Asian European Journal of Probability and Statistics
Vol. 1, No-2, (2024), pp. 153-167

Prediction of Heart Disease: A Comparative Study of Machine Learning Algorithms

Reham Al-Maqtari[a] and Sumi A.R[b]

[a] *Department of Statistics, University of Kerala, India* [b] *Department of Statistics, University of Kerala, India*

ARTICLE HISTORY
Compiled April 16, 2025

Received 10 April 2024; Accepted 04 January 2025

ABSTRACT
Heart disease remains one of the leading causes of morbidity and mortality worldwide, affecting millions of people each year (World Health Organization, 2023). It encompasses various cardiovascular conditions, including coronary artery disease (CAD), heart failure, and arrhythmias, with CAD being the most prevalent. Early detection and prevention play a crucial role in reducing the burden of heart disease. Traditional diagnostic methods rely on clinical assessments and patient history, but advancements in artificial intelligence (AI) and machine learning (ML) have enabled more accurate and efficient heart disease prediction. By analyzing multiple risk factors such as blood pressure, cholesterol levels, lifestyle habits, and genetic predisposition, predictive models help in early diagnosis, personalized treatment, and reducing healthcare costs. This article explores the significance of heart disease prediction, highlighting the role of ML-algoriths in improving cardiovascular healthcare. This paper compares eight machine learning algorithms in order to improve predictive accuracy and offer a reliable instrument for early diagnosis. We have compared the following models: Artificial Neural Network (ANN), Decision Tree (DT), K-Nearest Neighbor (KNN), Naive Bayes (NB), Support Vector Machine (SVM), Random Forest (RF), Gradient Boosting (GB) and AdaBoost, using a dataset of 270 patients representing 14 clinical and demographic attributes from the University of California Irvine's Machine Learning Repository. The Artificial Neural Network model obtained the highest accuracy at 96.67%, with a precision of 100%, recall of 91.67%, and F1 score of 95.65%. These results underscore the potential of machine learning algorithms to enhance the early diagnosis of heart disease, thereby assisting healthcare professionals in making better decisions.

KEYWORDS
Heart Disease Prediction, Machine Learning, Artificial Neural Network, Decision Tree, K-Nearest Neighbors, Naive Bayes, Support Vector Machine, Random Forest, Gradient Boosting, AdaBoost.

1. Introduction

Heart diseases are presently the leading cause of death around the world, with the World Health Organization (WHO) estimating around 17.9 million deaths annually in 2020 ([27]). Heart disease refers to a variety of heart and blood vascular illnesses

CONTACT Reham Ali-Maqtari[a]. Email:almaqtarireham@gmail.com

that can cause heart attacks and heart failure . Several risk factors cause heart disease. Those include excessive blood pressure, cholesterol, diabetes, obesity, and smoking. These conditions affect the cardiovascular system, raising heart disease chances. Each form of heart disease necessitates different therapies, such as regular exercise, low-fat and low-sodium diets, blood pressure medications, and surgical operations to address blockages or abnormalities. If left untreated, heart disease can cause serious consequences such as heart attacks, strokes, and sudden cardiac arrest, all of which can be deadly. Preventing heart disease includes regular exercise, keeping a healthy weight, managing stress, quitting smoking, and treating diabetes, high blood pressure, and high cholesterol. Early detection of heart conditions is crucial, and machine learning algorithms are among the various methods utilized to enhance diagnosis and identification. These approaches allow hidden details to be obtained by identifying correlations among characteristics within the dataset, and it is a promising strategy for heart disease classification Canlas [5], Helma et al. [14], Lee et al. [18]. To provide quality care, it's important to accurately diagnose patients and identify appropriate treatments, while avoiding incorrect diagnoses Parthiban and Subramanian [21].

1.1. Abbreviations

ML - Machine Learning
CVDs - Cardiovascular Diseases
KNN - K-Nearest Neighbor
SVM - Support Vector Machine
ANN - Artificial Neural Network
AdaBoost - Adaptive Boosting
FP - False Positive
FN - False Negative
NB - Naive Bayes
RF - Random Forest
TN - True Negative
TP - True Positive
DT - Decision Tree
RT - Regression Techniques
GB - Gradient boosting

1.2. Objectives

This paper aims at utilizing machine learning algorithms to create an efficient predictive model that would be useful for the early diagnosis of heart diseases. The main objectives of this paper are the following:
1. To build, train, and evaluate machine learning models, i.e. ANN, DT, KNN, NB, SVM, RT, GB, and AdaBoost, for predicting heart disease.
2. To compare the performance of these algorithms based on their performance in fitting the data as well as their prediction accuracies.
3. To analyse the performance of the various classification algorithms, in determining whether a patient has heart disease or not. whether a patient has a heart disease or not.

2. Previous studies

bThe examination method, typically using a stethoscope, was the primary and conventional approach used by physicians to diagnose heart illness since it allowed them to differentiate between abnormal and normal cardiac sounds Guraksin et al. [8]. Using stethoscopes, referred to as the auscultation method, to listen to these heart sounds, doctors were able to diagnose every cardiac condition Sinha et al. [26]. There are several disadvantages to the auscultation method that medical professionals employ to identify cardiac disease. The ability of physicians to distinguish between different heart sounds and classify them is linked to their training and experience, which they acquire via extensive exams Kandaswamy et al. [16]. Various machine learning techniques have been developed for the identification of cardiovascular diseases (CVD) in addition to the manual method.

Abdar et al. [1] investigated the performance of several machine learning algorithms, including C 5.0 DT, SVM, KNN, and ANN, in predicting heart disease. The study used a dataset from the University of California, Irvine (UCI) that had 270 records and 13 characteristics. The data was separated into two sets: training (70%), and testing (30%). The models were tested according to specificity, sensitivity, precision, and accuracy. The findings showed that the C5.0 DT method had the greatest prediction accuracy of 93.02%, followed by KNN at 88.37%, SVM at 86.05%, and ANN at 80.23%. The study highlights the significance of variables including Thal, chest pain kind, and slope in predicting cardiac problems. Their findings indicate the C5.0 DT model's capability to provide accurate and interpretable predictions, which can help clinical practitioners make more informed decisions Abdar et al. [1].

Agrahara [2] investigated the prediction of heart disease using a variety of machine learning algorithms, including NB, SVM, DT, Logistic Regression, RF, and KNN. The dataset, which included 1004 rows and 14 columns, was taken from Kaggle and the UCI repository for the study. When these algorithms' performances were evaluated, the research discovered that the DT algorithm performed the best, with an accuracy of 98.29%, making it the most accurate and important technique for predicting heart disease. Agrahara [2] illustrates the effectiveness of machine learning methods in healthcare applications by highlighting their usefulness in improving the accuracy of heart disease detection.

In a similar context, Arghandabi and Shams [3] compared various machine learning algorithms—DT, Logistic Regression, KNN, SVM, and GB — for their effectiveness in predicting heart disease. The research employed the UCI heart disease dataset, consisting of 303 entries and 14 characteristics. The dataset was divided into 73% for training and 37% for testing. The study encompassed data preparation, visualising feature correlations using heat maps, and executing algorithms with Python and Jupyter Notebook. The models were assessed according to accuracy metrics to determine the best effective algorithm for early-stage heart disease prediction. The findings revealed that KNN attained the maximum accuracy of 85.7%, establishing it as the most effective algorithm among those evaluated. Their work emphasises the significance of choosing the correct machine learning model to enhance the early identification and treatment of cardiovascular illness Arghandabi and Shams [3].

Singh and Kumar [24] conducted a research evaluating several machine learning algorithms —Logistic Regression (LR), RF, Extreme Gradient Boosting (XGBoost), and ANN — were evaluated for predicting heart disease. The Cleveland heart disease dataset from the UCI Machine Learning Repository was used in the study. After careful preprocessing and hyperparameter tuning, the study discovered that Extreme Gradient

Boosting (XGBoost) had the best accuracy of 93.02%, making it the most successful model of those examined. RF came second with an accuracy of 88%; neural network with 85%; logistic regression with 84%. This work emphasises the need of appropriate data preparation and hyperparameter tweaking to improve model performance as well as the possibilities of advanced machine learning models in precisely predicting cardiac disease Singh and Kumar [24].

In another research, Hasan [10] examined the usefulness of several machine learning algorithms, including KNN, DT, GBN, Logistic Regression, and RF, in predicting cardiac disease. The study used a dataset with 303 cases and 14 variables. The findings showed that the RF algorithm had the highest accuracy (98.17%), making it the most effective model for heart disease prediction. This study emphasises machine learning's potential for improving early diagnosis and treatment, as well as the need of reliable prediction models in the health care sector.

In the analysis carried out by Pandita et al. [20], various machine learning algorithms, including SVM, KNN, , RF, and Logistic Regression, were evaluated for predicting heart disease. The study utilized the Cleveland heart disease dataset from the UCI repository, which includes 303 records and 76 attributes. After cleaning and preparing the data to address missing values and outliers, the dataset was split into 73% for training and 37% for testing. The data was scaled to promote homogeneity. The models were implemented using Python in a Jupyter Notebook environment. The performance of the algorithms was judged based on accuracy scores. The findings indicated that KNN obtained the maximum accuracy at 89.06%, while Logistic Regression had the lowest accuracy at 84.38%. This work highlights the potential of machine learning models in accurately predicting cardiac disease, highlighting the necessity for appropriate data preparation and algorithm selection to enhance early diagnosis and treatment results Pandita et al. [20].

In the most recent study, Hammoud et al. [9] investigated the performance of seven machine learning methods — Logistic Regression, Support Vector Classifier (SVC), KNN, RF, DT, NB, and Gradient Boosting Classifier — in predicting coronary heart disease. The study used a dataset of 1190 patients and 12 variables obtained from three medical clinics (Cleveland, Statlog, and Hungary). The models were assessed using 10-fold cross-validation to determine accuracy, specificity, sensitivity, and F1- scores. The RF model achieved the highest accuracy at 94.96% post-tuning, highlighting its superior performance in diagnosing heart diseases.

3. Methodology

3.1. Data Source

The data used in this research is a secondary data. The data used here comes from the University of California Irvine's Machine Learning Repository, comprising 270 rows and 14 columns. The 14 attributes used in this study are: Age, Sex, CP (chest pain type), trestbps (resting blood pressure), chol (serum cholesterol), fbs (fasting blood sugar), restecg (resting electrocardiographic results), thalach (maximum heart rate achieved), exang (exercise-induced angina), oldpeak (ST depression induced by exercise), slope (slope of the peak exercise ST segment), ca (number of major vessels colored by fluoroscopy), thal (thalassemia), and target (diagnosis of heart disease). This is shown in Table (1) below:

Table 1. Information about the Attributes used.

SN	Attribute	Description
1	Age	age of the patient
2	sex	male=1; female=0
3	CP	chest pain type (4 values)
4	trestbps	resting blood pressure(in mm Hg on admission to the hospital)
5	chol	serum cholestoral in mg/dl
6	fbs	fasting blood sugar > 120 mg/dl (1 = true ; 0 = false)
7	restecg	cresting electrocardiographic results
8	thalach	maximum heart rate achieved
9	exang	exercise induced angina (1 = yes; 0 = no)
10	oldpeak	ST depression induced by exercise relative to rest
11	slope	the slope of the peak exercise ST segment
12	ca	number of major vessels (0-3) colored by flourosopy
13	Thal	3 = normal; 6 = fixed defect; 7 = reversable defect
14	target	0 = no disease and 1 = disease

3.2. Dataset Description

Descriptive statistics are essential for defining the features of data. It simplifies the material to facilitate our understanding. Table (2) describes the statistical measurements of the clinical attributes with their measures, such as count of records, minimum (min) value, maximum (max) value, mean, and standard deviation (Std). For example, the age attribute has 54.433333 as a mean value and 9.109067 as a standard deviation, and the maximum and minimum age numbers are 77 and 29 years, respectively. These statistical measurements are also calculated for the rest of the attributes.

Table 2. Dataset Description.

	Count	Mean	Std	Min	25%	50%	75%	Max
Age	270.0	54.433333	9.109067	29.0	48.0	55.0	61.0	77.0
Sex	270.0	0.677778	0.468195	0.0	0.0	1.0	1.0	1.0
Chest pain type	270.0	3.174074	0.950090	1.0	3.0	3.0	4.0	4.0
BP	270.0	131.344444	17.861608	94.0	120.0	130.0	140.0	200.0
Cholesterol	270.0	249.659259	51.686237	126.0	213.0	245.0	280.0	564.0
FBS over 120	270.0	0.148148	0.355906	0.0	0.0	0.0	0.0	1.0
EKG results	270.0	1.022222	0.997891	0.0	0.0	2.0	2.0	2.0
Max HR	270.0	149.677778	23.165717	71.0	133.0	153.5	166.0	202.0
Exercise angina	270.0	0.329630	0.470952	0.0	0.0	0.0	1.0	1.0
ST depression	270.0	1.050000	1.145210	0.0	0.0	0.8	1.6	6.2
Slope of ST	270.0	1.585185	0.614390	1.0	1.0	2.0	2.0	3.0
Number of vessels fluro	270.0	0.670370	0.943896	0.0	0.0	0.0	1.0	3.0
Thallium	270.0	4.696296	1.940659	3.0	3.0	3.07	7.0	7.0
Heart Disease	270.0	0.444444	0.497827	0.0	0.0	0.0	1.0	1.0

3.3. Histogram of Dataset

A histogram illustrates and analyses the distribution of data samples. Histograms can exhibit several forms, including uniform, normal, left-skewed, or right-skewed. Figure 1 below illustrates the normally distributed histograms that categorize all characteristics

within the specified value range. The x-axis denotes the characteristic of the attribute, while the y-axis indicates the value of that attribute.

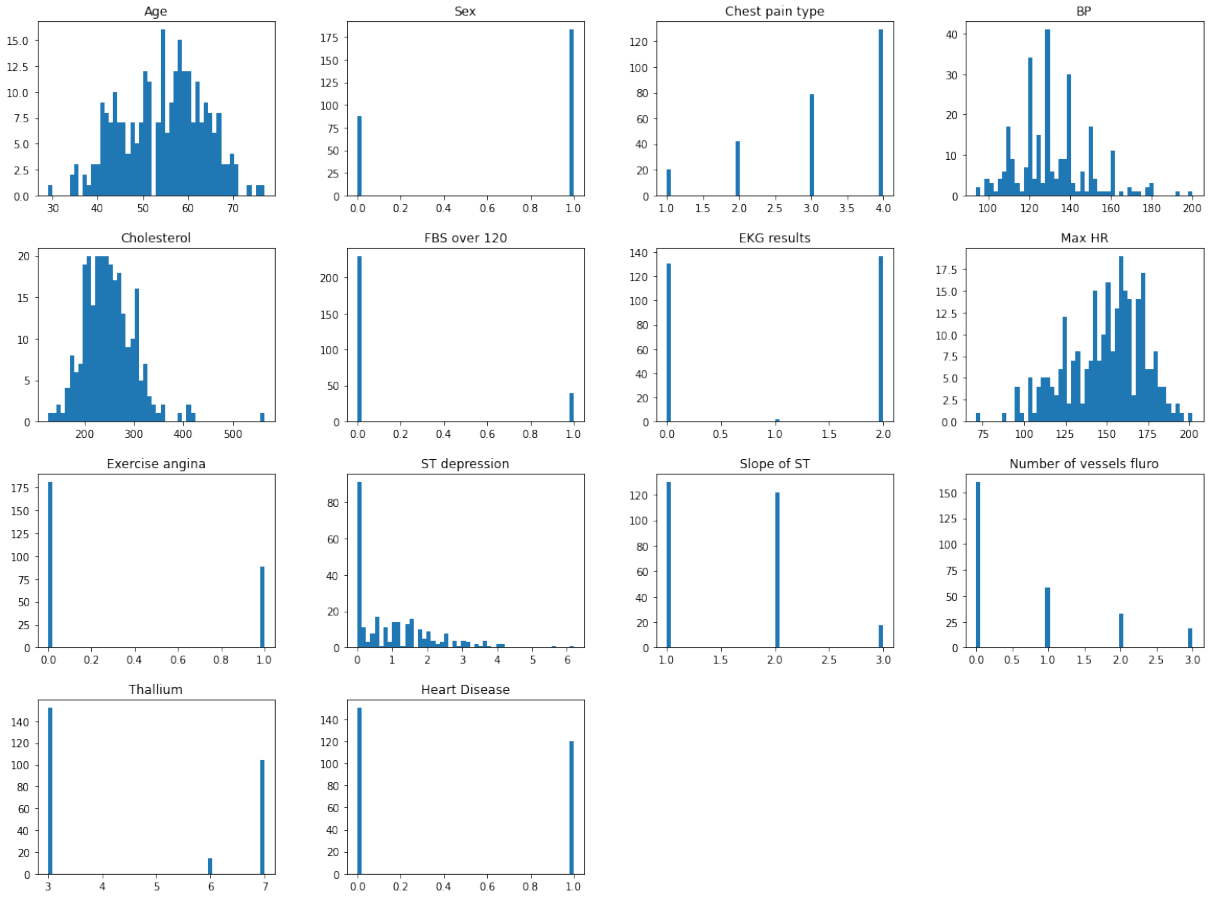


Figure 1. Histogram of attributes.

3.4. Correlation Coefficient Analysis

The correlation coefficient analysis (CCA) method is used to identify and plot the relationship among the dataset's attributes Hussain and Naaz [15]. A dataset is considered good if a strong association/relationship exists between the set of independent and dependent attributes. Fig.2 presents the CCA of all attributes used to predict disease, and the range of relationships exists between +1 to -1 within the x-axis and y-axis. The cell value indicates the degree of relationship between the intersecting attributes. For example, a strong positive correlation of 0.61 between ST Depression and Slope of the Peak Exercise ST Segment , highlighting a significant positive relationship between these variables.

3.5. Data Preprocessing

The predictive model's accuracy and performance are influenced by the quality of the dataset and the preprocessing techniques, in addition to the algorithms employed.

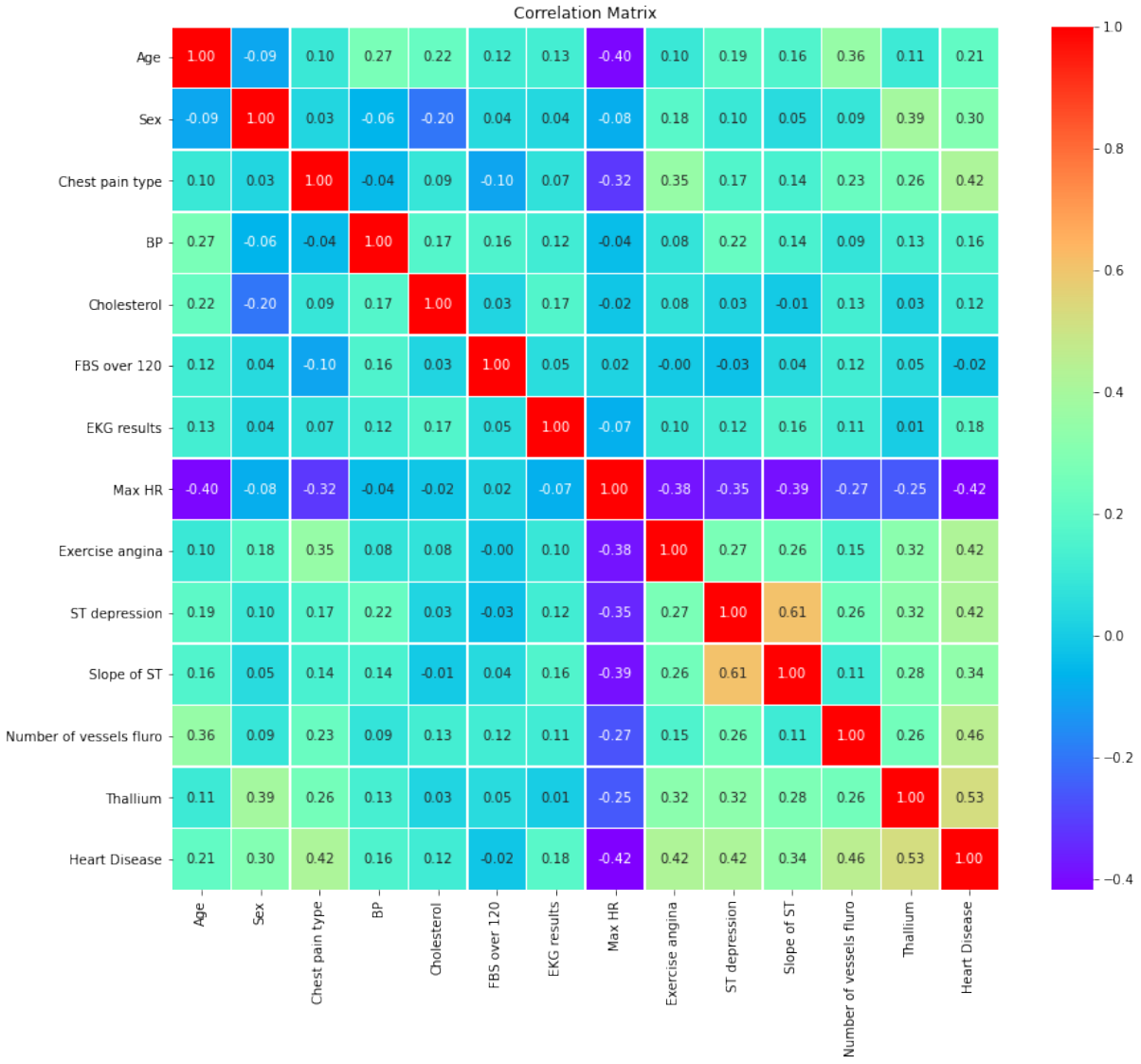


Figure 2. Correlation Coefficient Analysis.

Preprocessing is the term used to describe the procedures that are implemented on the dataset prior to the application of machine learning algorithms. Errors, missing data, redundancies, noise, and numerous other issues can render datasets unsuitable for direct use by the machine learning algorithm. The dataset's quantity is an additional consideration. Certain datasets possess numerous attributes that complicate the algorithm's ability to analyse, identify patterns, or generate precise predictions. Data preprocessing steps includes: data cleaning, data transformation, missing values imputation, data normalization, feature selection, and other steps depending on the nature of the dataset Garcıa et al. [7]. None of the previous steps were done in this paper because the dataset was complete, consistent, and already structured for analysis. The only preprocessing step involved was encoding the categorical values in the 'Heart Disease' column by replacing 'Presence' with '1' and 'Absence' with '0'.

3.6. Training and testing procedures

This study was conducted using a dataset that consists of 270 patients with 14 variables. We segmented the data into 80% for training and 20% for testing in order to evaluate the performance of different machine learning algorithms. This means that out of 270 , 54 patient records were used for testing, and 216 records were used for training. However, it was necessary to apply a data strategy again since the early testing accuracy was quite low. As a result, the data portioning changed to 11% testing and 89% training and therefore 30 patient records were available for testing while 240 were for training. This alteration attempted to get rid of the constraint of inadequate training dataset by providing enough information for the models to learn from the current data more effectively. Within the training stage, the models learned the patterns and relationships between the input features and the target variable predicting the presence or absence of heart disease. Cross-validation was used during training for a fair assessment and to minimize the chances of over fitting. The testing phase used the other 30 consecutive patient records helping to test the algorithms on the cases which were not previously encountered further optimizing the chances of real-world suitability. This ability training finally revealed that the ANN algorithm obtained the highest accuracy of 96 %. As a result of this in and out process of data segregation and analysis, the provisional models did not only outperform in the parameter specified but were also proven to be safe and transferrable to any other population also reliable and generalizable to broader patient populations.

3.7. Performance measures used in this research

• Some of the evaluation metrics used in data mining include accuracy, precision, recall as well as F-measure. These metrics make use of terms like True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) as noted by [25].
• Accuracy is defined as the number of accurately classified instance divided by a total number of instance in the dataset as in (1).

Accuracy = (TP+TN)/(TP+FP+FN+TN)... (1)

• Precision is the average probability of relevant retrieval as described in (2).

Precision= (TP)/(TP+FP) ... (2)

• The recall is defined as the average probability of complete retrieval as defined in (3).

Recall= (TP)/(TP+FN) ... (3)

• F1- Score is the calculated by using both precision and recall as shown in (4).

F1-Score =(2xRecall x Precision)/(Recall+ Precision) ... (4)

3.8. Machine Learning (ML)

Machine learning is a domain of artificial intelligence that allows computer systems to acquire knowledge and enhance performance via experience, without explicit programming. It entails the development of algorithms and statistical models that enable computers to execute specified tasks efficiently through the use of data. Machine learning encompasses many applications such as computer vision, natural language processing, predictive analytics, and autonomous systems Heaton [13]. Machine learning fundamentally relies on algorithms that identify patterns in data to facilitate predictions or judgements. These algorithms may be classified into three main categories: supervised, unsupervised, and reinforcement learning. Supervised learning algorithms are trained

on labelled datasets, where the expected result is known, enabling the system to learn the mapping from input to the appropriate output. On the other hand, unsupervised learning algorithms seek to identify concealed patterns within unlabelled data. Reinforcement learning entails an agent engaging with an environment and acquiring knowledge via a framework of rewards and punishments Russell and Norvig [23].

3.8.1. Artificial Neural Network (ANN)

An ANN is a series of algorithms that attempts to recognize underlying relationships in a set of data through a process that mimics the way the human brain operates. A "neuron" is a mathematical function that collects and classifies information according to a specific architecture. The network bears a strong resemblance to statistical methods such as curve fitting and regression analysis. A neural network contains layers of interconnected nodes. The layers can be single or multiple. In a single-layered neural network, each node is a perceptron and is similar to a multiple linear regression. The perceptron feeds the signal produced by a multiple linear regression into an activation function that may be nonlinear. In a multi-layered perceptron (MLP), perceptrons are arranged in interconnected layers. The input layer collects input patterns. The output layer has classifications or output signals to which input patterns may map (https://www.investopedia.com/terms/n/neuralnetwork.asp).

3.8.2. Decision Tree (DT)

A DT is a tree-based model that is mostly used for classification tasks but can also be used for regression. It is an example of supervised machine learning and is often used to solve classification problems. Hassan et al. [12] say that decision trees are better at making decisions than other widely used machine learning methods.

3.8.3. K-Nearest Neighbours(KNN)

A KNN algorithm sorts cases into groups based on how similar they are to other cases. A 'neighbour' is a case that is close to another one. As soon as a new case is added to the model, its distance from every other case in the model is found. By using this classification, the case is named as the closest neighbour, which means it is the most similar the other case. So, it puts the case in the group with the cases that are closest to it. It is also possible for the algorithm to keep figuring out numbers for a goal. In this case, the predicted value of the new case is found by taking the average or median goal value of the close neighbour Yazdani et al. [28].

3.8.4. Naive Bayes (NB)

According to Ning et al. [19], an NB is a widely used classifier method that mathematically applies the Bayes theorem. P(B/A) = (P(A/B)P(B))/(P(A))
The above Bayes theorem states that the dependent feature vectors, x_1 through x_j, and the supplied class variable, y, are related.

$$\mathrm{P}(y/x_1, \cdots, x_j) = \frac{p(x_1,\cdots,x_j/y)}{p(x_1,\cdots,x_j)}$$

NB's greatest benefit is that, in comparison to other machine learning algorithms, it consumes less computing time. Compared to numerical input variables, it performs better with categorical input variables. Furthermore, it assumes that each characteris-

tic is an independent variable, which makes it challenging to use in real-world settings Le et al. [17].

3.8.5. Support Vector Machine (SVM)

An SVM is a supervised machine learning method. It is commonly used for classification difficulties. It operates on the hyperplane concept, with the goal of classifying provided data by constructing a hyperplane Zhao et al. [29]. It uses a hyperplane in an n-dimensional space to distinguish between classes. As a consequence, an ideal hyperplane is developed iteratively in order to minimise mistakes. The goal here is to design a hyperplane with maximal margins that divides the dataset efficiently.

3.8.6. Random Forest (RF)

An RF is a data classification approach that uses group learning and is based on DT Breiman [4]. During the training stage, it generates a great number of trees as well as a forest of DTs. During the testing period, each tree in the forest forecasts the class label for each individual instance Hasan et al. [11]. When each tree predicts a class label, majority voting is employed to make the ultimate selection for each test dataset Quinlan [22]. The class label with the most votes is judged the best appropriate for the test data. This process is repeated for each data point in the data collection.

3.8.7. Gradient Boosting (GB)

The GB method begins by fitting a basic model, such as a DT, to the data. The residuals of the first model are then calculated, and a new DT is fitted to them. The new model is then integrated with the prior model by calculating the weighted average of their forecasts. This procedure is done repeatedly, with each new model fitting to the preceding model's residuals, until a predefined number of trees are achieved or the residuals cannot be reduced any more Zhao et al. [30].

3.8.8. AdaBoost

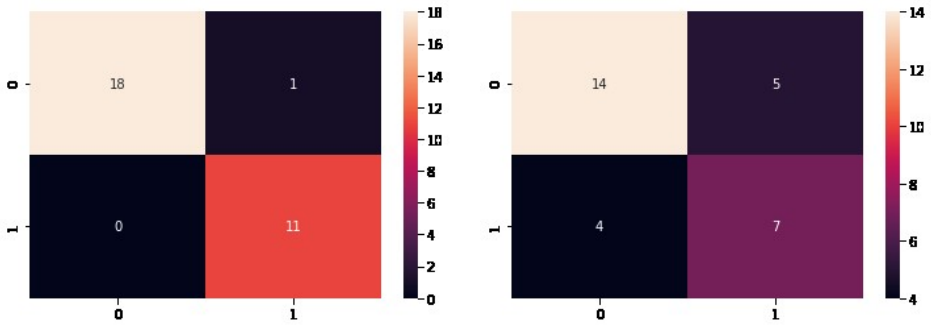
The AdaBoost algorithm is a type of boosting classification method that can transform a set of 'weak' classifiers into a 'strong' classifier. To train a base classifier from the first training data, these algorithms often start with a base classified algorithm whose classification skill is just slightly better than random guessing. The sample weight is then adjusted based on the base classifier's results, causing the erroneously categorised samples to receive more attention. Finally, the altered samples are utilised to train the next base learner. After iterations, weights are applied to the base learners to create the final classifier Freund and Schapire [6].

3.9. Findings

For better understanding, confusion matrices for each algorithm were used. A confusion matrix, also known as an error matrix, is a specific table layout that allows visualization of the performance of an algorithm. In this research the confusion matrix has two classes, so this is a 2*2 confusion matrix. Here:
Class 0 = Does not have heart disease.
Class 1 = Has heart disease .
Confusion Matrix for the cases of having heart disease and not having heart disease.

Table 3. Confusion matrix.

	0	1
0	True Negative(TN)	False Positive(FP)
1	False Negative(FN)	True Positive(TP)



(a) Confusion matrix ANN

(b) Confusion matrix DT.

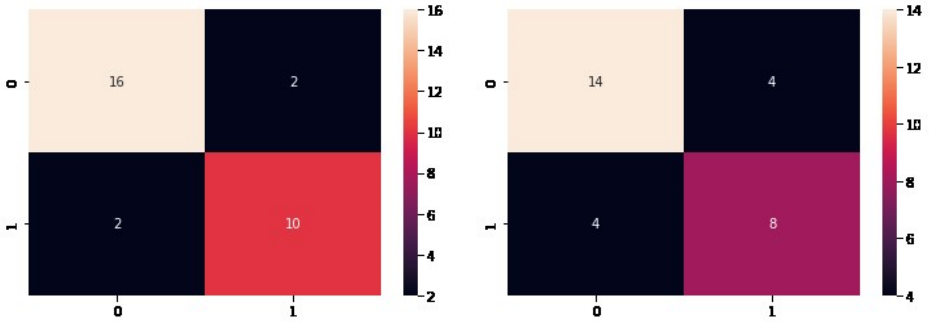
(c) Confusion matrix KNN

(d) Confusion matrix NB

(e) Confusion matrix SVM

(f) Confusion matrix RF

The confusion matrix for various machine learning algorithms is shown in Figure 3. The percentage of predicted values and the percentage of real values are represented in terms of true positives/negatives and false positives/negatives in this graph. The proposed ANN algorithm accurately estimates 11 percent (true positive) of heart disease cases, with just 0 percent (false positive) misclassification; for non-heart disease cases, the ANN algorithm gives 1 percent (false negative) of misclassification and 18 percent (true negative) of precise classification as illustrated in Figure.(a). Similarly, Figure.(b), Figure.(c), Figure.(d), Figure.(e), Figure.(f), Figure.(g) and Figure.(h) depict the confusion matrix of ANN, DT, KNN, NB, SVM, RF, GB, and AdaBoost algorithms respectively, with lower true positives/negatives and false



(a) Confusion matrix GB (b) Confusion matrix AdaBoost

Figure 3. Normalized Confusion matrix of ML algorithms

positives/negatives rate than the ANN algorithm.

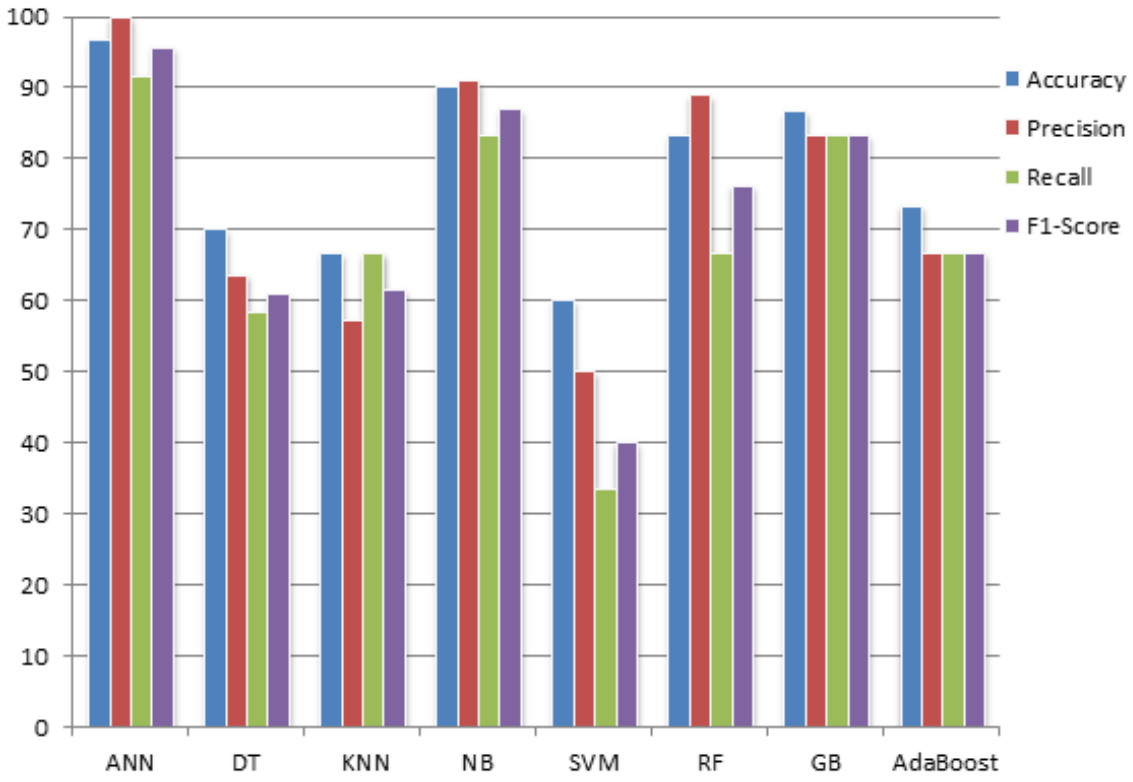


Figure 4. Visual comparison of the performance of the algorithms.

Figure 4 below shows performance of various evaluation parameters in the prediction of heart disease. The experimental results show the comparison of ANN, DT, NB,

KNN, SVM ,RF, GB and AdaBoost classifiers and evaluate the performance on the bases of accuracy, precision, recall and F measure. In all classifiers ANN got the highest accuracy of 96.666667.

3.10. Conclusion

The aim of this research paper is to analyze the performance of various classification algorithms and in doing so find the most accurate algorithm for predicting whether a patient would develop heart disease or not. This research was done using techniques of ANN, DT, KNN, NB, SVM, RF, GB, Ada Boost. On the UCI dataset. Dataset was split into training and test data and models were trained and the accuracy was noted using Python. A comparison of the performance of the algorithms is obtained, and presented in Table 4. The model Artificial Neural Network has been selected out of all the other models because it performs best with an accuracy of 96.666667, precision of 100.00, recall of 91.666667, and F1-score of 95.652174.
The model Neural Network has been selected out of all other model because it performs best with an accuracy 96.666667 , precision 100.00 , recall 91.666667 and f1 score 95.652174.

Table 4. Comparison of The Performance of the Algorithms.

Algorithm	Accuracy	Precision	Recall	F1-Score
ANN	96.66667	100.00	91.666667	95.652174
DT	70.000000	63.64	58.333333	60.869565
KNN	66.666667	57.14	66.666667	61.538462
NB	90.000000	90.91	83.333333	86.956522
SVM	60.000000	50.00	33.333333	40.000000
RF	83.333333	88.89	66.666667	76.190476
GB	86.666667	83.33	83.333333	83.333333
AdaBoost	73.333333	66.67	66.666667	66.666667

3.11. Prediction

Out of all the algorithms chosen, Artificial Neural Network performs the best. So, predicted Heart Disease possibility of 9 samples using the ANN are given below; where 0 indicates the absence of heart disease and 1 indicates the presence of heart disease .

	Age	Sex	Chest pain type	BP	Cholesterol	FBS over 120	EKG results	Max HR	Exercise angina	ST depression	Slope of ST	Number of vessels fluro	Thallium	predictions
0	70	1	4	130	322	0	2	109	0	2.4	2	3	3	1
1	67	0	3	115	564	0	2	160	0	1.6	2	0	7	1
2	57	1	2	124	261	0	0	141	0	0.3	1	0	7	1
3	64	1	4	128	263	0	0	105	1	0.2	2	1	7	1
4	74	0	2	120	269	0	2	121	1	0.2	1	1	3	0
5	65	1	4	120	177	0	0	140	0	0.4	1	0	7	0
6	56	1	3	130	256	1	2	142	1	0.6	2	1	6	1
7	59	1	4	110	239	0	2	142	1	1.2	2	1	7	1
8	60	1	4	140	293	0	2	170	0	1.2	2	2	7	1

Figure 5. Prediction

Acknowledgments
Authors would like to express their sincere thanks to both the anonymous referees for their valuable comments that have helped to improved the quality of this paper.

References

[1] Abdar, M., Kalhori, S. R. N., Sutikno, T., Subroto, I. M. I., and Arji, G. (2015). Comparing performance of data mining algorithms in prediction heart diseases.
[2] Agrahara, A. (2020). Heart disease prediction using machine learning algorithms. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 6(4):137–149.
[3] Arghandabi, H. and Shams, P. (2020). A comparative study of machine learning algorithms for the prediction of heart disease. *International Journal for Research in Applied Science and Engineering Technology*, 8(12):677–683.
[4] Breiman, L. (2001). Random forests. *Machine learning*, 45:5–32.
[5] Canlas, R. (2009). Data mining in healthcare: Current applications and issues. *School of Information Systems & Management, Carnegie Mellon University, Australia.*
[6] Freund, Y. and Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, 55(1):119–139.
[7] García, S., Ramírez-Gallego, S., Luengo, J., Benítez, J. M., and Herrera, F. (2016). Big data preprocessing: methods and prospects. *Big data analytics*, 1:1–22.
[8] Guraksin, G. E., Ergun, U., and Deperlioglu, O. (2010). Classification of the heart sounds via artificial neural network. *International Journal of Reasoning-Based Intelligent Systems*, 2(3-4):272–278.
[9] Hammoud, A., Karaki, A., Tafreshi, R., Abdulla, S., and Wahid, M. (2024). Coronary heart disease prediction: a comparative study of machine learning algorithms. *Journal of Advances in Information Technology*, 15(1):27–32.
[10] Hasan, R. (2021). Comparative analysis of machine learning algorithms for heart disease prediction. In *ITM Web of Conferences*, volume 40, page 03007. EDP Sciences.

[11] Hasan, S., Mamun, M., Uddin, M., and Hossain, M. (2018). Comparative analysis of classification approaches for heart disease prediction. In *2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)*, pages 1–4. IEEE.

[12] Hassan, M. M., Mollick, S., and Yasmin, F. (2022). An unsupervised cluster-based feature grouping model for early diabetes detection. *Healthcare Analytics*, 2:100112.

[13] Heaton, J. (2018). Ian goodfellow, yoshua bengio, and aaron courville: Deep learning: The mit press, 2016, 800 pp, isbn: 0262035618. *Genetic programming and evolvable machines*, 19(1):305–307.

[14] Helma, C., Gottmann, E., and Kramer, S. (2000). Knowledge discovery and data mining in toxicology. *Statistical methods in medical research*, 9(4):329–358.

[15] Hussain, A. and Naaz, S. (2021). Prediction of diabetes mellitus: comparative study of various machine learning models. In *International Conference on Innovative Computing and Communications: Proceedings of ICICC 2020, Volume 2*, pages 103–115. Springer.

[16] Kandaswamy, A., Kumar, C. S., Ramanathan, R. P., Jayaraman, S., and Malmurugan, N. (2004). Neural classification of lung sounds using wavelet coefficients. *Computers in biology and medicine*, 34(6):523–537.

[17] Le, C. C., Prasad, P., Alsadoon, A., Pham, L., and Elchouemi, A. (2019). Text classification: Naïve bayes classifier with sentiment lexicon. *IAENG International journal of computer science*, 46(2):141–148.

[18] Lee, I.-N., Liao, S.-C., and Embrechts, M. (2000). Data mining techniques applied to medical information. *Medical informatics and the Internet in medicine*, 25(2):81–102.

[19] Ning, B., Junwei, W., and Feng, H. (2019). Spam message classification based on the naïve bayes classification algorithm. *IAENG International Journal of Computer Science*, 46(1):46–53.

[20] Pandita, A., Vashisht, S., Tyagi, A., and Yadav, S. (2021). Prediction of heart disease using machine learning algorithms. *International Journal for Research in Applied Science and Engineering Technology*, 9(5):2422–2429.

[21] Parthiban, L. and Subramanian, R. (2008). Intelligent heart disease prediction system using canfis and genetic algorithm. *International Journal of Biological, Biomedical and Medical Sciences*, 3(3).

[22] Quinlan, J. R. (1986). Induction of decision trees. *Machine learning*, 1:81–106.

[23] Russell, S. J. and Norvig, P. (2016). *Artificial intelligence: a modern approach*. Pearson.

[24] Singh, A. and Kumar, R. (2020). Heart disease prediction using machine learning algorithms. In *2020 international conference on electrical and electronics engineering (ICE3)*, pages 452–457. IEEE.

[25] Sinha, P., Sinha, P., et al. (2015). Comparative study of chronic kidney disease prediction using knn and svm. *International Journal of Engineering Research and Technology*, 4(12):608–12.

[26] Sinha, R. K., Aggarwal, Y., and Das, B. N. (2007). Backpropagation artificial neural network classifier to detect changes in heart sound due to mitral valve regurgitation. *Journal of Medical Systems*, 31:205–209.

[27] WHO (2021). Health topics: Cardiovascular diseases. *Fact Sheet. Available online: http://www. who. int/cardiovascular_diseases/en/(accessed on 11 August 2024)*.

[28] Yazdani, A., Ebrahimi, T., and Hoffmann, U. (2009). Classification of eeg signals using dempster shafer theory and a k-nearest neighbor classifier. In *2009 4th International IEEE/EMBS Conference on Neural Engineering*, pages 327–330. IEEE.

[29] Zeinulla, E., Bekbayeva, K., and Yazici, A. (2020). Effective diagnosis of heart disease imposed by incomplete data based on fuzzy random forest. In *2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pages 1–9. IEEE.

[30] Zhao, P., Xue, J., Xu, X., Wang, L., and Chen, D. (2022). Logistic model and gradient boosting machine model for physical therapy of lumbar disc herniation. *Computational and Mathematical Methods in Medicine*, 2022(1):4799248.