved coupled by limited and selective features to produce high f1 score value in heart disease classification. In this paper it is analyzed different algorithms like KNN, DT, RF by comparing their f1 score levels on heart disease dataset collected from Kaggle. It also summarized the various prediction parameters obtained using each of the mentioned algorithms. In algorithms, methods such as the importance of features, hold out validation, 10-fold cross-validation, stratified 10-fold cross-validation, leave one out cross-validation are the result of effective performance and increase accuracy. In addition, feature importance scores for each feature was estimated for all the applied algorithms except MLP and KNN. These features were ranked based on feature importance score.

In order to Improvement of the performance of machine learning algorithms in predicting heart diseases, this paper highlight the importance of applying k-fold cross validation and feature importance score to compare the trained models generated from dataset with characteristics. Selecting good features and feature selection techniques has been effective in improving the results of algorithms, as shown in Fig 2 to 5. These help to extract the important features to train the machine learning models from the dataset.

The value under the ROC curve and evaluation criteria such as accuracy, sensitivity, accuracy and F1 score are compared to a number of machine learning classification algorithms to assess heart disease risk and actually predict heart disease to identify the best appropriate classifier. As a result of evaluation, Random Forest algorithm with F1 score value equal to 92% and AUC ROC equal to 95% have better

performance than other algorithms. However, the random forest model with limited features could be the best solution to improve the diagnostic accuracy of heart disease.

Some machine learning studies reported higher F1 score (100%) and accuracy (100%) for heart diseases prediction compared to the present study, which is likely due to using different databases. Similar to the database used in the current study, some studies used databases from specific medical or research centers.

Sultana et al., endorse coronary heart ailment prediction the usage of KStar, J48, SMO and Bayes internet and Multilayer perceptron using WEKA software [20]. Relying on the performance of various thing SMO (89% of accuracy) and Bayes internet (87% of accuracy) generate most beneficial performance than KStar, multilayer perceptron and J48 techniques the use of k- fold go validation. These algorithms have not been able to generate satisfactory results of performance. If one can improve the performance of accuracy, then it can help in better decision making to diagnosis heart disease.

Ali et al. research has been conducted using Cleveland dataset for heart diseases which involves 303 instances and used 10-fold cross validation, mentioning 13 attributes, putting 4 different algorithms, they resulted that Gaussian Naïve Bayes and random forest has given the maximum accuracy of 91.2% [21]. The usage of the same dataset of Framingham, Massachusetts, the experiments had been executed the usage of 4 fashions and had been skilled and tested with most accuracy ok- neighbors classifier: 87%, support vector classifier: 83%, selection tree classifier: 79% and random wooded area classifier: 84% [22]. For the heart disease prediction survey, Yadav et al. used a machine learning methodology with decision tree algorithm, Naïve Bayes, neural network, deep learning and SVM. CART, ID3, CYT, C5.0 and J48 are used to generate the outcome of the decision tree [23].

Shah et al. used the dataset which includes 303 instances and 76 attributes, of these 14 attributes were used for supervised learning algorithm which were decision tree, Naïve Bayes, KNN and random forest [24]. The results portray that the highest accuracy has achieved by KNN. Archana Singh et al. used dataset of 14 attributes for the models to trained and tested to get the maximum accuracy linear regression: 78%, decision tree: 79%, SVM: 83% and KNN: 87%. The results showed that the maximum accuracy has with KNN [25]. Pushkala et al. possess dataset of 75 attributes and 303 instances, out of which only 14 attributes are used for decision tree, random forest, SVM, KNN classifier, Gaussian Naïve Bayes. The Naïve Bayes gives the highest accuracy of 91% [26]. Mustafa et al. in their observation, ensemble mastering combines five classifiers version

approaches which includes assist vector system, synthetic neural network, naïve Bayesian, regression analysis and random forests to expect and diagnose the recurrence of cardiovascular disease. The dataset was taken from VCI data repository, 14 attributes were taken to train and test the models. The highest accuracy is 98.17% for the random forest algorithm [27].

Li et al. in their paper the classification algorithms were based on SVM, logistic regression, artificial neural networks, KNN, Naïve Bayes and decision tree. The model achieved accuracy of 92.37%, as compared to previous models [28]. Mohan et al. the primary aim of their paper was to find sufficient features by applying machine learning techniques like decision trees, language model, SVM, random forests, Naïve Bayes, neural networks, KNN. The proposed hybrid HRFLM method became used for combining the characteristics of random forests and linear technique. This model achieved the accuracy of 88.4% [29].

Prediction and evaluation of the prevalence of heart sickness and the use of data Mining strategies turned into counseled by Beyene et al. the key intention is expecting the occurrence of heart sickness in an effort to make an early computerized analysis of the sickness with a short end result. In healthcare establishments with specialists who lack enjoy and ability, the cautioned method is also crucial. It uses a ramification of medical traits, which includes blood sugar and heart price, as well as age and sex, to determine whether or not someone has heart disease. WEKA software is used to compute dataset analyses [30]. Kavitha et al. carried out a survey for the prediction of heart disease in their study it was found that most of the data was taken from Cleveland repository various machine learning classifiers are used to build heart disease prediction model. Based on the survey, the RF algorithms showed highest accuracy as compared to other models [31].

Kumar et al. used various machine learning techniques for the prediction of heart disease. The proposed model showed that random forests having highest accuracy achieved of 85.71% as compared to other classifier techniques [32]. Gavhane et al. in their paper, built a model for the prediction of heart disease using MLP which provide prediction results that gives the state of a user leading CAD to its users [33].

Kathiresan used supervised algorithms of machine learning to forecast heart disease. The learning algorithm is implemented on a variety of parameters in a real-time four-stage cloud system. The productivity is estimated to be over ten times higher, and cloud technology for forecasting and diagnosing heart disease has been integrated [34]. Sharma et al. summarized current state-of-the-art strategies and

available tools for disease prediction. Deep learning, a new branch of artificial intelligence, has shown promise in other fields, such as high-accuracy medical diagnosis [35].

CONCLUSION

Machine learning based solutions are widely used in healthcare sector for analyzing patients' data, predicting diseases and suggesting possible treatments. With a number of machine learning techniques available today, it is important to identify the most efficient and accurate technique especially in critical domains like healthcare. A comparative analysis of the various machine learning algorithms used in the heart disease prediction is presented.

Several techniques are discussed and compared to identify the best suited classifier for heart disease prediction. Many previous researches and studies related to heart disease prediction were identified and analyzed. Based on the identified researches and studies a performance analysis of various machine learning algorithms along with their F1 scores, accuracy and the value under the ROC curve used for prediction of heart diseases was done. The results showed that in most cases machine learning based approaches have shown significant potential to transform the healthcare sector and improve the entire process of disease predictions and suggesting treatments.

So, it used a heart disease dataset to test the utility of ML approaches to heart disease prediction and found that RF classification algorithm performs very well with an f1 score of 92%, accuracy of 92%, and ROC of 95% AUC. In addition, feature importance scores for each feature are estimated for all applied algorithms except MLP and KNN. These features were ranked based on the feature importance score. This study aimed to find the best ML techniques, among a number of well-accepted and easy-to-implement algorithms, and found that they performed well, at least for this dataset. This is an early stage of using ML approaches, but it shows that it can be a great aid to patient care.

AUTHOR'S CONTRIBUTION

All authors contributed to the literature review, design, data collection and analysis, drafting the manuscript, read and approved the final manuscript.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest regarding the publication of this study.

FINANCIAL DISCLOSURE

No financial interests related to the material of this manuscript have been declared.

REFERENCES

- Dwivedi AK. Performance evaluation of different machine learning techniques for prediction of heart disease. *Neural Computing and Applications*. 2018; 29(10): 685-93.
- Hasan R. Comparative analysis of machine learning algorithms for heart disease prediction. *ITM Web of Conferences*. EDP Sciences; 2021.
- Aradhana S, Jankisharan P, Virendra SK, Ashish M. Cardiovascular diseases prediction using various machine learning techniques. *IOP Conference Series: Materials Science and Engineering*. IOP Publishing; 2021.
- Soni J, Ansari U, Sharma D, Soni S. Predictive data mining for medical diagnosis: An overview of heart disease prediction. *International Journal of Computer Applications*. 2011; 17(8): 43-8.
- Li L, Wu Y, Ye M. Experimental comparisons of multi-class classifiers. *Informatica*. 2015; 39(1): 71-85.
- Vaghela C, Bhatt N, Mistry D. A survey on various classification techniques for clinical decision support system. *International Journal of Computer Applications*. 2015; 116(23): 11-7.
- Ali J, Khan R, Ahmad N, Maqsood I. Random forests and decision trees. *International Journal of Computer Science Issues*. 2012; 9(5): 272.
- Imandoust SB, Bolandraftar M. Application of k-nearest neighbor (KNN) approach for predicting economic events: Theoretical background. *International Journal of Engineering Research and Applications*. 2013; 3(5): 605-10.
- Baranidharan B, Pal A, Muruganandam P. Cardiovascular disease prediction based on ensemble technique enhanced using extra tree classifier for feature selection. *International Journal of Recent Technology and Engineering*. 2019; 8(3): 3236-42.
- Sayadi M, Varadarajan V, Sadoughi F, Chopannejad S, Langarizadeh M. A machine learning model for detection of coronary artery disease using noninvasive clinical parameters. *Life*. 2022; 12(11): 1933.
- Kwon K, Kim D, Park H. A parallel MR imaging method using multilayer perceptron. *Med Phys*. 2017; 44(12): 6209-24. PMID: 28944971 DOI: 10.1002/mp.12600 [PubMed]
- Tajmiri S, Azimi E, Hosseini MR, Azimi Y. Evolving multilayer perceptron, and factorial design for modelling and optimization of dye decomposition by bio-synthesized nano CdS-diatomite composite. *Environ Res*. 2020; 182: 108997. PMID: 31835116 DOI: 10.1016/j.envres.2019.108997 [PubMed]
- Azimi Y. Prediction of seismic wave intensity generated by bench blasting using intelligence committee machines. *International Journal of Engineering*. 2019; 32(4): 617-27.
- Stochastic gradient descent [Internet]. 2004 [cited: 22 Nov 2022]. Available from: https://en.wikipedia.org/wiki/Stochastic_gradient_descent
- Alalawi HH, Alsuwat MS. Detection of cardiovascular disease using machine learning classification models. *International Journal of Engineering Research & Technology*. 2021; 10(7): 151-7.
- Chowdhury MN, Ahmed E, Siddik MA, Zaman AU. Heart disease prognosis using machine learning classification techniques. *International Conference for Convergence in Technology*. IEEE; 2021.
- Ali MM, Paul BK, Ahmed K, Bui FM, Quinn JM, Moni MA. Heart disease prediction using supervised machine learning algorithms: Performance analysis and comparison. *Comput Biol Med*. 2021; 136: 104672. PMID: 34315030 DOI: 10.1016/j.compbimed.2021.104672 [PubMed]
- Yousefi S. Comparison of the performance of machine learning algorithms in predicting heart disease. *Frontiers in Health Informatics*. 2021; 10(1): 99.
- Dinesh KG, Arumugaraj K, Santhosh KD, Mareeswari V. Prediction of cardiovascular disease using machine learning algorithms. *International Conference on Current Trends towards Converging Technologies*. IEEE; 2018.
- Sultana M, Haider A, Uddin M. Analysis of data mining techniques for heart disease prediction. *International Conference on Electrical Engineering and Information Communication Technology*. IEEE; 2016.
- Ali M, Khan MD, Imran MA, Siddiki M. Heart disease prediction using machine learning algorithms [PhD Thesis]. BRAC University; 2019.
- Bhanot K. Predicting presence of heart diseases using machine learning [Internet]. 2019 [cited: 2 Nov 2020]. Available from: <https://towardsdatascience.com/predicting-presence-of-heart-diseases-using-machine-learning-6f00f3edb2c>
- Yadav R, Singh U. Survey on heart disease prediction by using machine learning technique. *International Journal of Creative Research Thoughts*. 2019; 7(1): 44-51.
- Shah D, Patel S, Bharti SK. Heart disease prediction using machine learning techniques. *SN Computer Science*. 2020; 1(6): 1-6.
- Singh A, Kumar R. Heart disease prediction using machine learning algorithms. *International Conference on Electrical and Electronics Engineering*. IEEE; 2020.
- Pushkala V, Agalya T, Angayarkanni SA. Comparative study of heart disease prediction using machine learning algorithms. *International Journal of Innovations in Engineering and Technology*. 2019; 12(4): 64-8.
- Jan M, Awan AA, Khalid MS, Nisar S. Ensemble approach for developing a smart heart disease prediction system using classification algorithms. *Research Reports in Clinical Cardiology*. 2018; 9: 33-

-
- 45.
28. Li JP, Haq AU, Din SU, Khan J, Khan A, Saboor A. Heart disease identification method using machine learning classification in e-healthcare. *IEEE Access*. 2020; 8: 107562-82.
29. Mohan S, Thirumalai C, Srivastava G. Effective heart disease prediction using hybrid machine learning techniques. *IEEE Access*. 2019; 7: 81542-54.
30. Beyene C, Kamat P. Survey on prediction and analysis the occurrence of heart disease using data mining techniques. *International Journal of Pure and Applied Mathematics*. 2018; 118(8): 165-74.
31. Kavitha BS, Siddappa M. A survey on machine learning techniques to predict heart disease. *International Journal of Computer Science & Communication*. 2020; 11(Special Issue): 48-53.
32. Kumar NK, Sindhu GS, Prashanthi DK, Sulthana AS. Analysis and prediction of cardio vascular disease using machine learning classifiers. *International Conference on Advanced Computing and Communication Systems*. IEEE; 2020.
33. Gavhane A, Kokkula G, Pandya I, Devadkar K. Prediction of heart disease using machine learning. *International Conference on Electronics, Communication and Aerospace Technology*. IEEE; 2018.
34. Kathiresan S. Analysis on cardiovascular disease classification using machine learning framework. *ICTACT Journal on Data Science and Machine Learning*. 2020; 2(1): 153-6.
35. Sharma H, Rizvi MA. Prediction of heart disease using machine learning algorithms: A survey. *International Journal on Recent and Innovation Trends in Computing and Communication*. 2017; 5(8): 99-104.