Scientific Research Publishing

Journal of Computer and Communications, 2022, 10, 1-18
https://www.scirp.org/journal/jcc
ISSN Online: 2327-5227
ISSN Print: 2327-5219

Performance Comparison of Various Machine Learning Approaches to Identify the Best One in Predicting Heart Disease

Enas M. Abd Allah[1], Doaa E. El-Matary[1], Esraa M. Eid[2], Adly S. Tag El Dien[2]

[1]Electronics and Communications Department, Al-Safwa High Institute of Engineering, Qalyubia, Egypt
[2]Electrical Engineering Department, Faculty of Engineering at Shoubra, Benha University, Cairo, Egypt
Email: inas.mostafa@alsafwa.edu.eg, doaa.elmatary@alsafwa.edu.eg, esraa.soliman@feng.bu.edu.eg, adlytag@feng.bu.edu.eg

How to cite this paper: Allah, E.M.A., El-Matary, D.E., Eid, E.M. and El Dien, A.S.T. (2022) Performance Comparison of Various Machine Learning Approaches to Identify the Best One in Predicting Heart Disease. *Journal of Computer and Communications*, **10**, 1-18.
https://doi.org/10.4236/jcc.2022.102001

Received: December 23, 2021
Accepted: February 8, 2022
Published: February 11, 2022

Abstract

Nowadays, machine learning is growing fast to be more popular in the world, especially in the healthcare field. Heart diseases are one of the most fatal diseases, and an early prediction of such disease is a vital task for many medical professionals to save their patient's life. The main contribution of this research is to provide a comparative analysis of different machine learning models to reach the most supporting decision for diagnosing heart disease with better accuracy as compared to existing models. Five models namely, K-Nearest Neighbor (KNN), Logistic Regression (LR), Random Forest (RF), Support Vector Machine (SVM), and Extreme Gradient Boost (XGB), have been introduced for this purpose. Their performance has been tested and compared considering different metrics for precise evaluation. The comparative study has proven that the XGB is the most suitable model due to its superior prediction capability to other models with an accuracy of 91.6% and 100% on two different heart ailments datasets, respectively. Both datasets were acquired from the heart diseases repositories where dataset_1 was taken from the University of California, Irvine (UCI) and dataset_2 was from Kaggle.

Keywords

Machine Learning, Healthcare, Heart Disease, Prediction

1. Introduction

Recently, Machine Learning plays a vital role in the sector of healthcare [1]. It is a method that allows machines to act like a human by repeating their behavior. It makes machines learn from their experience (*i.e.* training data) without being

programmed and then they can predict wanted elements.

Remote Healthcare technologies can also be used to insert decision support systems with mobile devices. It can collect data in real-time from patients and provide health services efficiently. It helps monitoring patients without visiting hospitals or health centers [2].

The heart is the critical part of the human body, which provides pure blood to all parts of the body. Without heart working healthy, people cannot live for a second. Typically, heart failure occurs when the heart cannot push the needed amount of blood to other parts of the human body to make the body work normally. Currently, heart ailments are increasing rapidly as there were 80% of people died due to a heart attack every year, according to a survey by the World Heart Organization (WHO). Heart ailment became one of the world's life-threatening human ailments [3].

Predicting and classification of heart ailment early play a critical role in treatment. When heart ailment could be expected earlier, more deaths of patients would be avoided and successful diagnosis would be known. Every day, there is a need to improve a system of medical diagnosis. The crucial points of medical diagnostics programs are to reduce cost with effective achievement for more reliable results. The development of the system for medical diagnosis based on machine learning to predict heart ailments produces a highly specific decision unlike the traditional way and reduces the cost of treatment [4].

Effective classification along with medicinal treatment reduces people's deaths. In this research, a comparative analysis of the UCI Cleveland dataset and another Kaggle heart disease dataset using a five supervised machine learning classification algorithms listed as KNN, Logistic Regression, Random Forest, SVM, and XG-Boost and selecting the best classifier to classify heart ailment with more accuracy.

The rest of this research is organized as follows: Section 2 presents the related works. Section 3 describes the methodology and the dataset information used in detail. The results of the proposed models are discussed in Section 4 and finally, Section 5 concludes the paper.

2. Related Work

In [1]: The authors explored and investigated different machine learning algorithms for the heart disease dataset. They trained and tested six models, which are Logistic Regression, Random Forest, XG-Boost, Support Vector Machine, Artificial Neural Network, and K-Nearest Neighbors. Random Forest was the most accurate algorithm used in this paper. The used dataset is taken from the UCI repository with 14 features and 303 instances. The performance, accuracy were shown as Random Forest 100% but over fitting occurs, XG-Boost 83%, Logistic Regression 83%, Artificial Neural Network 83%, Support Vector Machine 79.56%, and K Neighbors 71.69%.

In [2]: The authors proposed the development of the proposed hybrid system.

They used Support Vector Machine, Naïve Bayes, Logistic Regression, Random Forest, and Ada-Boost classifiers. The dataset of Cleveland heart disease was used in this research with 303 instances and 14 attributes. Random forest algorithm achieved the most accurate result with an accuracy 86.6%. It has been found that the system has given the most accurate results with a random forest classifier. The best performance, accuracy improvements using feature selection were Naïve Bayes (NB) 83.55%, SVM 84.46%, LR 85.07%, RF 86.60%, and Ada-Boost 86.59%.

In [3]: This paper applied data mining algorithms to predict heart disease. The authors implemented two algorithms, Naïve Bayes and (NB) tree, on different two datasets from the UCI repository to evaluate the performance. The first dataset was obtained from Cleveland Clinic Foundation with 14 attributes and 303 instances. The second dataset is taken from the public available platform named Heart Disease Dataset (Comprehensive) with 11 attributes and 1190 instances. The results were NB tree with 84.6% accuracy compared to Naive Bayes with only 80.58 % accuracy.

In [4]: The author studied the classification of heart disease by an automated medical diagnosis system with machine learning. Different machine learning classification techniques were used as Logistic Regression, Naïve Bayes (NB), Support Vector Machine (SVM), Decision Tree (DT), Random Forest, and K-Nearest Neighbor (KNN). The Cleveland dataset for heart disease classification was used in this research with 14 attributes and 303 instances. The result of performance, accuracy was Logistic regression with 76.31%, Multinomial Naive Bayes with 72.37%, Gaussian Naive Bayes with 84.21%, Bernoulli Naive Bayes with 77.63%, Linear Support Vector Classifier (SVC) with 89.47%, Decision tree classifier with 65.79%, Random forest classifier with 84.21% and K Neighbors Classifier with 84.21%.

In [5]: The authors studied machine learning algorithms for making a prediction of heart sickness using the Cleveland dataset which has 13 attributes, 1025 instances. They also analyzed the importance of the features of the dataset. Decision tree and Ada-Boost algorithms have been used to make the prediction with accuracy 97%, but a training set completely over-fitting the data for Decision tree and 89.88% for Ada-Boost.

In [6]: The author analyzed the Heart Disease dataset for 1025 patients collected from Cleveland, Hungary, Switzerland, and Long Beach with 14 attributes for heart disease classification. The algorithms used Naive Bayes, Stochastic Gradient Decent (SGD), SVM, KNN, Decision Table (DT), Ada-boost, and J-Ripper (J-Rip) classifiers to show the performance of them to best classify the heart disease cases. Accuracy of the selected algorithms (Naïve Bayes 83.122%, SGD 84.3902%, SVM 84.1951%, KNN 99.7073, Decision Table 93.6585%, Ada-boost 84.2927%, J-Rip 97.2683%).

In [7]: The authors studied several supervised machine learning algorithms that were applied and compared to achieve performance and accuracy of heart

sickness prediction. They used a dataset of heart disease obtained from Kaggle, this dataset contains 1025 instances and 14 attributes. They tested different classification algorithms (Logistic Regression, multilayer perceptron (MLP), KNN, Decision Tree, and Random Forest). The results of performance, accuracy were LR with 89.627%, MLP with 97.951%, KNN with 100.000%, DT with 100.000%, and RF with 100.000%.

According to the works discussed above, this research focuses on heartbeat rate. It produces a model that can predict heart diseases using machine learning algorithms with various algorithms listed as KNN, LR, RF, SVM, and XGB, then determine the performance of each algorithm to detect the best classifier model. It is demonstrated that all authors of the above researches have used only one dataset of heart diseases except one author has used two datasets but with lower performance accuracy achieved. In this study, two different datasets were used and compared to analyze the heartbeat rate with the superior accuracy of 91.6% and 100% for dataset_1 and dataset_2, respectively, which was supported with a better model for detecting the heart diseases more accurate than other models.

3. Methodology

The methods and materials used in this study will be explored through the following points:

3.1. Work-Flow of the Model

The workflow of the system has been implemented in different stages including Pre-processing of the dataset, Cross-Validation, Classification, and Performance Evaluation as depicted in Figure 1. Heart disease is diagnosed with the help of UCI and Kaggle datasets. Moreover, it is divided into a training and testing set.

3.2. Tools Used

- The Pandas tool is an open-source python package used to conduct this study, which is written in python or C. Tools for writing and reading data between in-memory data structures and various formats: Text files, Microsoft

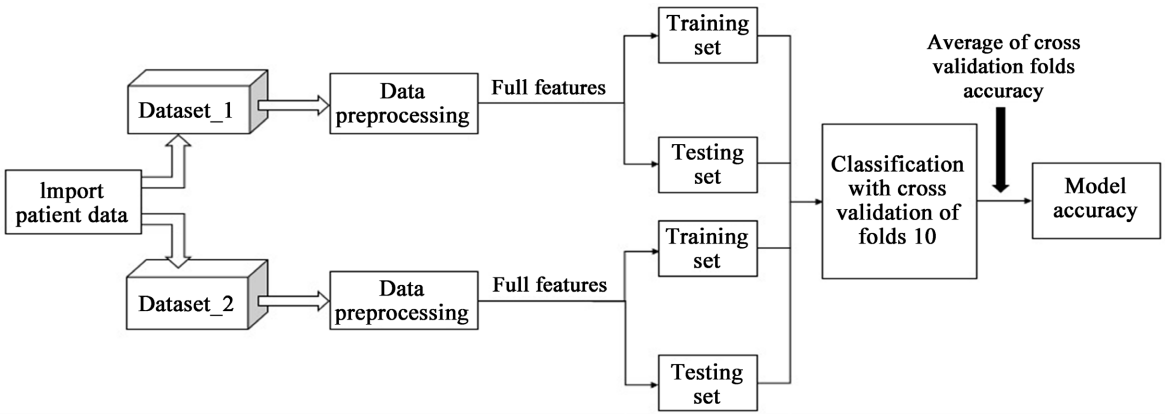

Figure 1. Work-flow of the model.

Excel, comma-separated values (CSV), structured query language (SQL) databases, and the fast Hierarchical Data Format 5 (HDF5) format [8].

- Matplotlib is a comprehensive library for creating animated, interactive, and static visualizations in Python used for machine learning [9]. In machine learning, it is useful to understand the vast amount of data through different visualization.

3.3. Dataset Description

Two heart disease datasets are used in this research. They are taken from the UCI and Kaggle repository respectively. The first dataset contains a total of 303 cases, 138 of which are healthy people and 165 have heart disease [10] while the other contains a total of 1025 cases, 499 of which are healthy people and 526 have heart disease [11] as depicted in **Figure 2**.

The two datasets have been selected with 76 attributes and preprocessed to produce 14 only for reducing the redundant variables. Four attributes are used to indicate common symptoms of the patient, and the remaining attributes are used to indicate ECG values. The attributes for both datasets are shown in detail in **Table 1**.

For a pictorial representation, histogram plotting has been created of age and sex attributes. Data of patients are grouped according to age and gender attributes with the absence and presence of heart disease for the two datasets as depicted in **Figure 3** and **Figure 4**.

Correlation is used to determine the relationship between two continuous, quantitative variables. The determination of relevant features is performed using the correlation technique. The correlation matrix is computed to detect the relationship between attributes of the dataset. This can improve the machine learnin cancelled weakly correlated attributes. The correlation matrix for the two datasets are plotted in **Figure 5** to well understand the correlation between the

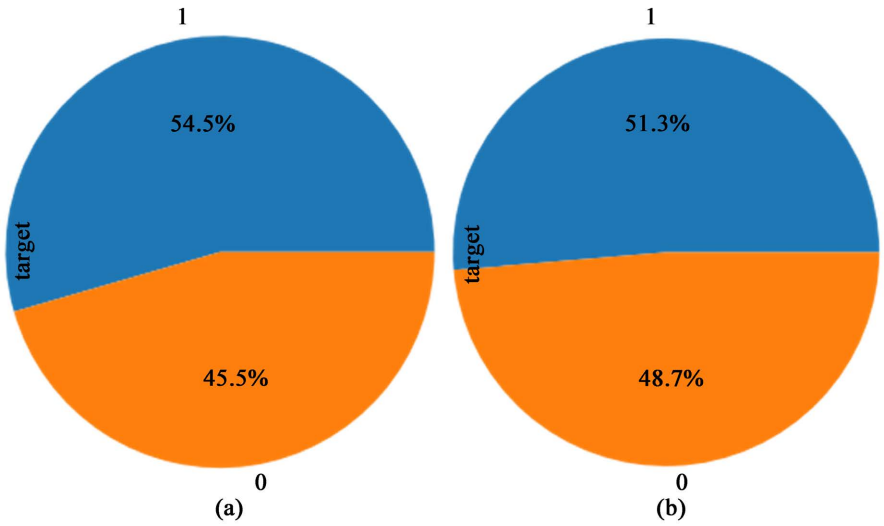

Figure 2. Percentage of people who has a heart disease for the two datasets which 0 → absence of heart disease, and 1 → presence of heart disease. (a) Dataset_1; (b) Dataset_2.

attributes. It is depicted in different colors, the dark color represents that the attribute are strongly correlated with another and light color performs a weakly correlated with another. Correlation range values from (−0.4 to +1.0). Positive correlation increases or decreases the column attributes together. Negative

Table 1. Detailed attributes of both datasets.

Attribute used	Attribute information
Sex	The patient's gender represented in binary form. Male = 1, Female = 0.
Age	The patient's age in years. Range → 29 years: 77 years.
Chest pain type (CP)	Chest pain. Range → 1:4 1 → Typical angina, 2 → Atypical angina 3 → Nonanginal Pain, 4 → No pain.
Resting blood pressure (Rest BP)	The patient's resting blood pressure, in (mm Hg), admitted in hospital. Range → 94:200
Serum cholesterol (Chol.)	Serum cholesterol, in (mg/dl). Range → 120:154
Fasting blood sugar (FBS)	The patient's fasting blood sugar, it is higher than 120 mg/dl → True = 1, False = 0.
Resting electrocardiographic results (Rest ECG)	The patient's resting electrocardiography records. Range → 0:2. 0 → Normal. 1 → ST-T wave abnormality. 2 → Probable or definite left ventricular hypertrophy.
Maximum heart rate achieved (HR)	The patient's maximum heart rate achieved. Range → 71:202
Exercise induced angina (Exang.)	Exercise induced angina, binary. 1 → Yes, 0 → No.
Old peak (OP)	ST depression induced by exercise, relative to the rest. Range → 0:6.2
Slope of peak exercise ST segment (Slope)	Measure the slope for peak exercise. Range → 1:3. 1 → Up sloping, 2 → Flat, 3 → Down sloping.
Number of major vessels colored by fluoroscopy (CA)	The number of major vessels colored by fluoroscopy. Range → 0:3, (value is related to the darkness of the color).
Thallium scan (Thal.)	Thallium heart Scan of the patient, (3, 6, 7). 3 → Normal, 6 → Fixed defect, 7 → Reversible defect.
Target (TRT)	Diagnosis of heart disease (angiographies disease status). 0 → Absence of heart disease. 1 → Presence of heart disease.

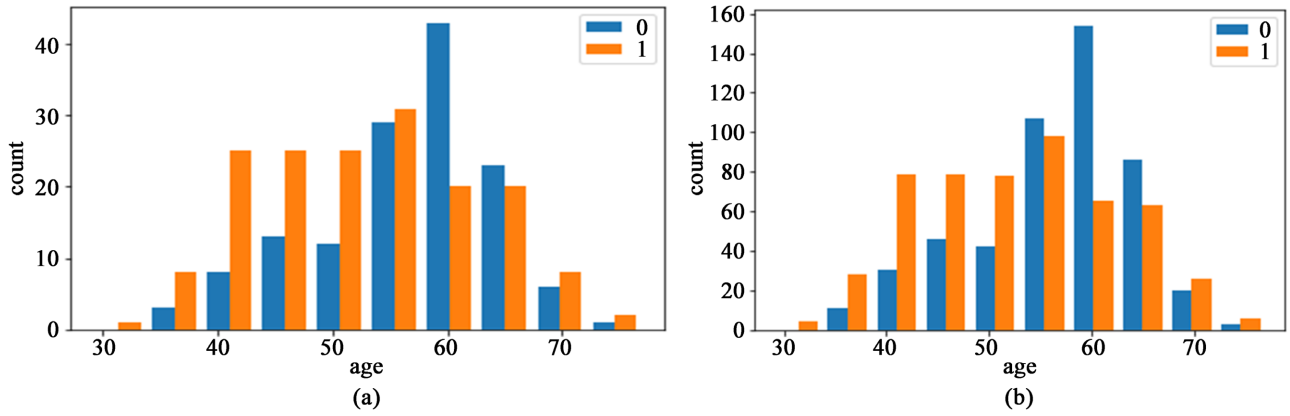

Figure 3. Group the age of the patients' data with the absence and presence of heart disease. (a) Dataset_1; (b) Dataset_2.

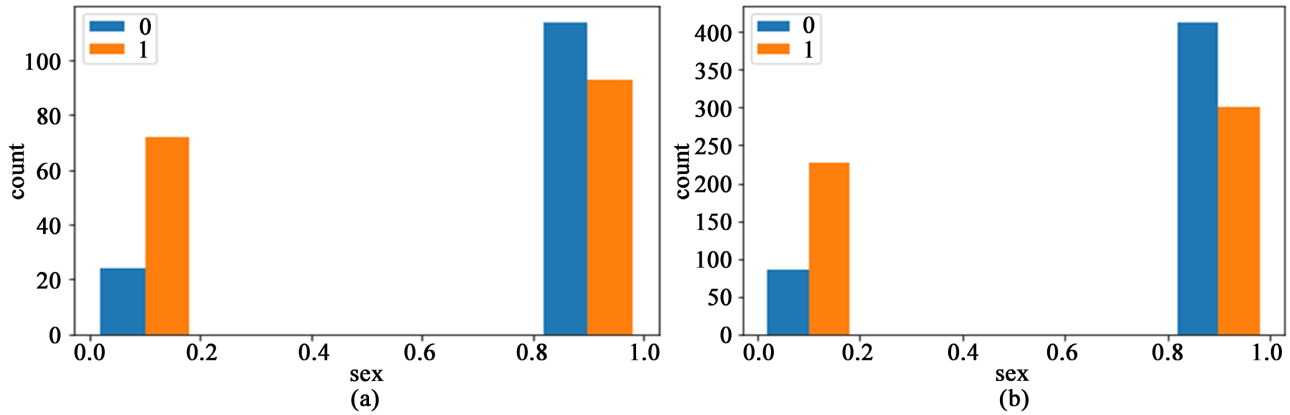

Figure 4. Group the gender of the patients' data with the absence and presence of heart disease. (a) Dataset_1; (b) Dataset_2.

correlation performs that one attribute will increase and another one decreases or *vice versa* [12].

3.4. Data Pre-Processing

Preprocessing data means the changes which are made on data before it is fed as an input to the algorithm. Data obtained from many sources is described as raw data, not suitable for analysis. In order to obtain better results, it is necessary to remove outliers, noise, and irregularities from the data, known as data cleaning as described below [13].

- Data Cleaning

The data that needs to be analyzed using algorithms of machine learning may be noisy, inconsistent and incomplete. It also deals with the missing values for attributes of interest as it changes the proper average value for the attribute. Likewise, invalid attribute values are cleared and filled manually with its mean value. Data is cleaned up by manipulating missing values, smoothing out noisy data and removing outliers [14].

The outlier is defined as a value that is more than 3 standard deviations from the mean. Then the outliers will be removed. **Figure 6** shows the outliers of some selected features for the two datasets.

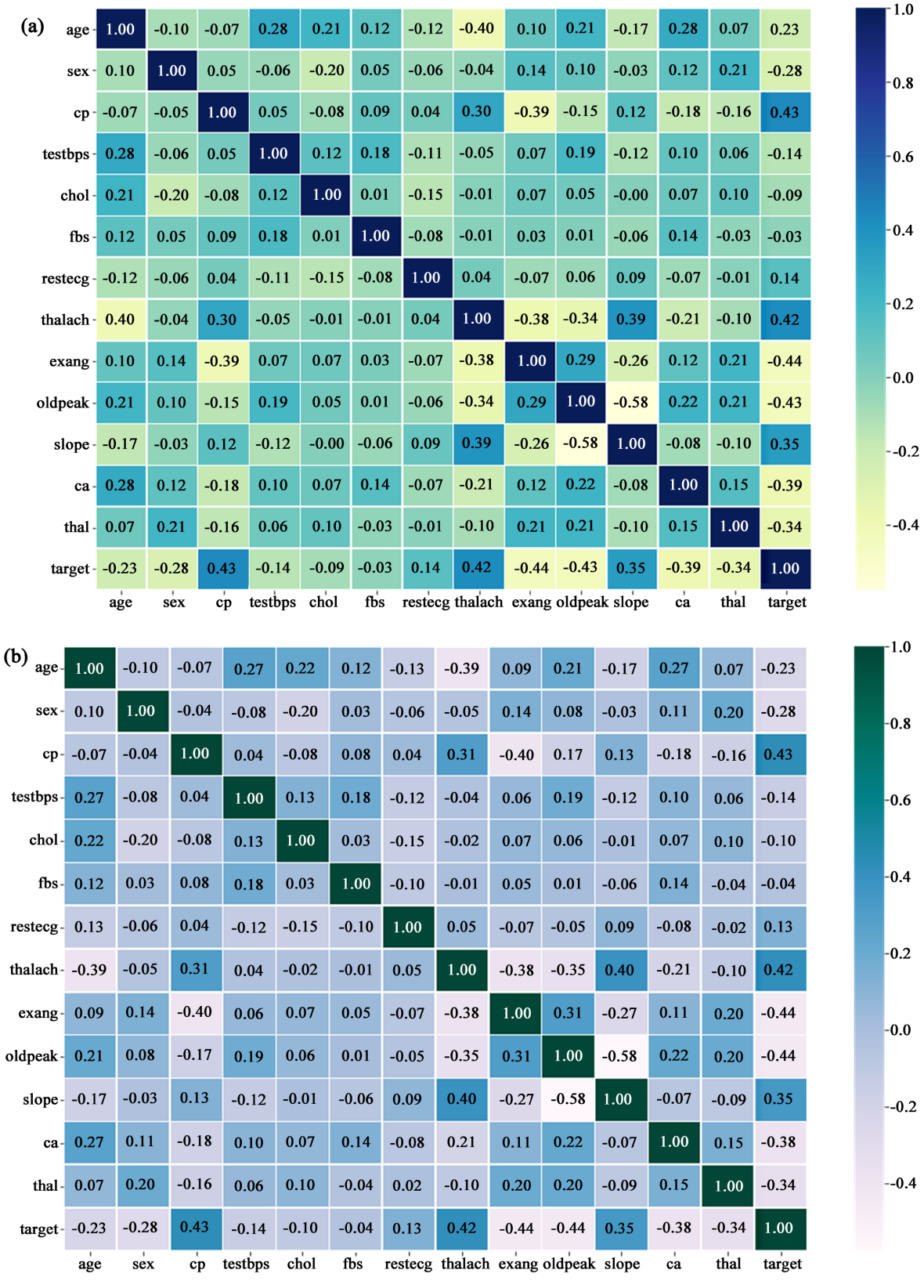

Figure 5. Correlation matrix of the various parameters in the two datasets of heart diseases. The color coding scale denotes the degree of Pearson correlation between variables with dark color being positively correlated and light color negatively correlated. (a) Dataset_1; (b) Dataset_2.

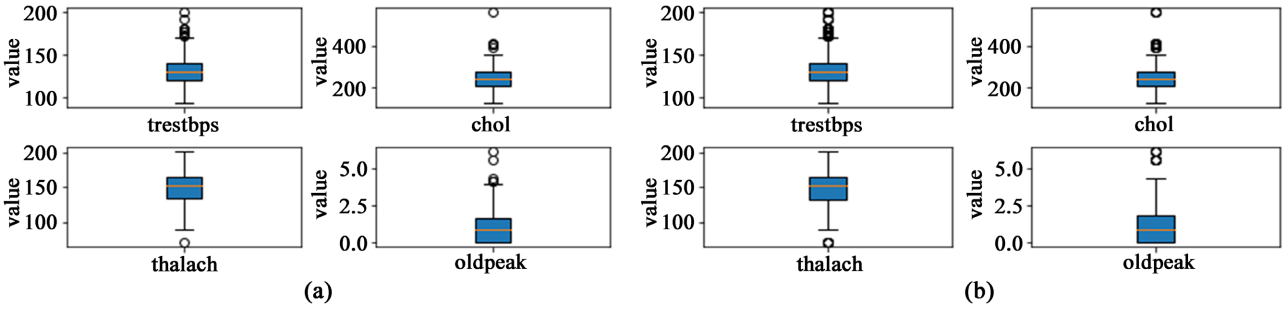

Figure 6. Box plot for outliers of selected features in the two dataset. (a) Dataset_1; (b) Dataset_2.

- Data Splitting

The dataset used in this research is splitting into 80% - 20%, which 80% of original data is considered as training dataset and 20% as testing dataset. Training dataset is used to train a model and testing dataset to check the performance of the trained model. For each algorithm the performance is analyzed and computed depending on different metrics used as F-measure scores, recall, precision and accuracy as described further. The various algorithms explored in this research are listed as below.

3.5. Machine Learning Algorithms

Many classification algorithms in machine learning are available, but it is complex to determine which one is superior to the others. It mainly depends on the nature of the dataset and the application used [14]. This research presents the detailed description of the five supervised classification algorithms (KNN, LR, RF, SVM and XGB) and how each algorithm work on the datasets. Firstly, each algorithm is trained with a percentage of the dataset, known as 'training set' and then tested on 'testing set' which is put away as 'invisible data' from evaluating the algorithm. The reason for choosing these five algorithms is that they are more suitable in its parameters for the two datasets used in this study and they can achieve high performance metrics measures as accuracy, precision, recall, and f-measure.

- K-Nearest Neighbor (KNN) Algorithm:

KNN, is the simplest and most popular supervised machine learning algorithm used for regression and classification. It is non-parametric learning algorithm [14]. It classifies a new sample based on the value of k, by the majority choice in the classification of the nearest k neighbors. It can be measured by the Euclidean distance "Equation (1)" as follow [4]:

$$D(x,y)=\sqrt{\sum_{i=1}^{n}\left(x_i-y_i\right)^2}=\sqrt{\left(x_1-y_1\right)^2+\left(x_2-y_2\right)^2+\cdots+\left(x_n-y_n\right)^2} \quad (1)$$

The best advantage of using this model is that it only needs a few tunes such as K and distance measurement for working to achieve high accuracy.

- Logistic Regression (LR) Algorithm:

LR is a classification supervised algorithm that is used for binary classification

problems. In these datasets the target attribute has the two types of binary numbers, (0) for healthy patients, and (1) for the patients who suffer from heart diseases [15]. It converts its output using the function of logistic sigmoid to return a probability value. The "Equation (2)" of logistic regression is as follows:

$$\varphi(x) = \frac{1}{1+e^{-x}} \tag{2}$$

- Random Forest (RF) Algorithm:

RF algorithm is a supervised machine learning technique, which can be used for both regression (input is a discrete data) and classification (input is a continuous data) tasks. Several trees, build a forest, in this algorithm. Each tree in the RF allows the class prediction and class with the most votes converts into the model's prediction. Higher accuracy is achieved when many numbers of trees are used [4]. The built model depended mostly on two important parameters of random forest, one of them is the maximum number of trees and the other one is max depth.

- Support Vector Machine (SVM) Algorithm:

SVM is a supervised machine learning algorithm that is used for both regression problems as support vector regression (SVR) and classification problems as support vector classification (SVC). It takes long time to process so it is suitable used for smaller dataset. It segregates the data two classes using a hyper plane. The hyper-plane is the decision limit which classifies the dataset into two classes. The accuracy of classification is improved by the maximum distance between data points of two classes [14]. The mathematical "Equation (3)" of hyper-plane describes as:

$$w \cdot x + b = 0 \tag{3}$$

where, x is the data point, w is the weighted vector and b is the scalar data.

- Extreme Gradient Boost (XGB) Algorithm:

Nowadays, XGB algorithm is the most common algorithm for machine learning. It is a supervised technique, that it has better solutions than other Machine learning algorithms regardless the form of the data (classification or regression). It is similar to the gradient boosting algorithms, but it is more effective [1]. It predicts accurately a target variable by merging an ensemble of estimations from a set of simpler models.

4. Results and Discussion

In this paper, four standard statistical measures, accuracy, precision, recall and F-scores are generated to estimate the performance of the classifier. The performance evaluation is based on computing the confusion matrix as depicted in Table 2. Confusion matrix is a table that often used to describe the performance of a classification model on a test data for which known true values. It is relatively simple to understand, but terminology related can be confusing.

True positive (TP): The number of healthy patients correctly predicted as healthy.

Table 2. Confusion matrix.

Matrix	Predicted cases	
	+	−
Actual cases	+TP	FP
	−FN	TN

False positive (FP): The number of unhealthy patients predicted to be healthy.

False negative (FN): The number of unhealthy patients that were correctly classified as unhealthy.

True negative (TN): The number of healthy patients that were incorrectly classified as unhealthy.

The efficiency of the classifiers in identifying cardiac disease could be measured from the confusion matrix evaluation and estimated below parameters [17].

Accuracy: is the classifier's ability to correctly predict that the class of the instances were be labeled for all the instances. It can be computed with "Equation (4)" [14]:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{4}$$

Precision: is the relation between the number of positive predictions and the total number of positive prediction class values. It can measure the exactness of the classifier as shows in "Equation (5)" [14]:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{5}$$

Recall: is the measure of positive prediction numbers divided by the number of positive class values in testing data. It is the completeness of the classifiers and it can be computed with "Equation (6)" [14]:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{6}$$

F-Measure: expresses the balance between the recall and precision. It is the harmonic mean of both precision and recall and it can be computed as shown in "Equation (7)" [14]:

$$\text{F-Measure} = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}} \tag{7}$$

In this research, KNN, LR, RF, SVM and XGB classifier algorithms are applied to the heart diseases datasets acquired from UCI and Kaggle repository respectively. Two datasets of heart disease are used. The first includes 303 instances and the second 1025 instances. Each instance is composed of 14 attributes like class attribute. The class attribute contains two numbers, such as, absence (0), presence (1). All the attributes of the dataset along with their range of UCI machine learning platform provides principles. For the input parameters listed in

Table 1, the classification algorithms are applied.

The classifiers with k (k = 10) fold cross validation are used for classification where the entire dataset is divided into k subsets where the k-1 subsets is used for training and the other one is used for model testing. Cross validation model repeats the process for 'k' times. Then results are analyzed and compared using Pandas and Matplotlib software. In Pandas and Matplotlib, Data pre-processing has been carried out as first step for all the 13 attributes then the optimal values for tuning parameters can be obtained. The Performance metrics of all 5-algorithms across both datasets with the best parameters considering their confusion matrix are shown in figures from Figures 7-11.

Figure 7 shows the performance metrics of KNN with Euclidean distance and varying K including the confusion matrix for both datasets. It can be noticed that the accuracy of KNN will be improved by increasing the value of K to be 7 for both datasets. Also the results indicate that the other metrics, recall, precision, and F-score, differs slightly (nearly the same) for all values of K but they reach to maximum records at K = 7. Hence K = 7 will be considered the best

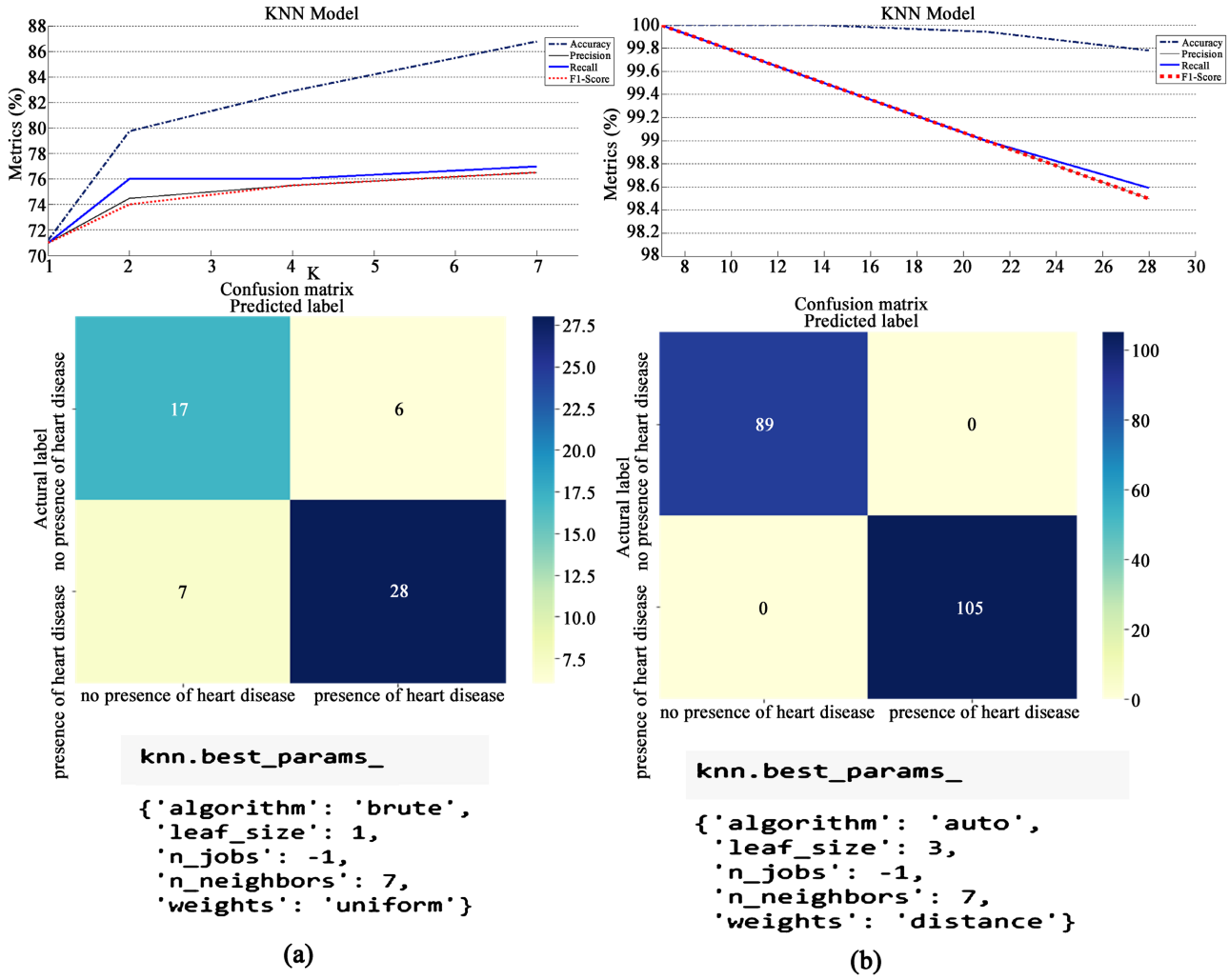

Figure 7. The performance metrics of KNN with best K for the two datasets. (a) Dataset_1; (b) Dataset_2.

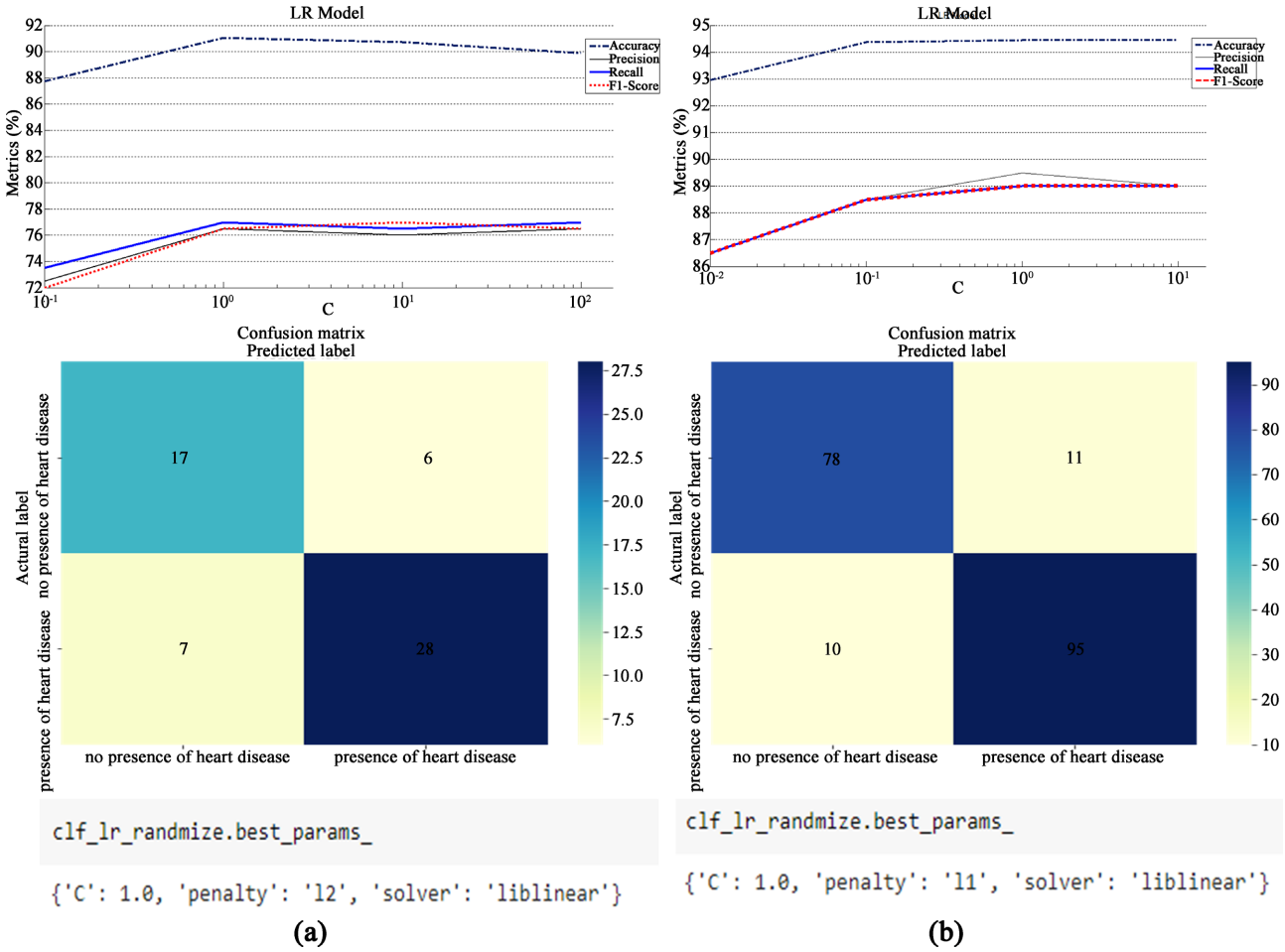

Figure 8. The performance metrics of LR with best C for the two datasets. (a) Dataset_1; (b) Dataset_2.

parameters in the rest of the study.

Figure 8 describes the performance metrics of LR with varying C parameter including the confusion matrix for both datasets. As it is shown above the accuracy of LR will be enhanced when the value of C is equal 1 for both datasets. The results of the other metrics, recall, precision, and F-score are also mostly similar to all values of C, but they are achieved maximum records at C = 1 that it will be the best parameters for this study.

Figure 9 depicts the performance metrics of RF with varying N and their confusion matrix for both datasets. It is clear that the accuracy of RF will be better when the value of N is equal to 566 for first dataset and 88 for the second dataset. Moreover the results for other metrics, recall, precision, and F-score, are slightly different from other values of N, but the maximum records are at N = 566 and N = 88 for the two datasets that they will be the best parameters for the two datasets, respectively.

Figure 10 presents the performance metrics of SVM with varying C with their confusion matrix for the two datasets. It can be shown that the accuracy of SVM will be increased as the value of C is equal 100 and 10 for both datasets, respectively. The results of the other metrics, recall, precision, and F-score, have

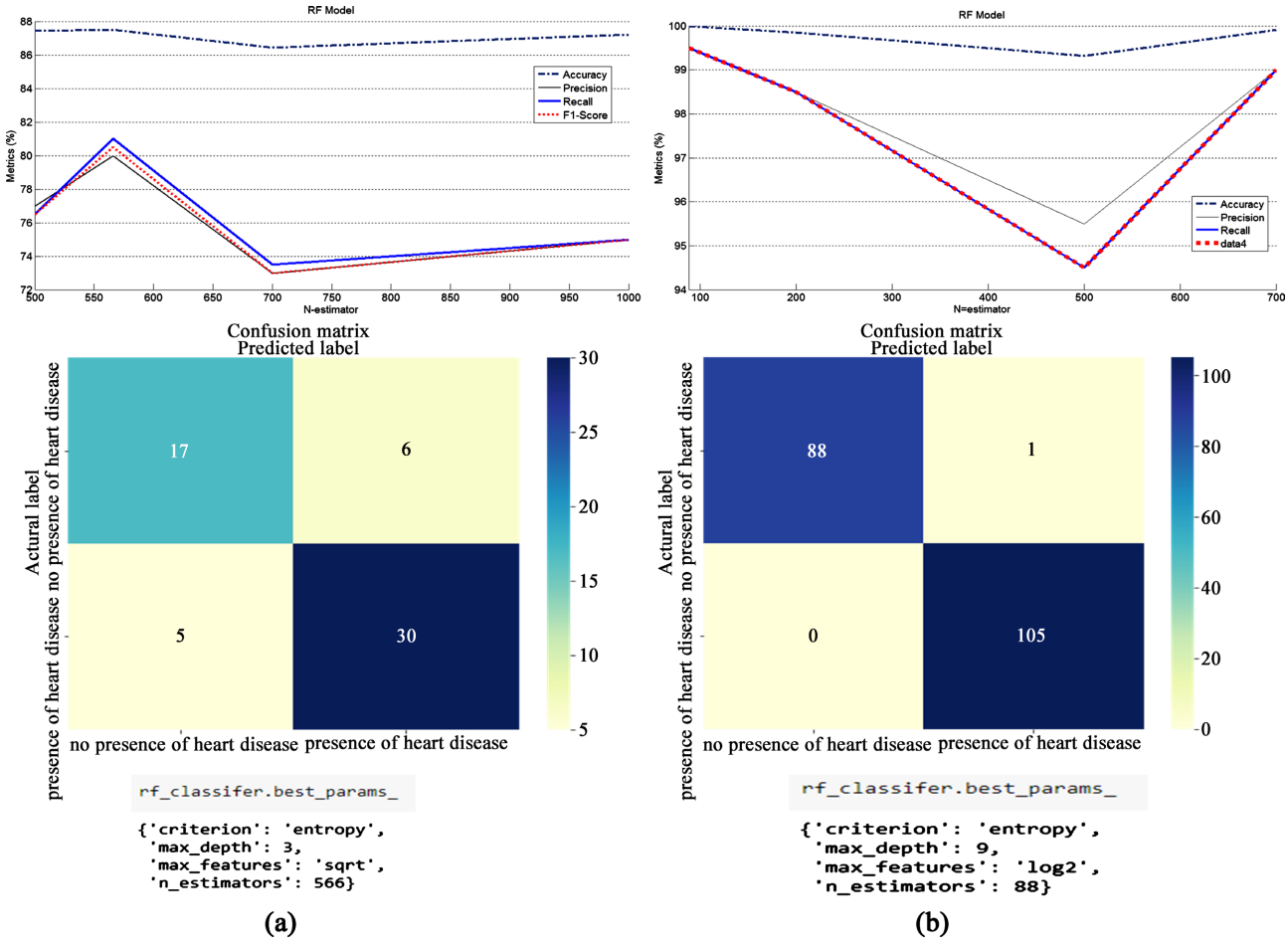

Figure 9. The performance metrics of RF with best N-estimator for the two datasets. (a) Dataset_1; (b) Dataset_2.

a little different for all values of C but the maximum records are shown when C = 100 for dataset_1 and C = 10 for dataset_2 where they will be the best parameters in this study.

Figure 11 indicates the performance metrics of XGB with varying Max-depth and the confusion matrix for both datasets. It can be observed that the accuracy of XGB will be the best when the value of Max-depth are 7 and 8 for the two datasets, respectively. Similarly the results of the other metrics, recall, precision, and F-score, are nearly the same for all values of Max-depth, but they achieved the maximum records when Max-depth is equal 7 and 8 for the two datasets, respectively. Hence, the best parameters in the rest of the study will be considered at Max-depth = 7 and Max-depth = 8.

The Performance metrics of all 5-algorithms across both datasets with the best parameters are shown in Table 3 and Performance accuracy of all five algorithms for the two datasets with the superior accuracy achieved with XGB model with 91.6% for dataset_1 and 100% for dataset_2 are shown in Figure 12.

It can be observed in Table 3 that the XGB model retains its higher performance on all four metrics with both datasets. This can be returned to the fact of binary classification task, so that the datasets are distributed fairly and then the

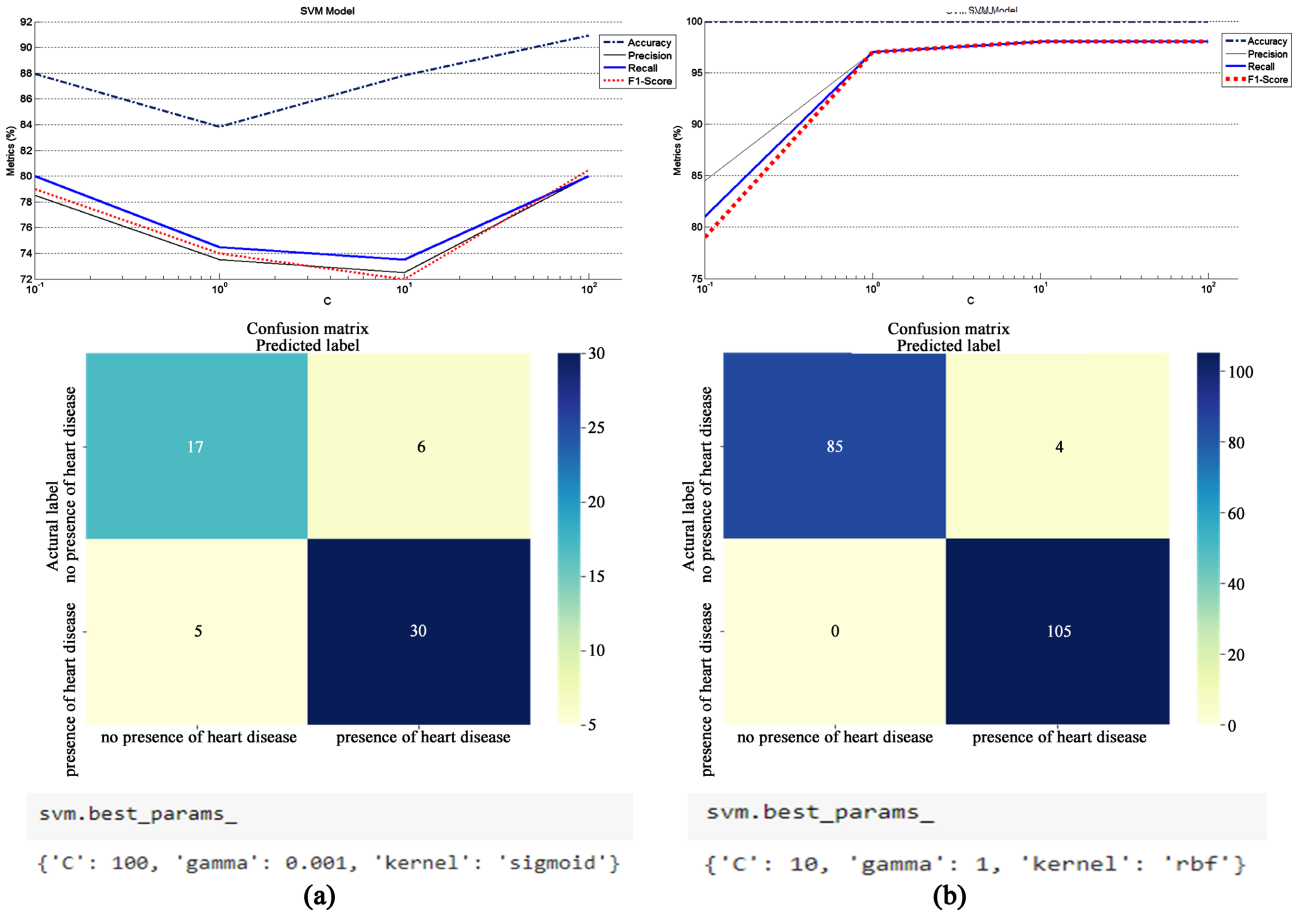

Figure 10. The performance metrics of SVM with best C for the two datasets. (a) Dataset_1; (b) Dataset_2.

accuracy performance carries on to other metrics.

It can be shown in Figure 12 that XGB model outperforms well other than machine learning models for both datasets. For detecting absence or presence of heart ailment, an accuracy of around 91.6% and 100% for the two datasets respectively are achieved which is better than the other algorithm.

Finally, a performance comparison of the proposed models with existing systems in term of accuracy are listed in Table 4. It can be clearly observed that this comparative research can achieve better results for the both datasets than other systems.

5. Conclusion

Heart ailment is life-threatening leading to deadly complications like heart attacks. To minimize this, the analysis research proposed to find out the best model that works well for both datasets selected where standard techniques for prognosis of heart diseases are tried. For this purpose, five machine learning algorithms were used. This was performed on two different datasets that related to heart disease datasets named "Cleveland dataset" both containing 14 features but different in number of recorded instances first one has 303 instances and second one has 1025 instances. The best results for each model on the two datasets are

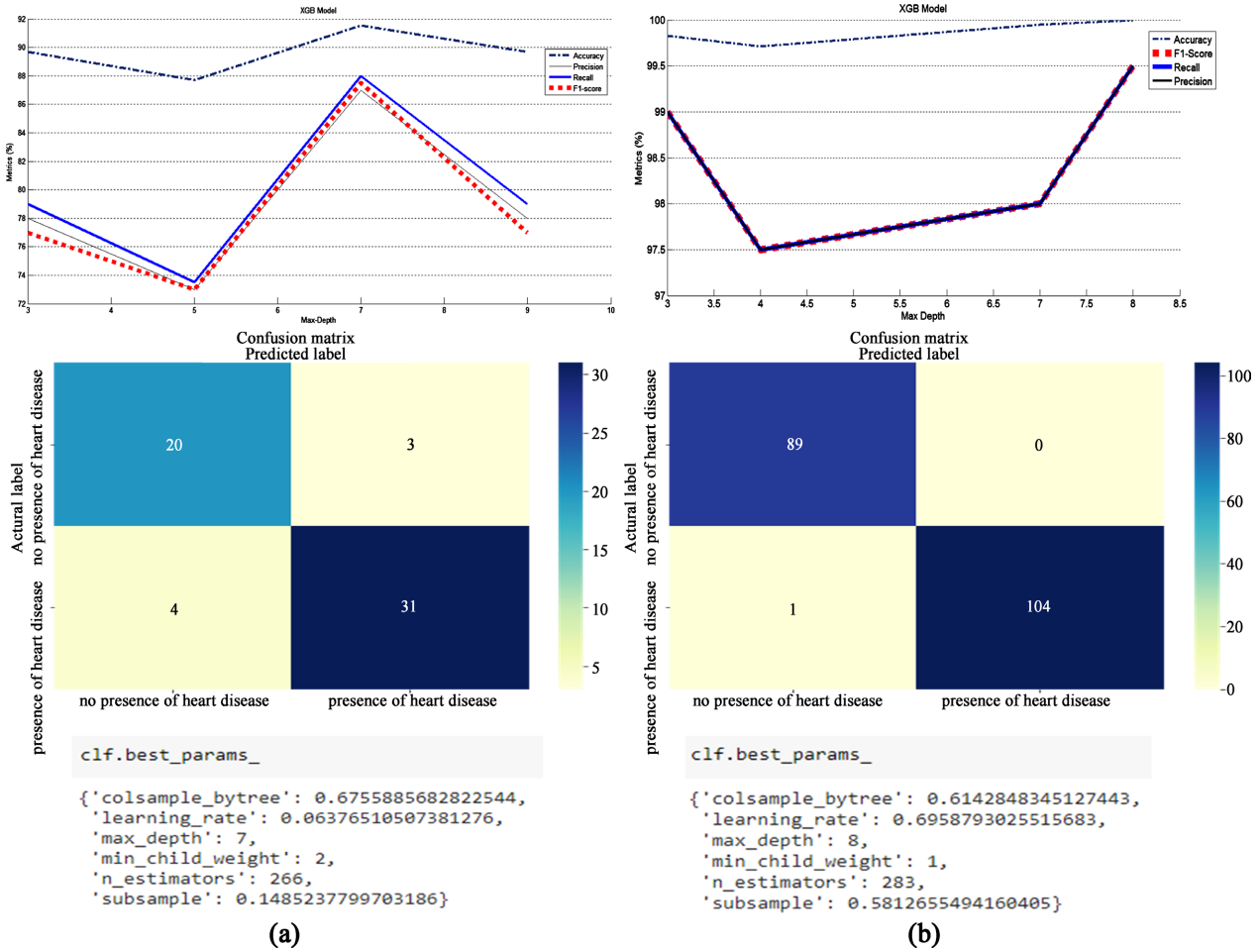

Figure 11. The performance metrics of XGB with best Max. depth for the two datasets. (a) Dataset_1; (b) Dataset_2.

Table 3. Performance metrics of all 5-algorithms across both datasets.

Algorithms	Accuracy (%)		Precision (%)		Recall (%)		F1-score (%)	
	Dataset 1	Dataset 2	Dataset 1	Dataset 2	Dataset 1	Dataset 2	Dataset 1	Dataset 2
KNN	86.8	100	76.5	100	77	100	76.5	100
LR	91	94	76.5	89.5	77	89	76.5	89
RF	87.5	100	80	99.5	80	99.5	80.5	99.5
SVM	90.9	100	80	98	80	98	80.5	98
XGB	91.6	100	87	99.5	88	99.5	87.5	99.5

Table 4. Comparison of different methods with previous researches.

	Method	Accuracy (%)	
		Dataset_1	Dataset_2
This Paper	KNN	86.8	100
	LR	91	94
	RF	87.5	100

Continued

	SVM	90.9	100
	XGB	91.6	100
Other Paper	KNN	84 [16], 84.21 [4]	99.7 [6], 100 [7]
	LR	84 [16], 85.07 [2], 76.31 [4]	87.36 [15], 89.6 [7]
	RF	82 [16], 86.60 [2], 84.21 [4]	89.14 [15], 100 [7]
	SVM	83 [16], 84.46 [2], 89.47 [4]	92.30 [15], 84.2 [6]
	Ada-Boost	82.34 [2]	89.88 [5], 84.3 [6]
	XGB	83 [1]	-

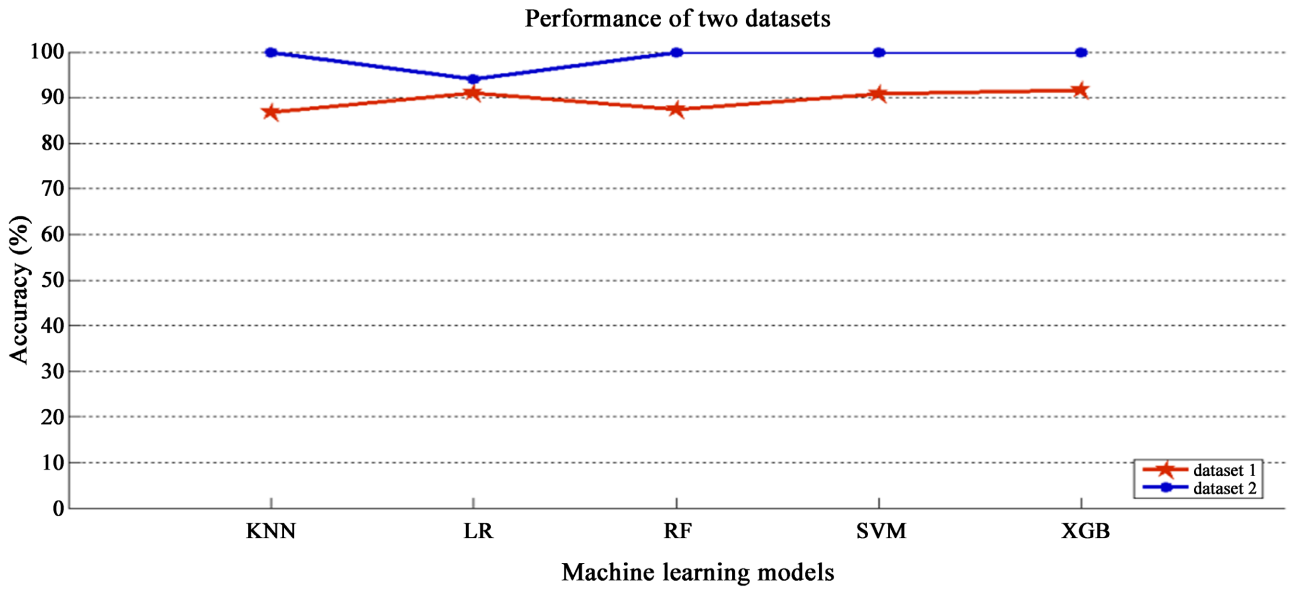

Figure 12. Performance accuracy of all five algorithms for the two datasets with the superior accuracy achieved with XGB algorithm with 91.6% for dataset_1 and 100% for dataset_2.

recorded. Finally, XGB model was the algorithm which achieved the best results with an accuracy of 91.6% and 100% for both datasets, respectively. The next scope is to use a large dataset with graphs of ECG for more accurate analysis and diagnosis of heart diseases.

Conflicts of Interest

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

References

[1] Beguma, S., Siddiqueb, F.A. and Tiwaric, R. (2021) A Study for Predicting Heart Disease Using Machine Learning. *Turkish Journal of Computer and Mathematics Education*, **12**, 4584-4592.

[2] Rani, P., Kumar, R., Sid Ahmed, N.M.O. and Jain, A. (2021) A Decision Support System for Heart Disease Prediction Based upon Machine Learning. *Journal of Re-*

liable Intelligent Environments, **7**, 263-275.
https://doi.org/10.1007/s40860-021-00133-6

[3] Sharmaa, C., Shambhub, S., Dasb, P. and Sakshid, S.J. (2021) Features Contributing towards Heart Disease Prediction Using Machine Learning. *Workshop on Advances in Computational Intelligence at ISIC* 2021, Delhi, 25-27 February 2021, 84-92.

[4] Patil, P.S. (2021) Automated Heart Disease Recognition System. *International Journal of Research in Engineering and Science*, **9**, 18-23.

[5] Choudhary, G. and Singh, S.N. (2020) Prediction of Heart Disease Using Machine Learning Algorithms. 2019 1*st International Conference on Innovations in Information and Communication Technology* (*ICIICT*), Chennai, 25-26 April 2019, 197-202.

[6] Almustafa, K.M. (2020) Prediction of Heart Disease and Classifiers' Sensitivity Analysis. *BMC Bioinformatics*, **21**, Article No. 278.
https://doi.org/10.1186/s12859-020-03626-y

[7] Ali, M.M., Paul, B.K., Ahmed, K., Bui, F.M., Quinn, J.M.W. and Moni, M.A. (2021) Heart Disease Prediction Using Supervised Machine Learning Algorithms: Performance Analysis and Comparison. *Computers in Biology and Medicine*, **136**, Article ID: 104672.

[8] Pandas. https://pandas.pydata.org

[9] Matplotlib 3.5.1 Documentation. https://matplotlib.org/stable/index.html

[10] UCI Machine Learning Repository.
https://archive.ics.uci.edu/ml/datasets/Heart+Disease

[11] Heart Disease Dataset. https://www.kaggle.com/johnsmith88/heart-disease-dataset

[12] Gawali, M.K. and Rambabu, C. (2020) IoT Model for Heart Disease Detection Using Machine Learning (ML) Techniques. In: Pawar, P.M., *et al.*, Eds., *Techno-Societal* 2020, Springer, Berlin, 399-409.

[13] Naik, A. and Naik, N. (2021) A Generalized Model for Cardiovascular Disease Classification Using Machine Learning Techniques. In: Chiplunkar, N.N. and Fukao, T., Eds., *Advances in Artificial Intelligence and Data Engineering*, Advances in Intelligent Systems and Computing, Vol. 1133, Springer, Berlin, 15-26.
https://doi.org/10.1007/978-981-15-3514-7_2

[14] Sathya, K. and Karthiban, R. (2020) Performance Analysis of Heart Disease Classification for Computer Diagnosis System. *International Conference on Computer Communication and Informatics* (*ICCCI* 2020), Coimbatore, 22-24 January 2020.

[15] Alotaibi, F.S. (2019) Implementation of Machine Learning Model to Predict Heart Failure Disease. (*IJACSA*) *International Journal of Advanced Computer Science and Applications*, **10**, 261-268. https://doi.org/10.14569/IJACSA.2019.0100637

[16] Doppala, B.P., Bhattacharyya, D., Chakravarthy, M. and Kim, T.H. (2021) A Hybrid Machine Learning Approach to Identify Coronary Diseases Using Feature Selection Mechanism on Heart Disease Dataset. Springer Science + Business Media, Berlin.
https://doi.org/10.1007/s10619-021-07329-y

[17] Bavani, B., Nirmala Sugirtha Rajini, S., Josephine, M.S. and Prasannakumari, V. (2021) Classification of Arrhythmia Disease Using Enhanced RNN Model. *Design Engineering*, No. 7, 4062-4072.