Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 15s (2025)
https://www.jneonatalsurg.com

Prediction Model for Detection of Heart Disease Stages Using Machine Learning Approaches

Gashaw Alemu[1], Eyasu Tafere[2], Rakesh K. Sharma[3]

[1, 2] Faculty of Computing and Software Engineering, Arba Minch University, Ethiopia

Email ID: gashaw.alemu2@gmail.com,

Email ID: metku218@gmail.com,

[3]Department of Computer Science and Engineering Technology, University of Maryland Eastern Shore, Princess Anne, MD 21853 (USA)

Email ID: rsharma@umes.edu

Cite this paper as: Gashaw Alemu, Eyasu Tafere, Rakesh K. Sharma, (2025) Prediction Model for Detection of Heart Disease Stages Using Machine Learning Approaches. *Journal of Neonatal Surgery*, 14 (15s), 1251-1265.

ABSTRACT

Health sector reports reveal that heart patients are exponentially growing concerning time. Although there are several types of medication and treatment mechanisms available in the health sector. However, researcher's efforts are trying to explore the applications of machine learning as an emerging area of research. The healthcare sector has accumulated enormous amounts of data that may contain some hidden insights or useful health indicators, which may later be useful for effective decision-making in treatment and medication processes. In this research, an effective heart disease prediction system model is developed using machine learning approaches. The designed system model for the smart prediction about the risk stage (risk level) of heart disease can be a new and alternative instrument in the medical sector for heart disease diagnosis. This research used 14 attributes like Sex, Age, Chest Pain, Family History, Past History, Cholesterol, Fasting Blood Sugar, Resting ECG, Slope ST, Heart Rate, Pulse Rate, Blood Pressure, CBC, and Diagnosis. The proposed model predicts the likelihood or the appropriate stage of the risk of heart patients. This mechanism can help doctors to minimize the later-stage risks and consequences. These hidden factors or patterns are required to be discovered and analyzed to reveal the hidden insights or patterns from the unused and unexplored data especially in Ethiopia hospitals. Initially, it was aimed to categorize the status of heart disease in terms of major or minor stages of risks. We used the KDD process model to find out and interpret the discovered patterns from data repositories. Decision trees (J48 and Random Forest), Bayes (Naïve Bayes), and ANN (Multilayer perceptron) algorithms were used for the classification of data mining tasks. After experimentation, the overall accuracy of the tested classifiers was achieved 90% + in approximation. It was revealed that the ANN (multilayer perceptron) classifier relatively produces higher classification accuracy (97.0%) than the other selected four classifiers. The classification was based on the obtained attributes and revealed that prediction rates are not uniform among all the classifiers and the selected attributes. Finally, the study concluded that machine learning can be used as a new technique to discover the hidden patterns or insights in the heart patient's massive amount of data to determine the risk stage and this can help in minimizing the risks at an early stage.

Keywords: *Artificial Neural Network, Decision Tree, Heart Disease, Machine Learning, Naive Bayes, Prediction, Risk Stage*

1. INTRODUCTION

Today the automation is taking place in health sectors of both the developed and the developing countries. Artificial Intelligence and other computing technologies are helping in transforming healthcare support and the services starting from diseases diagnosis, diseases detection, communication and collaboraton among the healthcare institutions [1] [2] [3]. The disease prediction using machine learning techniques, i.e., the term machine learning refers to the automated detection of meaningful patterns in the collected data. In the past few decades, it's become a standard tool in almost any task that needs information extraction from large data sets. It is also widely used in scientific applications such as bioinformatics, medicine, and astronomy. Machine learning techniques are concerned with endowing programs with the ability to "learn" and adapt new things from the specified data sets in the determined application domain [4]. The research topic concerns designing a prediction model of heart disease stage detection by processing a patient's dataset and information of Patients to whom we want to predict the possibility of the prevalence of heart disease. Accurate and on-time diagnosis of heart conditions is vital for coronary failure prevention and treatment. To classify healthy people and people with heart disease, noninvasive-based methods such as machine learning are reliable and efficient [5].

The diagnosis of heart disease is usually based on signs, symptoms, and physical examination of the patient. However, after identifying the heart disease absence it also needs to know the stage. The available heart disease dataset consists of both numerical and categorical data. Before further processing, cleaning, and filtering are applied to these records to filter the irrelevant data from the database [6].

Nowadays, people face various diseases including heart disease due to the different cases and living habits. So, the prediction of any disease at an earlier stage becomes an important task. However, the prediction based on symptoms becomes too difficult for doctors. The correct prediction of disease is the most challenging task. To overcome this problem, machine learning plays an important and efficient way to predict the disease [7] towards solving the above-mentioned problem. The questions answered in this research study are: Which factors cause heart disease stage? Which are the most influencing determinant attributes affecting the heart disease stage? Which machine learning technique or algorithm is more appropriate to construct a heart disease stage predictive model that can be used for heart disease stage prediction? Which patterns are more interesting to develop a heart disease stage detection system? The objective of this research is to design a predictive model for predicting heart disease stage (major or minor) that enhances and determines new knowledge by using machine learning algorithms and techniques.

The significance of the study is to investigate the potential applicability of machine learning techniques to develop a predictive model that could help healthcare centers identify and categorize patients with major heart disease or minor heart disease. This is significantly used to minimize patients' risk of death and predict models to enhance their performance in diagnosing them. The study outcome will help the patients to improve themselves in the future and experts to give some extra care for patients who have major heart disease or not. The limitations of this study were faced by the researcher where the patient’s data was not well recorded and organized, the support of experts in this domain was not sufficient, and the limitation of reference materials in this specific area.

2. REVIEW OF LITERATURE

In this study, the researchers conducted to identify factors that will affect the chest disease and predict the chest disease as either normal or have the disease by using the decision tree and ordinal regression approaches. The result of a study revealed 66.8% of the chest disease which were predicted to have chest disease while 33.2% were predicted to be normal. It is observed that a much larger percentage of the patients were likely to have the diseases and there is also a higher likelihood and less percentage of the analysis shows as prediction result is normal. The researcher investigated and predicted the patient’s diagnosis in the machine teaching field to discover new knowledge or patterns in the health center data records and found required information that enabled imaginative ways of contributing to both patients and health center professionals [8].

Another research study [9] based on an analysis of “Heart Disease Prediction using Different Data Mining Techniques” presents 13 attributes using J48, Random forest, Random Tree, Multilayer Perceptron & Naive Bayes classifiers. The researcher also used, initially, a data set of 909 records with 13 attributes. The analysis shows that a Neural Network with 15 attributes has shown the highest accuracy i.e. 100%. On the other hand, the Decision Tree also performed well with 99.62% accuracy by using 15 attributes. Moreover, in combination with the Genetic Algorithm and 6 attributes, the Decision Tree has shown 99.2% efficiency. However, the data set used by the research is not enough to get such accurate results. The researcher also cannot explain the data processing and analysis tools rather than explain the result. So it is better to define the tool they used to develop the prediction model.

The study [10] designed a predictive model for heart disease detection using data mining techniques from the Transthoracic Echocardiography Report dataset that is capable of enhancing the reliability of heart disease diagnosis using echocardiography. The researcher collected data from PGI, Chandigarh from the year 2008 to 2011 containing 7,339 instances. The models were built on the preprocessed Transthoracic Echocardiography dataset with three different supervised machine learning algorithms i.e. J48 Classifier, Naïve Bayes classifier, and Multilayer Perceptron (ANN) using WEKA 3.6.4 machine learning software. The most effective model to predict patients with heart disease appears to be a J48 classifier implemented on selected attributes with a classification accuracy of 95.56%. Here the researcher only used a supervised machine learning algorithm, however, there is also unsupervised machine learning, and was ignored.

M.A.Nishara Banu et al. [11] revealed that MAFIA (Maximal Frequent Itemset algorithm) and K-Means clustering are better performing. As classification is important for the prediction of disease, classification based on MAFIA and K-Means produced results with higher accuracy [11]. Researchers [12] applied several machine learning algorithms. In this study, the researcher got the best accuracy on the ANN classifiers. However, the researcher used these two algorithms without including other algorithms and comparing them. The attributes used for this study are enough however the researcher couldn’t explain the relevant results. After applying ANN on preprocessed and training datasets, the results showed that there are zero FN or FP entries. The analysis results suggested that the system predicts heart disease with 100% accuracy. For this study, the researcher also used the WEKA 3.6 version tool.

3. RESEARCH METHODOLOGY & TOOLS

This research used an applied experimental research design. It mainly focuses on a quantitative data analysis approach. The research used the KDD (Knowledge Discovery in Databases) modeling approach. The researchers have selected the KDD-DM processing method. The steps of the KDD process are data understanding, data selection, preprocessing, transformation, data mining, evaluation, and interpretation of the mined data. This method is very important to preprocess the data to get better accuracy and to develop the desired model to overcome the problem observed in the specified domain [9].

3.1 Methodology Used and Data Analysis

The step-by-step design approaches of the proposed system and the work-flow of the complete system are as follows:

- Collection and selection of different heart disease data to train various machine learning algorithms.
- Comparison of various data mining algorithm’s accuracy and performance in predicting heart disease stage.
- Selecting the best algorithm by comparing the performance characteristics of the models to develop an intelligent heart disease stage prediction system.
- Inputting all parameters of heart disease manually and they can input sensor data like heartbeat using a specific button. Hence, they can predict whether the heart disease stage is major or minor.
- Depending on the result of the prediction system the doctor can easily treat the patients.

3.2 Data Preprocessing -Therefore, this step targets data preprocessing to obtain consistent and clean data. To achieve higher efficiency, improve accuracy get the highest value of data mining, and discover new knowledge; integrity, cleanness, and consistency must be maintained. During this step according to their data type, missing values are filled with either mean or mode values. Therefore, it is checked if there are any noises and redundancy happens. In this study there was no outlier in the numeric attributes and the data was cleaned from such problems.

3.3 Transformation - The data is cleaned; it is needed to transform or consolidate data into suitable forms for mining strategies and transfer into a data mining capable format using attributes such as attribute construction, aggregation, and discretization. The data source may be found in different formats and storage therefore this needs data integration which is detecting and deciding data value conflicts. To make the data more suitable for machine learning tasks encoding of some attributes is performed and discretization is applied to the attributes with continuous value.

3.4 Data Source and Collection Method-Data were collected from the patient’s dead file/ heart disease patient repository/record stored in the different hospitals of Ethiopia, Southern Nations, Nationalities, and Peoples' Region (SNNPR) such as Arba Minch and Ottona Hospital. In other words, some records were found in the dead file which was not organized in the computer system. Some of the data were collected from hardcopy or files of the patients.

3.5 Machine Learning- This phase was engaged in searching the patterns of interest in a particular representational form, depending on the ML techniques Classification, clustering, association, & prediction, out of this classification technique was selected. The goal of the study is to classify the Level of heart disease, into major and minor heart disease Stages.

3.6 Interpretation /Evaluation of the Discovered Knowledge-After discovering the required pattern or knowledge from the dataset; the interpretation and evaluation of the discovered patterns was done. The interpretation was concerned with whether the discovered pattern is interesting or not and verifies it is knowledge or not. Here also the evolution was done whether the prediction model was correctly working or not.

Comparatively measuring the performance of each classifier and representing the result in suitable ways. The performance of the classifiers adopted in the study is measured and ultimately, evaluated based on their accuracy, TP rate, recall, and precision. To compare the experiments performed using the selected algorithms, experiment results are generated from the preferred algorithm, and detecting interesting models and interpretations were approved by the domain experts.

3.8 Cross-Validation for Testing and Training the Model-In k-fold cross-validation, the initial data are randomly split into k equally exclusive subsets or folds, D1, D2, D3… Dk, where each fold was almost equal in size. In the first iteration depending on the fold size, training and testing were executed k times. Based on iteration I partitioned Di was kept as the test set, and the remaining partitions were jointly used to train the model. Then, the subsets D2… Dk is used as the training set to get a first model, which was tested on D1; and as well the second iteration is accomplished (trained) on subsets D1, D3… DK and tested on D2; [13]. Hence, for classification, the accuracy computed was the overall number of correct classifications from the k iterations, divided by the total number of tuples in the initial data. In general, stratified 10-fold cross-validation is recommended for estimating accuracy (even if computation power allows using more folds) due to its relatively low bias and variance.

3.9 Performance Measurement of the Model-A confusion matrix is one of the performance measurements used to compare and measure the quality of classification algorithms. A confusion matrix is an n-by-n table that contains information related to how many instances are correctly and incorrectly classified for each class [14]. A confusion matrix contains information about real and predicted classifications done by a classification system. The data in the matrix are evaluated to know the performance of such systems [8]. A confusion matrix contains information about real and predicted classification done by a classification.

3.11 Selection of Machine Learning Tool-To conduct the proposed research different software and designing tools were also selected using the best fit strategy. The researchers used open-source tools for data analysis, designing a predictive model and prototype, and validating the developed model. So, the researchers selected the following tools from different possible tools by using systematic tool selection criteria.

This research study critically compared and analyzed different types of Software Tools and technologies presented in Figure **3.3** Edraw-Max for Designing Conceptual Diagrams, WEKA for Visualization and feature selection, Data analysis & predictive modeling development, and C# and Python to develop a prototype for sample functional implementation of the proposed prediction system.

In this study, an effective machine learning technique and algorithm were chosen from some available algorithms and in a Java-based open access data mining platform (WEKA) to detect the stage of the disease to decide the probability of having major or minor heart diseases from a large dataset that developed from the real collected data from the real sources. Then a predictive and continuous monitoring system design was proposed using the C# & python programming language system [15]. The step-by-step design approaches of the proposed system and the work-flow of the complete system have been listed below:

- Collection and build heart disease datasets to train and test various machine learning algorithms.
- Comparison of various algorithm's accuracy and performance in predicting heart disease using the dataset.
- Selection of the best algorithm from the performance characteristics of the models to develop a heart disease prediction system that will support experts in health care.

3.11 The WEKA tool- this research WEKA 3.9.0 software which is developed at the University of Waikato in New Zealand was selected and used. This software is available at www.cs.waikato.ac.nz/ml/Weka site [9]. WEKA 3.9.0 machine learning software is selected because of the reasons; first, it is open source, which means that it can be obtained for free without depending on a Particular institution or company. Second, WEKA 3.9.0 has a category of prediction and classification features which is essential for this research. Third, it is fully implemented in Java and runs on almost any platform [16].

3.12 Methods of Training and Testing-The classifiers were evaluated by a cross-validation test model using the number of folds. K-fold is a natural number used to check the performance of the model through k–k-times. K-Fold is suitable whether the size of the data is very large or not. This is because of its general tests on various datasets with dissimilar learning schemes. In 10-fold cross-validation, the learning scheme or dataset was randomly reordered and then split into n folds of equal size. In each iteration, one fold is used for testing and the other n-1 folds will be used for training the classifier [8]. Finally, the overall folds gave the estimate of accuracy for cross-validation. Hence during the experimentation, the training and testing model was done using two testing models, 10-fold cross-validation and percentage split.

3.13 Methods of System Performance Evaluation & Analytics-After creating the model, comparing the predictive accuracy of the classifiers for unknown tuples is often useful to evaluate the performance of predictive modeling. This method can show how frequently instances of particular classes are correctly classified as actual classes or misclassified as some other classes. The performance of each applied classifier was comparatively identified corresponding to their accuracy using the testing model.

3.14 Validation-In order to prove the validity of machine learning applied the algorithm to the dataset, results were analyzed from 2 perspectives: statistical validity, considering the duplicability with completely different cohorts and correctness of statistical values obtained (i.e., metrics), and intra-validity, concerning the clinical and real implications of the algorithms daily. This is often a pair-wise co-existence; none of the machine learning cardiac classification algorithms is applied within the dataset if there's no agreement from both sides. The following subsections describe how the metrics and clinical effectiveness are thought of.

For the validity of the algorithms applied to the collected dataset, a whole heart condition patient's data set should be split into three different percentage splits, known as the training set, validation set, and testing set, severally. These groups are units typically selected in such a way that subgroups share demographic distributions like age or sex, to represent a real-world scenario. A balanced distribution of control and pathologic subjects is also required. Once the machine learning model is trained and tested, different metrics are obtained to evaluate its performance. Accuracy measures the percentage of the algorithm classifying the input data correctly. It is a simple measure used in multiple scientific scenarios if there's no class imbalance. One of the drawbacks of using accuracy because the metric is that there's an information loss once activity False Positive and False Negative observations. Therefore, Specificity (Sp) and Sensitivity (Se) are widely used for measuring the performance of the algorithm, now taking into thought a possible category imbalance. To assess the performance of an algorithm and to know wherever there could be a miss-classification issue, a table report called Confusion Matrix is used.

This specific table layout is typically used to describe the performance of a supervised learning model. Every row of the matrix represents the instance predicted class whereas every column represents the instances in an actual class (or vice versa). From sensitivity, specificity, and the confusion matrix we extracted a performance plot representation called the receiver operating (ROC) curve.

It was created by plotting the true positive rate (TP rate) against the false positive rate (FP rate) at various threshold settings. In machine learning, the true-positive rate is also called sensitivity, recall, or probability of detection. Roc analysis is related in a very direct and natural way to the cost/benefit analysis of diagnostic decision-making. The Area under the ROC curve (AUC) is another metric used to measure algorithms' performance [10].

4. RESEARCH DESIGN AND MODELING

4.1 Designing Model

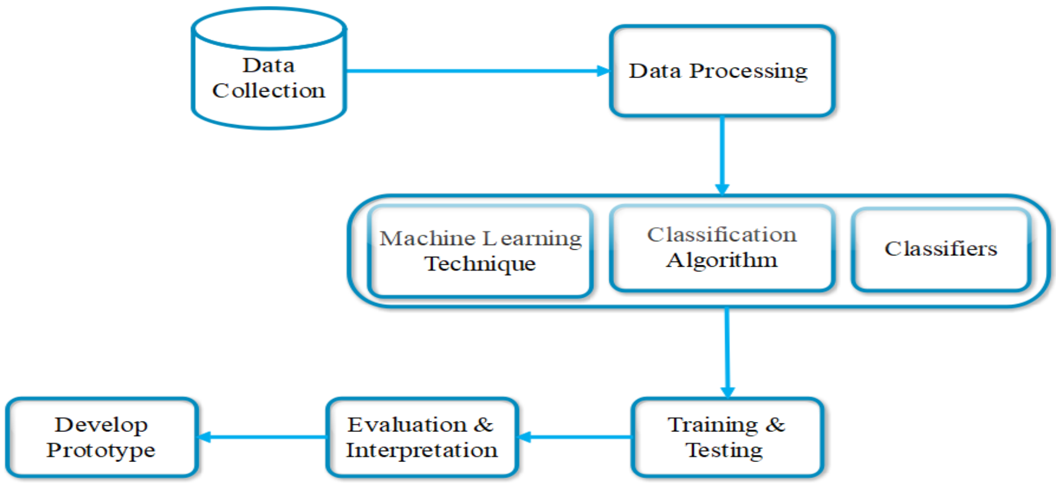

Figure 4.1: Basic Flow for Designing Model of Research Work

4.2 Data Understanding and Data - The aim of this study is also to understand the level (status) of heart disease patients to know whether he/she is in a serious heart disease problem or at the stage of a minor heart disease problem. To understand the domain problem, the data was collected from the patient repository (heart disease patient records) and the domain experts to measure the level of heart disease and improve the current diagnosis system. The researcher has used recorded files of patients from the year of 2012 to 2019 G.C. The data population size of heart disease patients' records was identified as a total of 5,050. However random data representation techniques were used to collect relevant data among this population size.

4.3 Data Preprocessing - To have a good classification and prediction, preprocessing is a very essential part of data mining and machine learning approaches. Here in this study, the main aim of preprocessing is that the source of data might contain incomplete information or data with noise data outliers, or missing values of the selected data.

4.4 How to handle missing values-

Missing data is a cell that remains vacant or which is without value so that it can affect the output of the model. Various options are available to handle missing values; the easiest one is to neglect those records with missing values. However, this might lead to a biased subset of the dataset. So, it needs other options to overcome the aforementioned problem [17]. Filling manually in the missing values is not possible if the dataset is huge. Filling all missing values with some global constant, this is also creating confusion during experimentation. Filling the missing values with a mean value which is numeric. Missing attribute values in the recorded data are most likely related to the absence of interesting information, lack of important data entry, and misunderstanding of data. The proneness of data for noise, inconsistency, and incompleteness is becoming dominant in the real world; this is mainly associated with its massive size and heterogeneous source of data [18]. Thus, it is substantial for the researcher to control missing values efficiently. If the missing values are not managed properly, handling missing values in the dataset to get quality data that importantly improves the overall quality of the patterns mined is very essential.

4.5 Data transformation- transformation in machine learning is essential and employed before applying different data mining techniques because the data are transformed or consolidated into forms appropriate for mining. Sex, Age, Chest Pain, Family History, Past History, Cholesterol, Fasting Blood Sugar, Resting ECG, Slope ST, Heart Rate, Pulse Rate, Blood Pressure, and CBC which are used to classify the heart disease patient into major or minor heart disease. The data set obtained from domain experts indicates that the heart disease patient's status was measured and considered into two categories called major or minor heart disease. Based on this category it is needed to group the patient status to be more appropriate for predictive analysis and relatively accurate for predicting heart disease. So, in this research, the patient's status is classified as major and minor heart disease based on the collected data and built dataset and the current status of the patient.

In data transformation, the data are consolidated into suitable forms for discovery (mining), so the strategies for data transformation involve the following: Smoothing, Attribute construction, aggregation, Normalization, and Discretization.

4.6 Data Format-The data was originally stored in an Excel worksheet to make it work in WEKA. In the Excel worksheet, missing values were replaced with the most frequent value if the data type is nominal and if the data type is numeric it is filled with the mean value. Numeric attributes also are merged into a certain group (nominal). The next step was to save the Excel file by the extension name .csv (comma-separated values) which is a format where commas are placed between values in adjacent columns. After all, the selected dataset was imported from MS Excel format to create a WEKA understandable format (.arff) for preprocessing and experimentation.

4.7 Selection of Modeling Techniques-The most powerful predictive modeling methods include decision trees, Neural Networks, Support Vector Machines, Gene Expression Programming, Symbolic Regression, K-means clustering, Linear Discriminant Analysis, Linear Regression models and Logistic Regression models. To accomplish this research, the researcher used four machine-learning classification algorithms. To conduct the four experiments decision trees (J48 and Random Forest) algorithm, Bayes (NaiveBayes) algorithm, and function (multilayer perceptron) ANN algorithm were used.

During the conduct of experiments; many classifiers were applied but, because of their high accuracy, precision, and recall the three classifiers mentioned above were selected. Also, they can handle both continuous and categorical variables that can perform classification as well as regression. It automatically handles interactions between variables and identifies important variables. The WEKA machine learning tool was integrated and implemented using J48, Random Forest, Naive Bayes, and ANN (Multilayer perceptron) [12].

4.8 Selecting and Evaluating the Attributes- To do this researchers have used, the CfsSubsetEval method. This method is used to assign a value to each subset of attributes by searching Best First style in the WEKA tool. During experimentation, all the default parameters that were already set in the WEKA tool were used and percentage split 50%, 66% & 80 % were used for training tests and building models. However, a 66% percentage split is selected for experimentation due to more accuracy value and model performance than the other percentage splits.

4.9 Visualized Data-In machine learning, visualization is one of the techniques used/adopted for viewing data in graphical means, and to communicate data to identify relations and biases hidden in unstructured data sets through graphical representation. Showing collected data can lead to better understanding and models to provide statistical or other means of analysis hence, in this research, the following graph shows the graphical representation of reprocessed data. This data is ready for experimentation.

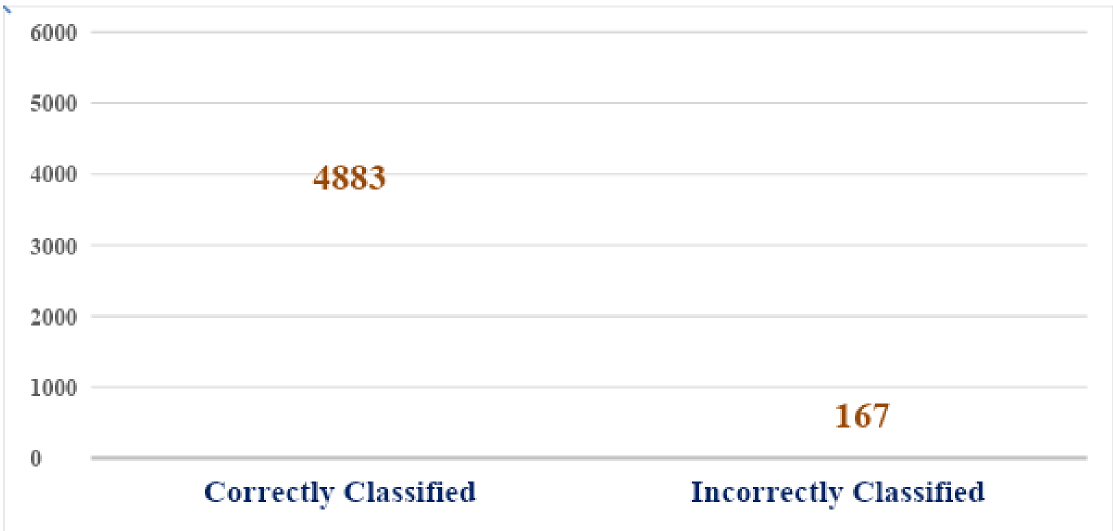

Figure 4.2 Visualizing preprocessed data with their instances

4.10 Experiment Set Up- The experiment of the study followed the setup of the following experiment such as:

- The outcome of test validation of 10-fold cross-validation and percentage split for each experiment.
- The accuracy of each experiment (experiments 1, 2, 3, and 4) including its TP rate, Recall, and Precision result.

- Determining the effect of classification by using different classifiers with selected attributes.
- Using best first attribute selection methods to select determinant attributes.

4.11 Experiment -The experimentation is performed with a Decision tree, Bayes, and Function machine learning classification algorithms. In this experiment, classifier algorithms like ANN (Multilayer Perceptron), J48, Random Forest, and Naive Bayes were applied. The experiment was designed to evaluate the performance of all the selected classifiers for building model, testing, and training data sets. The experimentation was done with the classifier's default parameters using the test mode. The default 10-fold cross-validation and different percentage split test options are conducted for training and testing the classification model. For the percentage split test option, 5,050 total datasets were split into 50%, 66%, and 80% were used for the training set and the remaining is for the test set. Among the selected percentage split 66% is more accurate than the other percentage split, and the best classifier is ANN.

Experiment I

4.11.1 Classification Building and Testing Model Using J48 Algorithm- To get the best performance of the model, the research was conducted using different cross-validation values on both trainings. So as the k value increases the accuracy value increases and ultimately, after it reaches 10-fold, it decreases slightly. Hence the 10-fold cross-validation test option was used due to being relatively more accurate and suitable for testing. By experimenting, the model was generated with 8 numbers of leaves the size of the tree 15, and the time taken to build the model: 0.11 seconds. After stratified cross-validation (pruned), the correctly classified instance was 4883 with an accuracy of 96.6931 % and the incorrectly classified instance was 167 with 3.3069 % incorrectly classified as well. The following table 4.4 shows the results for the accuracy of the applied J48 classifiers.

Table 4. 4 Experimental Results for J48 Algorithm

Classifier	Test options	No of leaves	Size of the tree	Correctly classified	Incorrectly classified
J48	Cross-validation 10-fold	8	15	4883(96.6931 %)	167(3.3069 %)
	Percentage split 50%	8	15	2444(96.7921 %)	81(3.2079 %)
	Percentage split 66%	8	15	1660(96.6803 %)	57(3.3197 %)
	Percentage split 80%	8	15	973(96.3366 %)	37(3.6634 %)

As it is summarized in Table 4.4 the classifier used a percentage split of 50%, 66%, and 80%, and the tree used all attributes. Hence the 50% percentage split was comparatively selected to test the model and measure the classification accuracy which was correctly classified at 96.7 % with 2444 trained instances and generally the result of this percentage split was more accurate than the other percentage split.

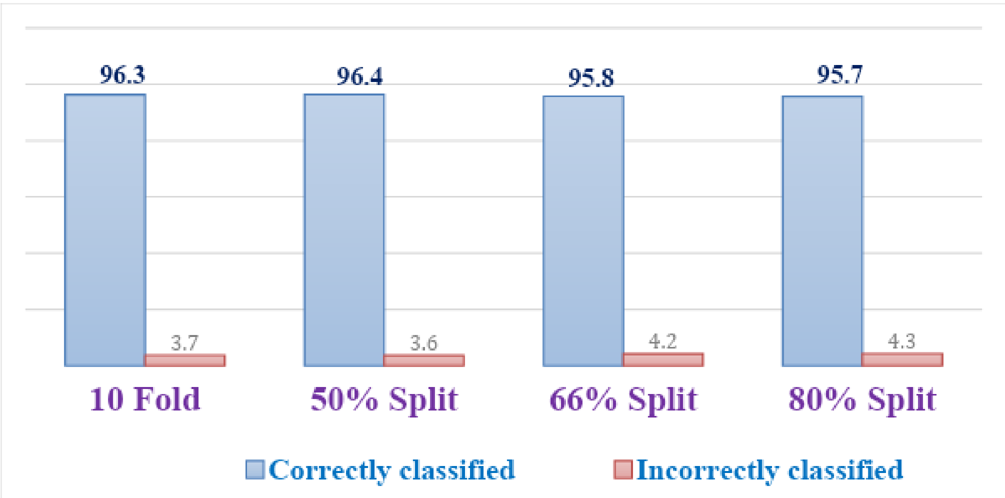

Figure 4. 5 Experiment Results using the J48 Algorithm

Table 4. 5 Experimental Results for J48 Algorithm using Detailed Accuracy by Class

Actual	10-Fold			Percentage Split								
				50%			66%			80%		
	TP rate	Precision	ROC Area	TP rate	Precision	ROC Area	TP rate	Precision	ROC Area	TP rate	Precision	ROC Area
Major Heart Disease	0.964	1.00	0.984	0.965	1.00	0.984	0.964	1.00	0.985	0.960	1.00	0.982
Minor Heart Disease	1.00	0.679	0.984	1.00	0.709	0.984	1.00	0.712	0.985	1.000	0.694	0.982
Weighted Average	0.967	0.978	0.984	0.968	0.977	0.984	0.967	0.976	0.985	0.963	0.975	0.982

Experiment II

4.11.2 Classification Building and Testing Model Using ANN Algorithm

Table 4. 6 Experimental Results for ANN Algorithm

Classifier	Test options	Correctly classified	Incorrectly classified
ANN	Cross-validation 10-fold	4885(97.0%)	165(3.0%)
	Percentage split 50%	2417(95.7%)	108(4.3%)
	Percentage split 66%	1658(96.6%)	59(3.4%)
	Percentage split 80%	973(96.3%)	37(3.7%)

As it is summarized in Table 4.6 the classifier used a percentage split of 50%, 66%, and 80%, and the tree used all attributes. Hence the 66% percentage split was comparatively selected to test the model and measure the classification accuracy which was correctly classified 96.6% with 1658 trained instances and generally the result of this percentage split was more accurate than the other percentage split.

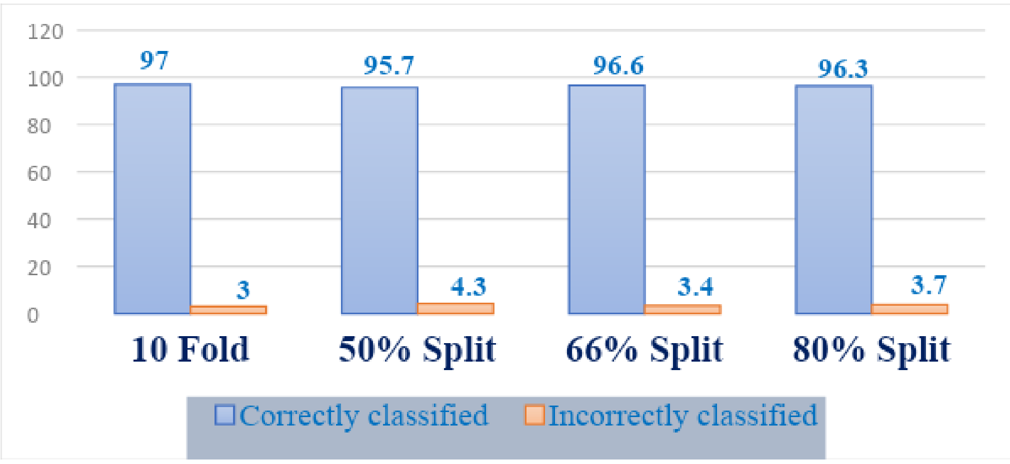

Figure 4. 6 Experiment Results using ANN Algorithm

Table 4. 7 Experimental Results for ANN Algorithm using Detailed Accuracy by Class

Actual	10-Fold Cross Validation			Percentage Split								
				50%			66%			80%		
	TP rate	Precision	ROC Area	TP rate	Precision	ROC Area	TP rate	Precision	ROC Area	TP rate	Precision	ROC Area
Major Heart Disease	0.966	0.999	0.988	0.979	0.974	0.987	0.964	0.999	0.986	0.960	1.000	0.985
Minor Heart Disease	0.986	0.685	0.988	0.695	0.741	0.987	0.986	0.709	0.986	1.000	0.694	0.985
Weighted Average	0.967	0.977	0.988	0.957	0.956	0.987	0.966	0.975	0.986	0.963	0.975	0.985

Experiment III

4.11.3 Classification Building and Testing Model Using Random Forest Algorithm

Table 4. 8 Experimental Results for Random Forest Algorithm

Classifier	Test Options	Correctly Classified	Incorrectly Classified
Random Forest	Cross-validation 10-fold	4864(96.3%)	186(3.7%)
	Percentage split 50%	2434(96.4%)	91(3.6%)
	Percentage split 66%	1645(95.8%)	72(4.2%)
	Percentage split 80%	967(95.7%)	43(4.3%)

As it is summarized in Table 4.8 the classifier used a percentage split of 50%, 66%, and 80%, and the tree used all attributes. Hence the 50% percentage split was comparatively selected to test the model and measure the classification accuracy which was correctly classified at 96.4% with 2434 trained instances and generally the result of this percentage split was more accurate than the other percentage split.

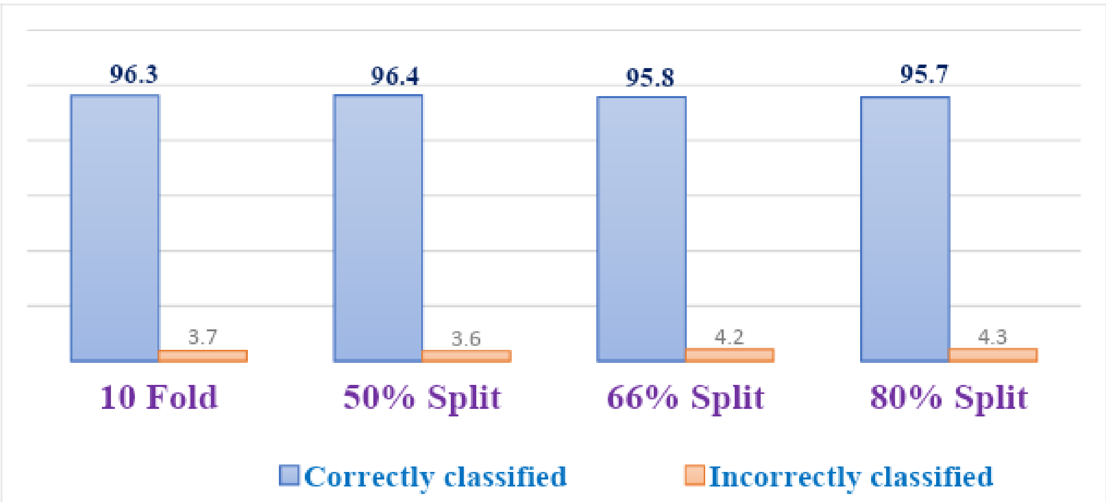

Figure 4. 7 Experiment Results using Random Forest Algorithm

Table 4. 9 Experimental Results for Random Forest Algorithm using Detailed Accuracy by Class

Actual	10-Fold			Percentage Split								
				50%			66%			80%		
	TP rate	Precision	ROC Area	TP rate	Precision	ROC Area	TP rate	Precision	ROC Area	TP rate	Precision	ROC Area
Major Heart Disease	0.969	0.992	0.988	0.976	0.984	0.988	0.975	0.980	0.986	0.969	0.985	0.985
Minor Heart Disease	0.890	0.681	0.988	0.817	0.745	0.988	0.773	0.732	0.986	0.833	0.707	0.985
Weighted Average	0.963	0.970	0.988	0.964	0.966	0.988	0.958	0.959	0.986	0.957	0.962	0.985

Experiment IV

4.11.4 Classification Building and Testing Model Using Naïve Bayes Algorithm

Table 4. 10 Experimental Results for Naïve Bayes Algorithm

Classifier	Test Options	Total Number of Instances	Correctly Classified	Incorrectly Classified
Naïve Bayes	Cross-validation 10-Fold	5050	4761(94.3%)	289(5.7%)
	Percentage split 50%	2525	2415(95.6%)	110(4.4%)
	Percentage split 66%	1717	1628(94.8%)	89(5.2%)
	Percentage split 80%	1010	953(94.4%)	57(5.6%)

As it is summarized in table 4.10 the classifier used a percentage split of 50%, 66%, and 80%, and the tree used all attributes. Hence the 50% percentage split was comparatively selected to test the model and measure the classification accuracy which was correctly classified at 95.6% with 2415 trained instances and generally the result of this percentage split was more accurate than the other percentage split.

Figure 4. 8 Experiment Result using Naïve Bayes Algorithm

Table 4. 11 Experimental Results for Naïve Bayes Algorithm using Detailed Accuracy by Class

Actual	10-Fold Cross Validation			Percentage Split								
				50%			66%			80%		
	TP rate	Precision	ROC Area	TP rate	Precision	ROC Area	TP rate	Precision	ROC Area	TP rate	Precision	ROC Area
Major Heart Disease	0.943	0.995	0.970	0.958	0.995	0.972	0.948	0.995	0.972	0.943	0.995	0.969
Minor Heart Disease	0.938	0.554	0.970	0.939	0.654	0.972	0.950	0.620	0.972	0.952	0.602	0.969
Weighted Average	0.943	0.964	0.970	0.956	0.968	0.972	0.948	0.965	0.972	0.944	0.963	0.969

5. RESULT AND DISCUSSION

5.1 Performance Evaluation and Interpretation

5.2 Experimental Results

The result of those selected algorithms is analyzed and discussed which showed an accuracy level of more than 90%. The algorithms are applied to the data set by using 10-fold Cross-Validation on WEKA version 3.9.3 for measuring the performance of the machine learning algorithms. The following measures are calculated for performance analysis of the selected algorithms:

Table 5. 1 Experimental Results

Experiment	Classifier/Algorithm	Test Option	Precision	Recall	Accuracy
Ex-1	ANN	10-Fold	97.7	96.7	97.0
		PS	97.5	96.6	96.6
Ex-2	J48	10-Fold	97.8	96.7	96.69
		PS	97.7	96.7	96.6
Ex-3	NaiveBayes	10-FoldSsS	96.4	94.3	94.3
		PS	96.8	95.6	95.6
Ex-4	RandomForest	10-Fold	97.0	96.3	96.3
		PS	96.6	96.4	96.4

To compare the best model; precision, recall, and accuracy (correctly classified instances) of the classes were used for different classifiers based on their test model of 10 FCV. The following Figure 5.1 shows how to compare and contrast each classifier based on their precision, recall, and accuracy.

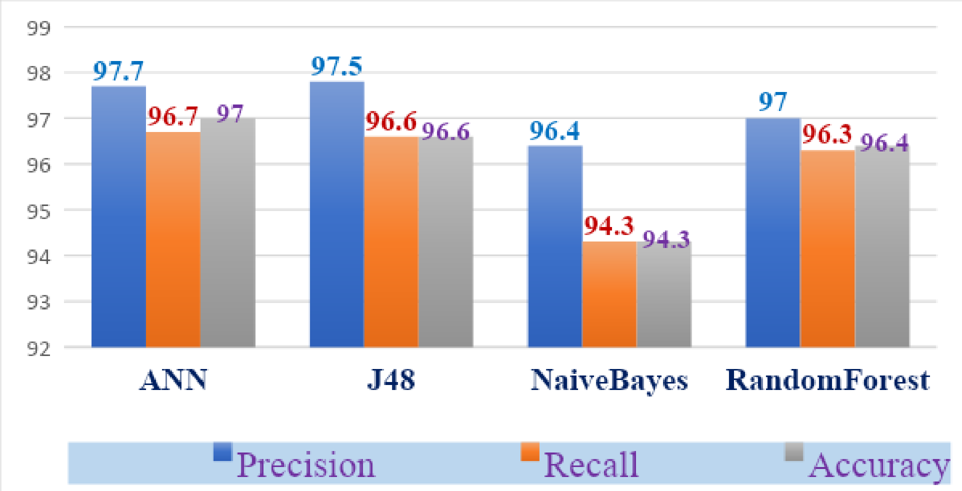

Figure 5. 1 10-Fold Comparison of the Classifiers

As you see in above figure 5.1 the cross-validation in each algorithm is stated the in figure. The algorithm ANN cross-validation is showing more accuracy than other algorithms.

To compare the best model; precision, recall, and accuracy (correctly classified instances) of the classes were used for different classifiers based on their test model of percentage split (50%).

5.3 Performance Comparison Training Test of Applied Algorithms

The results for the performance comparison of the obtained classification algorithms for the experimentation were summarized and presented in Fig. 5.2

Table 5. 2 Performance Comparison

Experiment	Classifier/ Algorithm	Time Taken to Build Model	Correctly Classified	Incorrectly Classified
Ex-I	ANN	14.45 Second	97.0%	3.0%
Ex-II	Naive Bayes	0.24 Second	94.3%	5.7%
Ex-III	J48	0.03 Second	96.69%	3.31%
Ex-IV	RandomForest	0.31Second	96.3%	3.7%

5.4 Proposed Prediction System Architecture

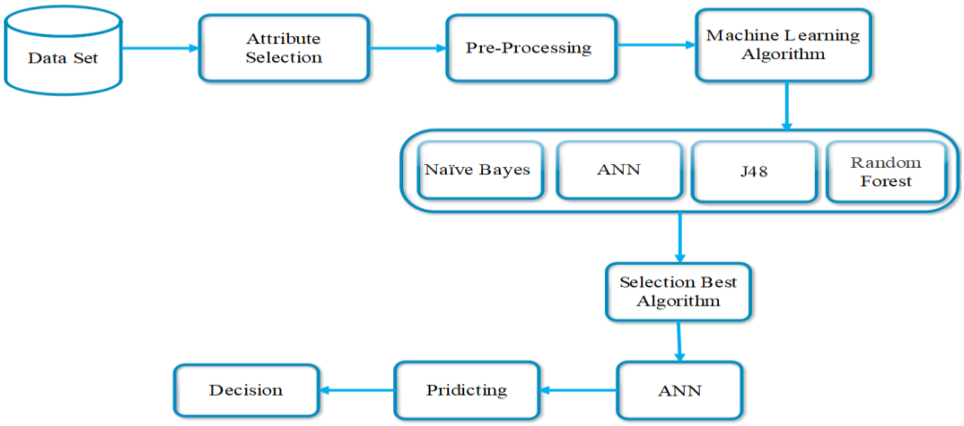

Figure 5. 3 System we suggest for the problem

5.2 Prototypical Representation of the Predictive Model

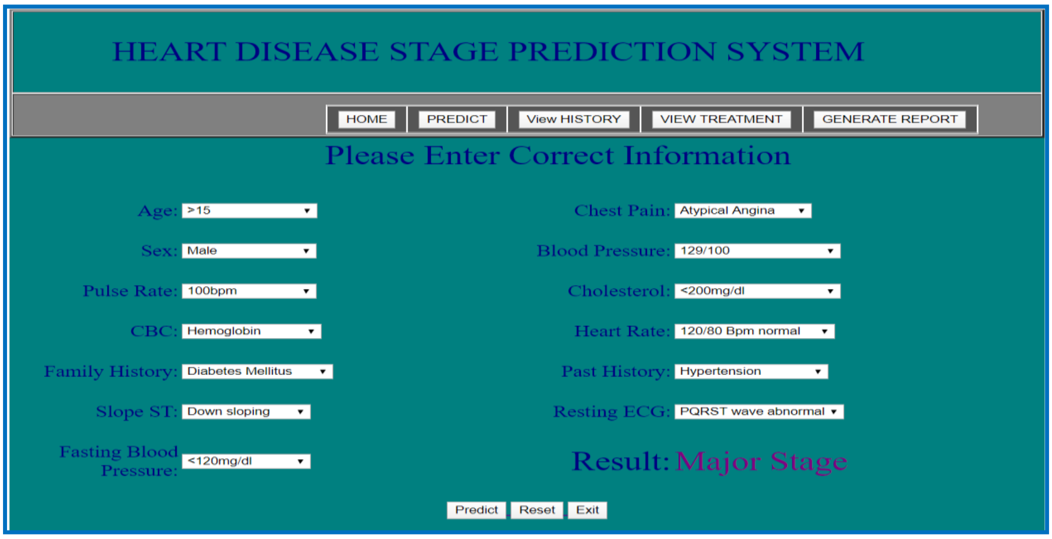

Figure 5. 5 Screenshot of Major Stage of Heart Disease

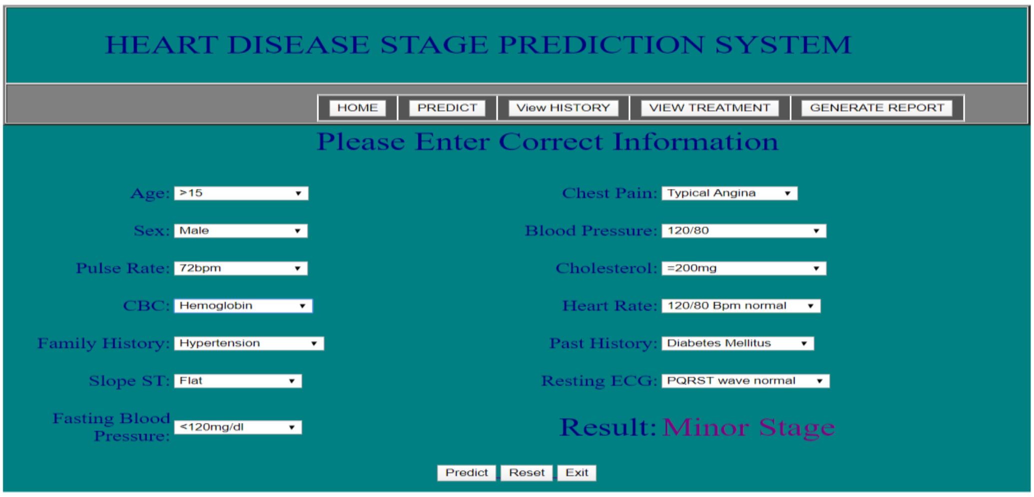

Figure 5. 6 Screenshot of Minor Stage of Heart Disease

6. CONCLUSION AND RECOMMENDATION

6.1 Conclusion

In this research, a machine learning model was trained to determine whether a patient has a major or minor heart disease. This machine learning model used a dataset with 14 selected attributes to conduct experiments using different classifiers. After experimentation, the prediction system model proved that it is better than traditional diagnosis. Also, it can identify the stage of the heart disease in the very simplest manner of the patient through machine learning algorithms. This study applied four machine learning algorithms to the dataset and out of those classifiers/algorithms the accuracy of ANN (97.0%) is found better than other algorithms. Based on this model it can be concluded that the proposed model can identify the early stage occurrence or appearance of the disease. This can help doctors to provide early-stage treatments to patients for better and more advanced care.

6.2 Recommendation

In this research, efforts have been made to apply machine learning techniques to predict the stage of heart disease. Thus, based on the results of the research the following recommendations are forwarded as they are important issues for further research directions in predictive modeling and analytics of heart disease.

In this research work, an attempt has been made to do the analytic and predictive modeling of heart disease stages by using some set of parameters/attributes. However as there are different types of parameters/attributes in heart disease, it is better to use some additional attributes to get better accuracy. For data sampling, the researcher has used 5050 datasets that were collected from different hospitals. However, if the data size is increased and data is collected from different data sources, the research could be able to provide some new patterns/ ideas and better heart disease stage predictive accuracy.

Even though there are many machine learning techniques, this research used only the classification technique with three classifications (Decision tree, Bayes, and Function) which is integrated into the WEKA 3.9 tool. But, other machine learning methods like rule-based, and vector machines were not considered for testing and predicting in this research. It might be important to predict the stage of heart disease using different algorithms and even to understand better classifiers while experimenting. For this study, the researcher has used the WEKA 3.9 tool, but there are many other machine learning tools like MatLab and Rapid Miner, which may be more sophisticated to develop, model, and analyze the data.

REFERENCES

[1] A.A.D.P. Sharma, “Towards Smart and Green Features of Cloud Computing in Healthcare Services: A Systematic Literature Review.” J Inf Syst Eng Bus Intell 2023;9:161–80. https://doi.org/10.20473/jisebi.9.2.161-180.

[2] S.H.M.S. D.P. Alade IO, “GBR-GSO based machine learning predictive model for estimating density of Al2N3, Si3N4, and TiN nanoparticles suspended in ethylene glycol nanofluids”. Eur Phys J Plus. 2022;137:587.

[3] A.A.D. P. Sharma, “Enhancing Healthcare Information Systems in Ethiopian Hospitals: Exploring Challenges and Prospects of a Cloud-based Model for Smart and Sustainable Information Services,” Int. J. Inf. Technol. Comput. Sci., vol. 16, no. 5, pp. 1–22, Oct. 2024, doi: 10.5815/ijitcs.2024.05.01.

[4] S. B.-D. S. Shalev-Shwartz, "Understanding machine learning From Theory to Algorithms," Cambridge University Press., 2014.

[5] T. M. S. H. H. K. Rajesh N1, " "Prediction of Heart Disease Using Machine Learning Algorithms," International Journal of Engineering & Technology, vol. vol. 7, 2018.

[6] 1. S. S. a. G. S. P.-J. Poornima Singh, " "Effective heart disease prediction system using data mining techniques," Int J Nanomedicine, 2018.

[7] P. S. E. F. D.-O. O.E. Taylor1*, "A Model to Detect Heart Disease using Data Mining Algorithm," International Journal of Computer Sciences and Engineering Open Access, Vols. Vols. Vol.-7, no. no. Issue-11, ,, Nov 2019.

[8] V. M. C. C. I. Z. R.-E. B. B. S. E. P. a. K. L. C. Martin-Isla1, "Image-Based Cardiac Diagnosis With Machine Learning," Fronteries, 24 January 2020.

[9] R. H. A. 1. a. M. K. S. Ma’aitah, "Deep Convolutional Neural Networks for Chest Diseases Detection," Journal of Healthcare Engineering, 2018.

[10] D. K. A. S. G. Tripti Mishra, "Students’ EmployabilityPrediction Model through Data Mining," International Journal of Applied Engineering Research, Vols. vol. 11, pp. 75-83, 2016.

[11] M. A. N. B. a. B. Gomathy, "Disease Forecasting System Using Data Mining Methods," International Conference on Intelligent Computing Applications (ICICA), 2014.

[12] M. R. R. M. R. I. M. H. I. S. Nashif, "Heart Disease Detection by Using Machine Learning Algorithms and a Real-Time Cardiovascular Health Monitoring System," World Journal of Engineering and Technology, vol. 6,

2018.

[13] D. D. D. Dr. David L. Olson, "Advanced Data Mining Techniques," USA: Springer, p. 2008.

[14] M. A. Imran, "n Depth Analysis of Lung Disease Prediction Using Machine Learning Algorithms," no. "I,", June 2020. .

[15] T. M. A. M. J. T. Saeide Kakavand, "Prediction the Loyal Student Using Decision Tree Algorithms," International Journal of Information and Communication Technology Research, vol. 4, 2014.

[16] D. K. D. J. Rub, "Predicting the Performance of Students in Higher Education Using Data Mining Classification Algorithms," International Journal for Research in Applied Science & Engineering Technology (IJRASET), vol. 2, no. 11, p. 17.

[17] P. R. B. T. Ms.Jyoti Chandrashekhar Bambal1, "Designing a Disease Prediction Model using Machine Learning," IOSR Journal, 2015.

[18] X. Zhu, "A novel data mining process for industry practice based on agile methods and visualization," AGILE MINING, no. , November 2017.