

PREDICTIVE ANALYSIS OF HEART DISEASE USING SELECTED MACHINE LEARNING META - ALGORITHMS

Bamanga, Mahmud Ahmad

Computer Science Department, Federal University of Lafia, Nigeria.
(mabamanga@gmail.com)

Ahmadu, Asabe Sandra

Computer Science Department, Modibbo Adama University, Yola, Nigeria.
(ahmaduasabe@mautech.edu.ng)

Musa, Yusuf Malgwi

Computer Science Department, Modibbo Adama University, Yola, Nigeria. (yumalgwi@mautech.edu.ng)

Kamalu Aliyu Babando

Computer Science Department, Taraba State Polytechnic, Jaling, Nigeria. (babando75@yahoo.com)

Abstract

Machine learning techniques are commonly used in clinical decision support systems for the identification and prediction of various diseases. Since heart disease is the leading cause of death for both men and women around the world. Because the heart is such an important part of the human body, it is one of the most pressing medical concerns. To improve the ability to diagnose and predict heart disease in humans, several researchers have developed intelligent medical decision support systems. However, there are few studies that look at the capabilities of ensemble methods in developing a heart disease detection and prediction model. In this study, the researcher look at how to use ensemble model, which proposes a more stable performance than the use of base learning algorithm and these leads to better results than other heart disease prediction models. The University of California, Irvine (UCI) Machine Learning Repository archive was used to extract patient heart disease data records. To achieve the aim of this study, the researcher aggregate Naïve Bayes, Support Vector Machine and Decision Tree with Adaboost and Bagging. The ensemble model is a superior solution in terms of high predictive accuracy and diagnostics output reliability, according to the results of the experiments. An ensemble heart disease prediction model is also presented in this work as a valuable, cost-effective, and timely predictive option with a user-friendly graphical user interface that is scalable and expandable. From the finding, the researcher suggests that Bagging is the best ensemble classifier to be adopted as the extended algorithm that has the highest prediction accuracy of 91% compared to adaboost with a prediction accuracy of 88% in the implementation of heart disease prediction.

1 INTRODUCTION

In modern healthcare information system, detecting and predicting solution model for heart diseases is extremely challenging. Any clinical attempt made through modern computing to provide economical solutions and options is worth taking to medical experts. Therefore, the need for computerized models capable of performing various tasks such as data analysis, predictive and detective analysis using effective data mining tools is vital. The application of such tools can assist medical practitioners to make prompt and right medical decision. A lot of computation applications are designed purposely to assist physicians in making effective decision about the various diseases in patients. Despite distinct soft computing innovative effort in place, predicting and detecting heart diseases models stills remained an issue of grave medical concern. Data mining techniques provide lot of algorithm to detect and predict disease but the challenges associated with each algorithm is still the matter that needs proper consideration (Sarwar, Sharma and Gupta, 2015). Moreover, artificial intelligence plays a key role in every aspect of modern classified intelligence. In medical perspective, artificial intelligence assist medical doctors and other health practitioners in performing expert diagnosis as well as detection or prediction of diseases in a patient more efficiently and accurately. Besides, artificial intelligence algorithms have great potential for exploring the hidden patterns in the datasets of the various disease related subjects by adjusting the data mining model for utilizing such patterns for clinical diagnosis (Nikookar and Naderi, 2018).

This research work is proposes to developed a predictive analysis of heart disease using machine learning meta-algorithms. The sample of patient database needed to carry out the detection and prediction of heart using common machine learning algorithms. This research will demonstrates the capability of computerizing three popular classifiers which are Support Vector Machine (SVM), Random Tree and Naïve Bayes of the collective dataset captured in the application. The model would also use Adaboost and Bagging as extended ensemble models or fuser's classifiers to finally perform decisive detection and prediction of heart diseases.

2 RESEARCH MOTIVATION

The fact that human life is dependent on the proper functioning of the heart is the driving force behind this research. The heart is a crucial part of our bodies, and heart disease has become the leading cause of death globally. In most nations around the world, heart disease is still the leading cause of mortality. Early detection and treatment of heart disease can pave the way for excellent heart health for the rest of one's life. In the medical field, artificial intelligence aids medical doctors and other health practitioners in more efficiently and reliably completing expert diagnosis, as well as detection or prediction of diseases in patients.

Furthermore, the goal of this research is to strengthen the research of scholars who have made significant contributions to the subject of study.

3 STATEMENT OF THE PROBLEM

The World Health Organization estimates that 12 million people die each year as a result of heart disease (Soni, Ansari, Sharma and Soni, 2011). The unfortunate loss of lives continued to rise, necessitating the need for medical professionals to devise adaptable remedies to the epidemic. A traditional heart disease prediction approach is daunting, especially in many underdeveloped countries where doctors are scarce and data is extensive (Chadha and Shubhankar, 2016). Data mining useful methods for identifying and detecting diseases in patients are incorporated into machine learning procedures.

Our research intends to address the need for a better method to managing detection and prediction of cardiac illnesses using various techniques to create specific detection outputs.

4 AIM AND OBJECTIVES

This research work is aimed at developing ensemble model for predicting heart disease using different machine learning meta-algorithm. The specific objectives are:

- i. To define the dataset to be used in the study
- ii. To design a model for heart disease predictive analysis using python
- iii. To perform evaluation of selected classifiers on the datasets using Support Vector Machine, Naïve Bayes and Decision Tree with Adaboost and Bootstrap Aggregating
- iv. To validate the proposed model using K-Fold Cross-Validation.
- v. To compare the performance of the selected extended classifiers on their predictive accuracy.

5 SIGNIFICANCE OF THE STUDY

- i. The development of heart disease prediction model will improve in modern healthcare challenges.
- ii. The finding of this research will be of immense benefit to the physicians and other medical practitioners and organization toward renewing the effort made in the process of detecting and prediction heart diseases as the application would lay the basic ensemble models to use.
- iii. The adaption of this ICT innovation will guarantee medical expertise compliance and support for smart heart diseases decision system without much related challenges.
- iv. The implementation of this model will help the patient to begin early heart disease treatment.

6 Summary of Related Work

S/N	AUTHOR	TITLE	METHOD USED	PROBLEMS IDENTIFIED	SOLUTION PROFFERED
1.	Chen, Huang, Hong, Cheng, and Lin, (2011)	Heart Diseases Predicting System	C++ AND C #	The research applied clinical raw data and when tested, The accuracy of the proposed method for prediction is near to 80%	The research used only one classification algorithm (i.e. neural network)
2.	Manpreet, Martins, Joanis, and Mago, (2016)	Building a Cardiovascular Disease Predictive Model using Structural Equation Model & Fuzzy Cognitive Map	Simulation using JAVA	The using Structure Equation Simulation and achieved 80% of tested ingredient using Fuzzy Cognitive Map	The dataset was mainly used for training rather than the real implementation
3.	Hannan, Alane, March, Salvatore, Park and Ram, (2010).	Prediction of Heart Disease Medical Prescription using Radial Basis Function	MATLAB	Obtained results show that radial basis function can be successfully used (with an accuracy of 90 to 97%) for prescribing the medicines for heart disease	Data set was tested once using Radial Basis Function without others functions.
4.	Mujawar and Devale (2015)	Prediction of Heart Disease using Modified k-means and by using	MATLAB	It finally gives a suitable number of clusters. Naive Bayes's creates	Highly depend on structured dataset.

		Naive Bayes		a model with predictive capabilities. This predictor has 93 % accuracy in predicting a heart disease and 89% accuracy in cases where it detected that a patient doesn't have a heart disease	
5.	Prajakta, Vrushali, Kajal, and Prajakta, (2016).	Intelligent Heart Attack Prediction System Using Big Data	Hadoop and Apache Mahout	Provide prototype intelligent heart attack prediction system that uses big data and data mining modelling techniques	Depend on big data and lacks sophisticated predictive models and risk calculation for a given data set
6.	Sudhakar and Manimekalai (2014).	Study of Heart Disease Prediction using Data Mining	Comparative Tool	Performs analyses and comparison between the works of different classification algorithms on a heart disease database.	Highly depend on rich information or data set
7.	Kamal and Kanwal (2014),	Review of Heart Disease Prediction using Data Mining Classifications	Apache Mahout	Discovered that Naïve Bayes appears to be the most effective model for disease prediction followed by neural network and Decision trees.	Tested only three classifiers at a time without other classifiers in place.
8.	Dhanashree, Mayur, and Shruti, (2013)	Heart Disease Prediction System Using	MATLAB	Proves that the accuracy has been	The research used only one

		Naive Bayes		obtained by altering the number of occasions within the given dataset normally using coaching and testing phases.	classification algorithm (i.e. Naiyes Bayes)
9.	Das, Ibrahim, and Abdulkadir, (2009).	Effective diagnosis of heart disease through neural networks ensembles	MATLAB	the diagnosis of heart disease have achieved 89.01% accuracy, and 80.95% sensitivity,	No centralized means of gathering the data set used.
10.	Samuel , Grace, Arun, Peng, and Guanglin, (2017)	An integrated decision support system based on ANN and fuzzy_AHP for heart failure risk prediction	Expert System	For detection of heart diseases problem, the hybrid model achieved 92.31% accuracy	The novel hybrid model only detect but have no capability of prediction

Table 1.1: Summary of Related Work (Field Work, 2021)

7 Source of Data and Data Collection

The sources of data for this study will be medical records of patients diagnosed with heart disease problem from Specialist Hospital, Yola Adamawa State. To generalize it all, the researcher will also use data from UCI Machine Learning repository.

8 System Design

System design encompasses the processes of designing and specifying the architecture of the system, system modules, user interface and data needed to satisfy the system requirement (Alan, Barbara and David, 2015). To design a system, the functional parts need to be evaluated at each phase or development. Therefore, we specify our data and users at every phase of development. We break down the system into application architecture, data flow diagram (DFD), the application interface, database model and the code. The development tools used in this research are presented in table 3.1.

Development	Programming Languages	Packages / Libraries	IDEs	Table 1.2: Tools
	Python,	Numpy, Pandas, Scikitlearn, Matplotlib, Flask,	Jupyter Lab	

9 Proposed Ensemble Model for Heart Disease prediction

9.1 70% Trained Data

This is the validated data the researcher used to trained and fit the model we created to predict heart disease.

9.2 30% Test Data

This is the ratio that the researcher utilized to put heart disease prediction model to the test.

9.3 ML Base Algorithms

The base learning algorithms utilized in machine learning approaches for the prediction and analysis of a complicated problem include the Nave Bayes, Decision Tree, and Support Vector Machine in which the researcher employed.

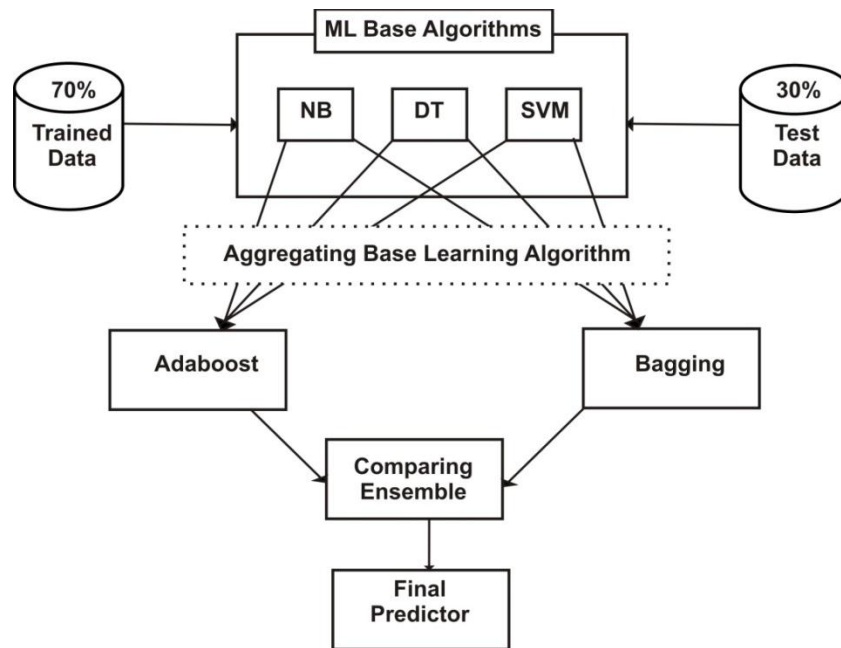


Figure 1.1: Proposed Ensemble Model for Heart Disease prediction

9.4 Aggregating Base Learning Algorithms

In this phase, the researcher combines the three base learning algorithms into a single powerful ensemble classifier in order to improve the model's prediction outcome.

9.5 Comparison of Adaboost and Bagging

During this phase, the researcher compared the ensemble classifiers in order to choose one with a superior prediction capacity.

9.6 Final Predictor

One of the extended classifiers with higher performance metrics was chosen by the researcher to be employed as the classifier during the diagnosis of patient data.

10 RESULTS

In this research, the researcher adopt Naïve Bayes, Decision Tree and Support Vector Machine as base learning algorithms in machine learning technology which contains a set of outputs and combines them into a strong and to enhanced prediction of the model by comparing the two extended classifiers or meta-algorithm.

The implementations of this ensemble model were applied on a dataset from UCI Machine Learning Repository for the training and testing of our model for prediction of heart disease. The data collected contains 303 instances of patients diagnosed of heart disease and each clinical instance of patience consists of 14 raw attributes.

11 PERFORMANCE EVALUATION OF THE MODEL

This is the crucial stage of the machine learning process. The researcher consider Accuracy, Precision, Sensitivity (Recall), and F1_Score as the measures use in evaluating the model considering its confusion matrix.

11.1 Confusion Matrix

Confusion Matrix is $N \times N$ matrix that is used in evaluating the outcome of our model, with N denoting the number of target. The matrix compares the values of the real target to our models' which provide us with a comprehensive result of how well our model is performing.

Table 1.2: Confusion Matrix for Binary Classification

	Predicted Positive	Predicted Negative
Actual Positive	TP	TN
Actual Negative	FP	FN

Two prediction scenarios exist as shows in table 4.1. A prediction is positive (meaning patients diagnosed with heart disease) and negative prediction, meaning patients diagnosed without heart disease.

Hence, from the above matrix, TP value will be the patients diagnosed with the disease while FN values were the patients diagnosed without the disease.

In view of the above, the researcher highlights the following terminologies:

- i. **True Positive (TP):** Is the scenario in which a case is predicted positive and they are actually positive.
- ii. **True Negative (TN):** A situation where the models were predicted negative and they are actually positive.
- iii. **False Positive (FP):** Patient predicted positive with disease and they are actually negative.
- iv. **False Negative (FN):** A cases were predicted negative but they are truly negative.

Table 1.3: Confusion Matrix of Bagging

	Predicted Positive	Predicted Negative
Actual Positive	81	9
Actual Negative	4	57

Table 4.2 was used in the computational parts of performance measures as follows;

- i. **Accuracy:** One metric for assessing classification models is accuracy. Informally, accuracy refers to the percentage of correct predictions made by our model. The following is the formal concept of calculating accuracy:

$$Accuracy = \frac{\text{Number of correct prediction}}{\text{Total number of prediction}} \quad 4.1$$

The following formula can be used to measure accuracy in terms of positives and negatives for binary classification:

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

i.e Accuracy = 91%

- ii. **Precision:** Is the proportion of the model outcomes that are relevant to the instances. The number of true positives divided by the number of true positives plus the number of false positives refers to as precision.

$$Precision = \frac{\text{Number of True Positive}}{\text{Number of True Positive} + \text{Number of False Positive}} \quad 4.2$$

The following formula can be used to measure precision in terms of positives and negatives for binary classification:

$$Precision = \frac{TP}{TP + FP}$$

i.e Precision = 95%

- iii. **Recall (Sensitivity):** This is a measure of the proportion of actual positive cases that got predicted as positive (or true positive) (or true positive). Sensitivity is often termed as Recall.

$$Recall = \frac{TP}{TP + FN} \quad 4.3$$

i.e Recall = 90%

- iv. **F1_Score:** This combines the precision and recall of a classifier into a single metric by taking their harmonic mean. It is primarily used to compare the performance of two classifiers.

$$F1_{Score} = 2 * \frac{(Precision * Recall)}{(Precision + Recall)} \quad 4.4$$

$$F1_{Score} = 2 * \frac{(95 * 90)}{(95 + 90)}$$

$$i.e F1_{Score} = 93\%$$

- v. **Macro average:** This method is used when you want to know how the system performs overall across multiple datasets.

- a. **Precision Macro average:**

$$Macro_{Avg}(Precision) = \frac{(Precision_{class0} + Precision_{class1})}{2} \quad 4.5$$

$$Macro_{Avg}(Precision) = 91\%$$

- b. **Recall Macro average:**

$$Macro_{Avg}(recall) = \frac{(Recall_{class0} + Recall_{class1})}{2} \quad 4.6$$

$$Macro_{Avg}(recall) = 92\%$$

- c. **F1_Score Macro average:**

$$Macro_{Avg}(FScore) = \frac{(FScore_{class0} + FScore_{class1})}{2} \quad 4.7$$

$$Macro_{Avg}(FScore) = 92\%$$

- vi. **Weighted Average:** The weighted average is a statistic measure that accounts for the various degrees of numbers in a dataset.

- a. **Precision Weighted Average:**

$$Weighted_{Avg}(Precision) = \frac{SUMPRODUCT(precision_{range}, weights_{range})}{SUM(weights_{range})} \quad 4.8$$

$$Weighted_{Avg}(Precision) = \frac{13796}{151}$$

$$Weighted_{Avg}(Precision) = 91\%$$

- b. **Recall Weighted Average:**

$$Weighted_{Avg}(Recall) = \frac{SUMPRODUCT(Recall_{range}, weights_{range})}{SUM(weights_{range})} \quad 4.9$$

$$Weighted_{Avg}(Recall) = \frac{13773}{151}$$

$$Weighted_{Avg}(Recall) = 91\%$$

c. **F1_Score Weighted average:**

$$Weighted_{Avg(FScore)} = \frac{SUMPRODUCT(FScore_{range}, weights_{range})}{SUM(weights_{range})} \quad 4.10$$

$$Weighted_{Avg(FScore)} = \frac{13860}{151}$$

$$Weighted_{Avg(FScore)} = 91\%$$

Table 4.3: Bagging Classification Report

	Precision	Recall	F1_Score	Support
Class0	95	90	93	90
Class1	86	93	90	61
Accuracy			91	151
Macro Average	91	92	91	151
Weighted Average	92	91	91	151

Table 4.3 present classification report of bagging and the accuracy of the model is estimated at 85%, Precision 95%, Recall 96% and f1_Score at 90%.

Table 1.2: Confusion Matrix of Adaboost

	Predicted Positive	Predicted Negative
Actual Positive	79	11
Actual Negative	7	54

Table 4.3 was used in the computation of performance measures in the following way;

- vii. **Accuracy:** One metric for assessing classification models is accuracy. Informally, accuracy refers to the percentage of correct predictions made by our model. The following is the formal concept of calculating accuracy:

$$Accuracy = \frac{\text{Number of correct prediction}}{\text{Total number of prediction}} \quad 4.11$$

The following formula can be used to measure accuracy in terms of positives and negatives for binary classification:

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

i.e Accuracy = 88%

- viii. **Precision:** Is the proportion of model outcomes that are relevant to the instances. The number of true positives divided by the number of true positives plus the number of false positives refers to as precision.

$$Precision = \frac{\text{Number of True Positive}}{\text{Number of True Positive} + \text{Number of False Negative}} \quad 4.12$$

$$Precision = \frac{TP}{TP + FP}$$

$$i.e \text{ Precision} = 92\%$$

- ix. **Recall (Sensitivity):** This is a measure of the proportion of actual positive cases that got predicted as positive (or true positive) (or true positive). Sensitivity is often termed as Recall.

$$Recall = \frac{TP}{TP + TN} \quad 4.13$$

$$i.e \text{ Recall} = 88\%$$

- x. **F1_Score:** This combines the precision and recall of a classifier into a single metric by taking their harmonic mean. It is primarily used to compare the performance of two classifiers.

$$F1_{Score} = 2 * \frac{(Precision * Recall)}{(Precision + Recall)} \quad 4.14$$

$$i.e \text{ F1}_{Score} = 90\%$$

- xi. **Macro average:** This method is used when you want to know how the system performs overall across multiple datasets.

- a. **Precision Macro average:**

$$Macro_{Avg}(Precision) = \frac{(Precision_{class0} + Precision_{class1})}{2} \quad 4.15$$

$$Macro_{Avg}(Precision) = 87\%$$

- b. **Recall Macro average:**

$$Macro_{Avg}(recall) = \frac{(Recall_{class0} + Recall_{class1})}{2} \quad 4.16$$

$$Macro_{Avg}(recall) = 88\%$$

- c. **F1_Score Macro average:**

$$Macro_{Avg}(FScore) = \frac{(FScore_{class0} + FScore_{class1})}{2} \quad 4.17$$

$$Macro_{Avg}(FScore) = 88\%$$

- xii. **Weighted Average:** The weighted average is a statistic measure that accounts for the various degrees of numbers in a dataset.

- a. **Precision Weighted Average:**

$$Weighted_{Avg}(Precision) = \frac{SUMPRODUCT(precision_{range}, weights_{range})}{SUM(weights_{range})} \quad 4.18$$

$$Weighted_{Avg}(Precision) = 88\%$$

- b. **Recall Weighted Average:**

$$Weighted_{Avg}(Recall) = \frac{SUMPRODUCT(Recall_{range}, weights_{range})}{SUM(weights_{range})} \quad 4.19$$

$$Weighted_{Avg}(Recall) = 88\%$$

c. **F1_Score Weighted average:**

$$Weighted_{Avg(FScore)} = \frac{SUMPRODUCT(FScore_{range}, weights_{range})}{SUM(weights_{range})} \quad 4.20$$

$$Weighted_{Avg(FScore)} = 88\%$$

Table 4.5: Adaboost Classification Report

	Precision	Recall	F1_Score	Support
Class0	92	88	90	90
Class1	83	89	86	61
Accuracy			88	151
Macro Average	87	88	88	151
Weighted Average	88	88	88	151

Table 4.5 shows Adaboost current classification report, with the model's accuracy evaluated at 88%, Precision 92%, Recall 88%, and f1 Score 90%.

11.2 Comparison Chart for Extended Classifiers

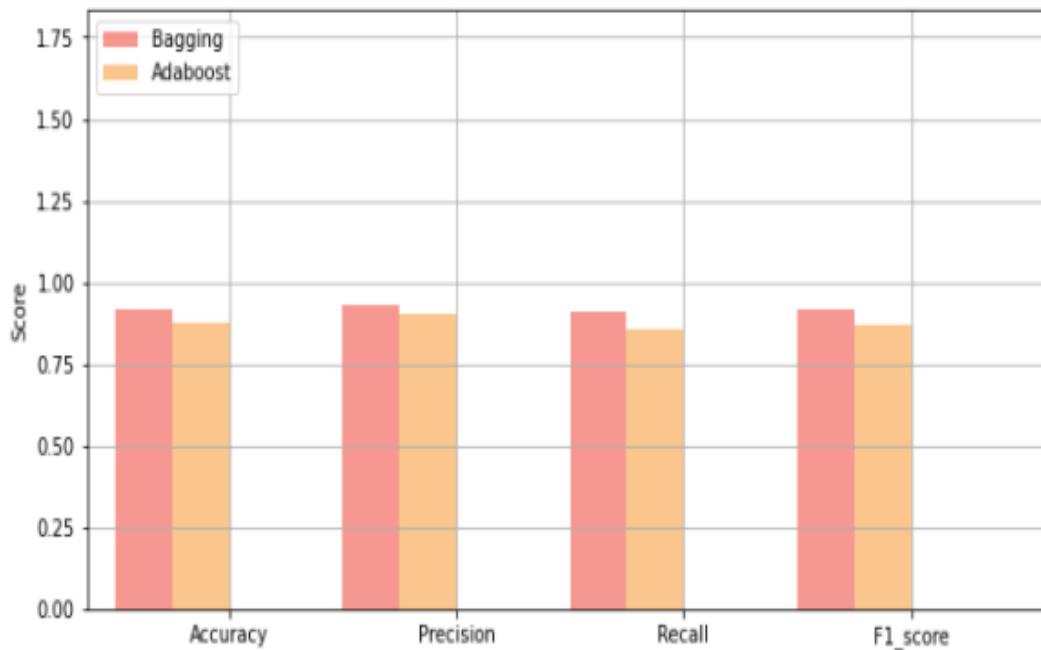


Figure 1.2: Comparison Chart for Extended Classifiers

Figure 4.5 depicts the performance of the extended classifiers utilized and their predictive accuracy, from which the researcher chose bagging as the classifier with the best predicting performance for the study's implementation.

12 DISCUSSION OF RESULTS

Ensemble technique is a well-proven methodology used in research for achieving extremely accurate data classification by integrating several machine learning algorithms in order to get more dependable and accurate prediction outcomes.

As shown in figure 4.5, from the two ensemble classifiers used in this research; Bagging has higher performance measures with an accuracy of 91%, precision of 95%, recall of 90%, and F1 Score of 93%, whereas Adaboost has an accuracy of 88%, precision of 92%, recall of 90%, and F1 Score of 90%. As a result, our model was implemented using bagging.

13 FINDING OF THE STUDY

The study's findings suggest that Bagging is the best ensemble classifier to be adopted in the implementation of ensemble model for heart disease presence in patients. The study model developed; is dependable, efficient, accurate, and can be used in modern health care facilities for the prediction of heart disease.

14 SUMMARY

This research work is aimed at developing an ensemble model for heart disease prediction. In the implementation process, the researcher uses dataset from UCI Machine Learning repository.

The Researcher develops the ensemble model for heart disease predictive analysis by employing python as a programming language. The dataset used in this research contains 303 instances and 14 attributes. In the evaluation; the researcher used of Naïve Bayes, Decision Trees, and Support Vector Machine as a base machine learning algorithms, along with Adaboost and Bagging as extended classifiers that combines the weaker algorithm which has been presented as a solution for heart disease prediction.

The model was validated using the 10-fold cross-validation methodology and it had an accuracy of 91%, precision of 95%, recall of 90%, and F1 Score of 93%. Additionally, a comparison of extended classifiers was conducted, with the outcome revealing that of the two meta-algorithms used; Bootstrap Aggregating is the best. As a result, the researcher employs bagging as part of an ensemble strategy for predicting heart disease.

15 CONCLUSION

Prediction of heart disease has been the most promising approach that will help in the suitable reduction in the fertility rate and will assist in the development of the alerts on the large scale.

The proposed model for heart disease prediction using ensemble method discussed in this research was successfully implemented and simulated using python. The model captures symptoms and risk factors to diagnose a patient with heart disease and presented.

Finally, based on the findings of this research, the researchers concluded that an ensemble method of predicting heart disease produces better prediction and achieves greater performance

than any one contributing model and the classification report of the study shows that bagging outperforms adaboost classifier.

16 RECOMMENDATIONS

The researcher makes the following recommendations based on the study's findings and the substantial benefits of heart disease prediction using ensemble method:

- i. That the suggested ensemble model for heart disease predictive analysis is recommended for usage in the medical industry in order to address today's healthcare concerns.
- ii. That the implementation of this model will aid medical practitioners in anticipating the presence of heart disease in patients in a very efficient and effective manner, as well as ensuring the medical expert's judgment system is free of challenges.
- iii. That the remarkable comparative results of this study suggest that the bagging classifier should be used on any medical illnesses where prediction conditions are difficult to predict.
- iv. It is also suggested that scholars broaden their research to encompass aspects not addressed in this study.

REFERENCES

1. Alan, D., Barbara, H. W. & David, T. (2015). *Systems Analysis & Design: An Object-Oriented Approach with UML* (5th Edition). East Carolina University. ISBN 978-1-118-80467-4.
2. Chadha, R. & Shubhankar, M. (2016). Prediction of heart disease using data mining techniques. *CSI Transactions on ICT*. 4. 10.1007/s40012-016-0121-0.
3. Chen, A. H., Huang, S. Y. Hong, P. S., Cheng, C. H. & Lin, E. J. (2011). Heart disease prediction system. *Computing in Cardiology, Hangzhou*, Vol. 5(8), PP. 557-560.
4. Das, R. Ibrahim, T. & Abdulkadir, S. (2009). Effective diagnosis of heart disease through neural networks ensembles. *Expert Systems with Applications*, Vol.36(4), PP. 7675-7680.
5. Dhanashree, S. M., Mayur, P. B. & Shruti, D. D. (2013). Heart Disease Prediction System using Naive Bayes. *International Journal of Enhanced Research in Science Technology & Engineering*. Vol.2(3), PP. 2319-7463.
6. Hannan, S., Alane, R., March, S., Salvatore, T., Park, J., & Ram, S. (2010). Prediction of heart disease medical prescription using radial basis function. *IEEE International Conference on Computational Intelligence and Computing Research, Coimbatore*, 2010, PP. 1-6.
7. Kamal, K. & Kanwal, G. (2014). Review of Heart Disease Prediction using Data Mining Classifications. *International Journal for Scientific Research & Development* Vol.2(4), PP. 2321-0613.
8. Kerzner, B. (2009). Clinical Investigation of Feeding Difficulties in Young Children: A Practical Approach. *Clinical Pediatrics*, Vol.48(9), PP. 23-29.
9. Manpreet, S., Martins, L., Joanis, P., & Mago, V. (2016). Building a Cardiovascular Disease predictive model using Structural Equation Model & Fuzzy Cognitive Map. 1377-1382. 10.1109/FUZZ-IEEE.2016.7737850.

10. Mujawar, S. H. & Devale, P. R. (2015). Prediction of Heart Disease using Modified K-means and by using Naive Bayes. *International Journal of Innovative Research in Computer and Communication Engineering*. Vol.3(10), DOI: 10.15680/IJIRCCE.2015. 0310123.
- Nikookar, E. & Naderi, E. (2018). Hybrid Ensemble Framework for Heart Disease Detection and Prediction. *International Journal of Advanced Computer Science and Applications*, Vol.9(10), www.ijacsa.thesai.org
- Prajakta, G., Vrushali, G., Kajal, K. & Prajakta, D. (2016). Intelligent Heart Attack Prediction System Using Big Data. *International Journal of Recent Research in Mathematics Computer Science and Information Technology*. Vol.2(2), PP. 73-77.
- Sarwar, A., Sharma, V. & Gupta, R. (2015). Hybrid ensemble learning technique for screening of cervical cancer using Papanicolaou smear image analysis. *Personalized Medicine Universe*, Vol.4(2), PP. 54-62.
- Soni, J., Ansari, U., Sharma D. & Soni, S. (2011). Predictive data mining for medical diagnosis. An overview of heart disease prediction. *International Journal of Computer Applications*, Vol.17(8), PP. 43-48.
- Sudhakar, K. & Manimekalai, M. (2014). Study of Heart Disease Prediction using Data Mining. *International Journal of Advanced Research in Computer Science and Software Engineering*. Vol.4(1), www.ijarcsse.com