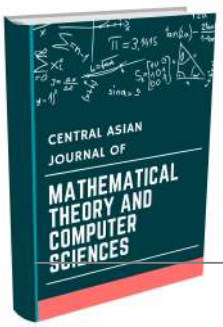

CENTRAL ASIAN JOURNAL OF MATHEMATICAL THEORY AND COMPUTER SCIENCES
https://cajmtcs.centralasianstudies.org

Volume: 04 Issue: 5 | May 2023 ISSN: 2660-5309

HEART DISEASE DIAGNOSIS WITH TREE STRUCTURAL NAÏVE BAYES

Hussein. M Jebur[1], Zainab marid Alzamili[2], Ali Hasan Ali[3]
[1,3]College of Information Technology, Imam Ja'afar Al-Sadiq University, Baghdad, Iraq
[2]Education Directorate of Thi-Qar, Ministry of Education, Iraq
[1]hussein.mankhi@sadiq.edu.iq
[2]zainab.alzamili@utq.edu.iq, [3]ali.hasaan@sadiq.edu.iq

Abstract

Among the world population, the Disease of Heart is one of the biggest mortality and morbidity causes. This disease's precise prediction and early detection might decline rate of mortality rate certainly. Learning machines are utilized to consider several problems in the science of information. In Fortune, one of efficient methods for classification is naïve bayes (NB) is because of the ability of it for learning inherent features of data. Although, generally such method groups data with just one single that makes this less efficient relatively in several classes for big classification issue. In the article, we present the tree structural naïve bayes (Tree-NB) that classifies big classification in small classifications with utilizing structure of tree. The particular classifier is adjusted after division for every small classification. By several classifiers that are employed, Tree-NB is able to complement each other in performance of classification as well as one classifier issue is solved. As all several classifiers are end-to-end frameworks, automatically Tree-NB is able to learn nonlinear relationship among output and input data with no extraction of feature. For verifying our model validity, we compare modern methods with Tree-NB by utilizing dataset of UCI. Experimental results illustrate that Tree- NB is able to obtain the higher performance in less time of training. Average Tree- NB accuracy is 1.19 % higher than the other modern methods also it possesses higher average recall and precision.

ARTICLEINFO

Article history:
Received 25 Mar 2023
Revised form 26 Apr 2023
Accepted 27 May 2023

Keyword: *Heart disease, naïve bayes, machine learning, data mining.*

1- Introduction

Nowadays, everybody is so much busy in their work and lives that they do not have time for considering themselves. People experience anxiety, depression, stress as well as a lot of other issues most of the time because of the hectic lives of themselves. Taking account of such things as the basic factors, they are having severe diseases and getting sick. A lot of diseases exist like tuberculosis, heart disease, cancer and so on that causes people death per year however the highest deaths rate in medical fled is with having disease of heart [1].

One of the blood vessels and heart disease causes, this is assumed that by 2030 the amount will raise up to 23 million. In several decades, in population diseases of cardiovascular have been disease leading causes main section. In domain of healthcare, because of large data value (huge data) which is created of several sources and fields like high throughput instruments, networks of sensor, mobile application, IOT, streaming machines, advanced healthcare systems, processing and collecting data is becoming so usual nowadays [2].

Several data mining methods have been integrated to diagnose disease and achieve multiple probabilities. Taking account of the prediction of heart disease multiple systems are being suggested that are being extended with several algorithms and techniques means. Obtaining the service of quality at the cost-effective price set the challenging and prime issue for establishments of healthcare. To offer the services of quality at par, patients precise diagnosis should be with the efficient medicines dosage. Clinical treatment and diagnosis with low quality is able to be in inadequate and undesired outcomes [3]. Various data mining methods' kinds are used for data mining prediction such as Decision tree, Neural Network, algorithm of KNN, Naïve Bayes. Naïve Bayes is used for predicting heart diseases probability. Most of the frameworks of classification according to the ML use just one classifier usually that influences performance on big classification issue.

In the article, we present the model of classification called tree structural naïve bayes (Tree-NB). Such model shares classes of data in nodes in structure of tree after that adjusts the particular classifier for every node. This is able to develop performance by several classifiers about big classification issue. Uniformly such classifiers utilize model of naïve bayes that is able to can learn raw data features automatically. Model of naïve bayes inputs data which is preprocessed in module of naïve bayes to learn the feature after that achieves outcomes of classification. Tree-BN model classification process is that data begins from tree structure root node also via classification with several classifiers, lastly data class which is predicted is able to be achieved. As all several classifiers are models of ML, Tree-NB is able to group data with no sets of feature that are designed manually.

This article is briefed as following. At first, we present the model of naïve bayes also defines particular model structure. Secondly, we achieve outcomes that small classification effect is better rather than that of big classification through experiments. Additionally, we verify proposed model validity with comparing that with modern techniques on public set of data. At last, one end-to-end model is Tree-NB that prevents designing sets of feature impact on effect of classification.

This article remainder is classified as below. Part two defines the related work. Part three defines particular proposed model framework details. Part four illustrates experiments on results analysis as well as public set of data. At last, final outlook and remarks are mentioned in part five.

2- Related work

Various works have been performed given the diagnosis of heart disease by utilizing various techniques of data mining. Methods, algorithms and set of dataset are utilized by writers also supervised the outcomes with work of future is performed in realizing the effective medical diagnosis methods for different diseases [4]. ML is the technology of data analysis which trains the computers for acting like humans. This utilizes the computational methods for directly extracting information from data. ML algorithm The performance is developed based on data quality and increasing prediction process of disease [5].

Prasad and Muibideen [6] presented the network model of Bayesian for the prediction of heart disease in human being. Such model was created by utilizing package of bnlearn in R. This paper aim is comparing Bayesian classifiers effectiveness in predicting heart diseases. They utilized two various Bayesian classifier

implementations: Naïve Bayes, Bayesian Belief Network. Bayesian Belief Network generated graphical dependencies representation among features. Achieved model aids us for identifying normal conditional independencies and dependencies among features. Bayesian Belief Network performed better rather than Naïve Bayes in heart diseases prediction.

Porkodi and his colleagues [7] presented the great and efficient system for predicting system of heart disease according to systems of ML. Such system is classified with utilizing different algorithms of classification like DT, NB, LR, RF, KNN, SVM. Proposed technique which is algorithmic solves feature selection issue also raises classification accuracy. Additionally, proposed algorithmic technique with that could utilize the non-invasive clinical data for assessing the severity and diagnosis of heart Disease. New multiple method implementation aids for developing EDA diagnosis accuracy. Outcomes prove that in healthcare domain this is able to be implemented easily and has more accuracy for identifying disease of CAD.

Vanitha Guda and his colleagues [8] present the method which targets to find the main attributes with using techniques of ML causing to develop accuracy in cardiovascular disease prediction. Model of prediction is described by several known techniques of classification and various features combinations. They present the improved level of performance by high accuracy via the model of prediction for heart disease with multiple methods.

Kavitha and his colleagues [9] utilized Cleveland dataset of heart disease also methods of data mining like classification and regression are utilized. Techniques of ML and Decision Tree, Random Forest are used. New ML model method is designed. In implementation, three algorithms of ML, which are one, are utilized. Random Forest, 2. Decision Tree and 3. Multiple model (decision tree and random forest Hybrid). Experimental results illustrate the level of as 88.7 percent via prediction model of heart disease with multiple model. Interface is designed for getting input parameter of user for predicting heart disease, for what they utilized the multiple Random Forest and Decision Tree model.

Mohan and his colleagues [10] present the new method targets to find the specific attributes with using ML methods causing to develop cardiovascular disease prediction accuracy. Model of prediction is described by several known classification techniques and features different combinations. They generate the improved level of performance by the level of accuracy as 88.7 percent via model of prediction for heart disease by multiple random forest by the linear model (HRFLM).

Khourdifi and his colleagues [11] exploited method of Fast Correlation-Based Feature Selection (FCBF) for filtering the redundant attributes for developing classification of heart disease quality. After that, they carry out the classification according to various algorithms of classification like SVM, Random Forest, K-Nearest Neighbour, Naïve Bayes, Multilayer Perception | ANN optimized by PSO integrated approaches of with Ant Colony Optimization (ACO). Proposed mixed approach is used for dataset of heart disease; outcomes show proposed hybrid method robustness and efficacy in processing different data kinds for the classification of heart disease.

Singh and Kumar [12] compute ML algorithms accuracy to predict heart disease, such algorithms include linear regression, SVM, decision tree, k-nearest neighbor with utilizing repository dataset of UCI in order to test and train. This is concluding that knn is best between them with accuracy of 87 percent.

3- Proposed method

In the article, we present the tree structural naïve bayes (Tree-NB) that classifies big classification in small classifications with utilizing structure of tree. The particular classifier is adjusted after division for every small classification. By several classifiers applied, Tree-NB is able to complement each other in performance of classification, one classifier issue is solved. As all hybrid classifiers are end-to-end frameworks, Tree- NB is able to learn nonlinear relationship automatically among output and input data with no extraction of feature.

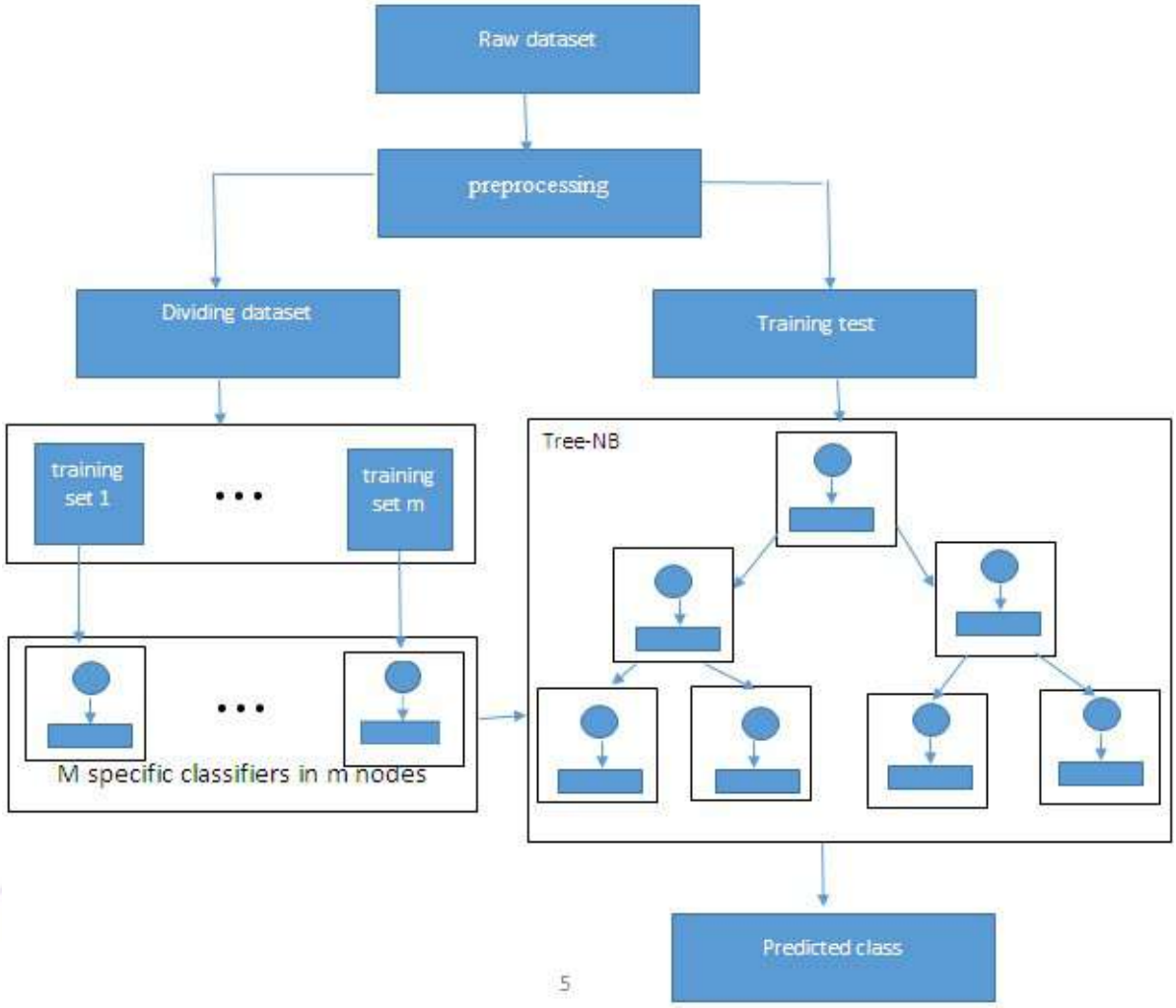

Fig. 1. proposed Tree-NB framework.

- **Step of Pre-processing**

Here, Dataset is preprocessed with changing symbolic valued features to numeric then using algorithm of discretization.

- **Step of classifying set of data**

Randomly sharing split file in one set of testing and training process in proportion is Data divide.

Raw set of data is shared for achieving *m* various sets of data for special classifier in *m* nodes, *m sets of* training are formed after procedure of preprocessing that are utilized for training classifiers in relating nodes in model of Tree-NB. In process of testing, raw set of data is preprocessed for transforming the set of testing. After that, set of testing is input in trained model of Tree-NB for achieving predicted class also assessing proposed model performance.

- **Step of Tree structural naïve bayes (Tree-NB)**

For obtaining better performance of classification, basic Tree-NB idea is leveraging structure of tree with classes of data in nodes and set of classifier for every node for implementing classification of data. Classes of data are shared in nodes in tree. Whole classes are featured for node of root as well as the other nodes, classes are featured based on the situations. After that, the particular classifier is adjusted for every node that the classifier utilized model of NB for learning data features that are time-related. After division for structure of tree structure, input model data are first in tree structure first layer sent to node of root node, secondly, data are judged with particular classifier for determining that node in second layer data must be sent also recursively solve.

➢ **Step of Naïve Bayes classifier**

The algorithm is obtained from classifier of Bayes that is robust in dividing irrelevant features as well as points of noise. Additionally, models of prediction are created quickly by the algorithm. Aware of base of data D with n features (a_1, a_2, . . ., a_n), m labels of class (c_1, c_2, . . ., c_m). sample x with form x=(a_1, a_2, . . ., a_n) is regarded to class ci if this has the highest conditional probability (Formula 1) [13]:

$$P(C_i \mid X) > P(C_k \mid X) \qquad k = 1 \text{ to } m \text{ and } k! = i \tag{1}$$

Probabilities, $P(C_i|X)$, are computed by utilizing theorem of Bayes regarding Formula 2 [13]:

$$P(C_i \mid X) = P(X / C_i) P(C_i) / P(X) \tag{2}$$

As Formula 2 denominator is equal for whole terms, this just remains for us for maximizing numerator. $P(C_i)$ is computed as below [13]:

$$P(C_i) = \frac{S_i}{S} \tag{3}$$

where S_i is samples number given class I, S is whole records or samples number. For decreasing computational complexity, P(X|C) is computed as below [13]:

$$P(X \mid C_i) = \Pi_{k=1\ldots n} P(X_k \mid C_i) \tag{4}$$

where X_k is a_k feature value for record X. If a_k is deterministic, $P(X_k|C_i)$ is computed as [13]:

$$P(X_k \mid C_i) = \frac{S_{ik}}{S} \tag{5}$$

where S_{ik} is records number which is labeled with class i that values equal to X_k, Si is records' number which are labeled by class i. even If a_k is going on, probability is calculated by utilizing distribution of Gaussian.

➢ **Step of divided rules and Tree choice**

For guaranteeing that every classification in structure of tree is small classification, the binary tree kind is accepted thus node degree in tree is not greater rather than 2. So, except for last classes classification in nodes of leaf, whole rest are 2 classifications. While classes' number in node is equal/more rather than 4, this is able to be shared in 2 classifications; on the other hand, this is able to be classified with no division directly. At last, classes' number in nodes of leaf is three/two that guarantees classifier in every node carries out 2/3 classification. For traffic classification of network with N traffic classes number, due to big classification issue is taken into consideration, this is estimated that N is equal/ greater rather than 4. Particular rules which are shared are as below [14]:

(1) For node of root, define whether N/2 is the integer or not. Even if this is, classes' number in the left and right side of it, nodes of child is N/2; on the other hand, classes' number in right and left nodes of child is (N-1)/2, (N+1)/2, respectively for normalization. We estimate i presents node level, allow i = 2.

(2) i layer node is sequentially traversed from left -right. We estimate classes' number in node j is M_{ij}. Even if M_{ij} is less rather than 4, division is stopped and node is marked; on the other hand, after defining whether $M_{ij}/2$ is an integer or not, we carry out similar operation as (1).

(3) division is completed if whole present tree leaf nodes are marked; on the other hand, adjust $i = i + 1$ then turn (2).

According to classes' amount, tree height is able to be allocated based on the rule which is shared. Estimated tree height is k, nodes' number in k layer is at most $2k-1$, classes' number in every node of leaf is at most 3, thus maximum classes amount in tree with height k is $2k-1 \times 3$. Min classes number is max classes number in tree with height k-1 + 1 that is $2k-2\times3+1$. Thus, while traffic classes amount is N, if k is tree height, integer k satisfies inequality that, $2k-2 \times 3 + 1 \leq N \leq 2\ k-1 \times 3$. But, due to classes number in nodes of leaf is three/two, while $2k-1 \times 2 \leq N \leq 2\ k-1 \times 3$, nodes' amount in the k layer derives the max value. for N division by various amounts.

Cosine is utilized for judging similarity among classes. With comparing similarities among various classes, this is able to be allocated that the classes are featured to similar node. After procedure of preprocessing, one packet of data in raw set of data is able to be presented by an array which is two-dimensional that is able to be flattened out in a vector that is dimensional. After the preprocessing procedure with averaging whole classes' data, class vector is able to be achieved, difference among vectors is measured with cosine for judging similarity among the classes.

- **Step of classification of data by Tree-NB**

In the part, we use publicly accessible UCI dataset of heart disease also teach five model of classification for data with proposed 2-height Tree-NB. Particularly, in second 2-height Tree-NB possesses two nodes with three classes, named as Class A as well as Class B. They are whole featured in second layer to the nodes. Whole classes are featured to the nodes of root. At last, the particular classifier is adjusted for every node in a tree which is shared.

Several models require for being taught while the process of training after division. Affected with various results of classification, various sets of training are required for training models. Abovementioned process needs 3 classifiers that are recorded as Classifier 1 (A, B classification), Classifier 2 (A classification) and Classifier 3 (B classification).

Tree-NB model process of evaluation by set of testing is defined as below: (1) this is able to be judged that data related to class A/class B via Classifier 1. (2) last data class which is predicted is able to be achieved via Classifier 2/Classifier 3. (3) Comparing actual and predicted class, we achieve file to assess model of Tree-NB.

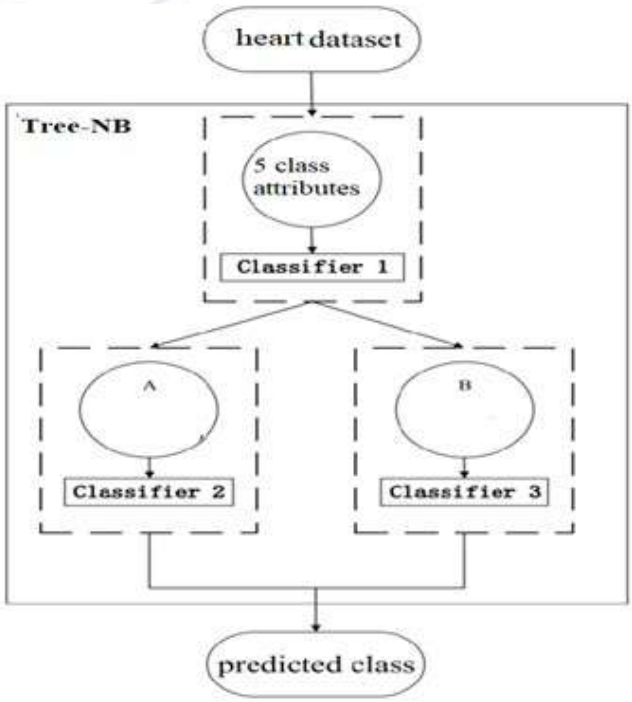

Fig. 2. Structure of Tree-NB.

4- Experiment and Result

In the part, we have attempted for comparing proposed method performance with the other methods of ML for prediction of heart disease [6–8, 10]. Whole factors that have been taken into consideration for experiment are mentioned in subsection below.

4.1. Dataset description

A set of data which is used in most articles of exploration is dataset of heart disease which is gotten of UCI (University of California, Irvine C.A) Center for canny, AI frameworks. This includes 4 bases of information from 4 clinics. Each set of data possesses same highlights' amount that is 14, still different records' quantities [15]. The most used dataset by Al scientists is the Dataset of Cleveland due to including less missing ascribes rather than sets of data as well as possessing much more records. Domain of "num" alludes heart disease existence in sick persons. This is the number which is esteemed from zero-four (no presence). Dataset of Cleveland includes 303 samples. Various features of set of data with the amounts of them are defined in table 1.

Table: 1: dataset attributes

No	Name	Description
1	Age	Age in Years
2	Sex	1=male, 0=female
3	Cp	Chest pain type(1 = typical angina, 2 =atypical angina, 3 = non-anginal pain, 4 = asymptomatic)
4	trestbps	Resting blood sugar(in mm Hg on admission to hospital)
5	Chol	Serum cholesterol in mg/dl
6	Fbs	Fasting blood sugar>120 mg/dl(1= true, 0=false)
7	restecg	Resting electrocardiographic results(0 = normal, 1 = having ST-T wave abnormality, 2 = left ventricularhypertrophy)
8	thalach	Maximum heart rate
9	exang	Exercise induced angina
10	oldpeak	ST depression induced by exercise relative to rest
11	slope	Slope of the peak exercise ST segment (1=upsloping, 2=flat, 3= downsloping)
12	Ca	Number of major vessels colored by fluoroscopy
13	thal	3= normal, 6=fixed defect, 7= reversible defect
14	Num	Class(0=healthy, 1=have heart disease)

4.2. Evaluation metrics

Here, we utilize the issue which is two-classification for specifying metrics. In the issue which is two-classification, four terms exist: false positive (FP), false negative (FN), true positive (TP), true negative (TN). True positive is a term that predict the positive sample accurately as the positive sample; false positive is a term which is predict a negative example wrongly as the positive sample; true negative is a term which predict a negative example accurately as the negative sample; false negative is a term which predict a positive sample wrongly as the negative sample. Matrix of confusion is the standard format to express effect of classification where the columns show the classes which are predicted also rows show real actual classes. Table. 2 illustrates the common matrix of confusion in an issue which is two-classification.

Table 2: Confusion Matrix

Actual Label	Predicted Label	
	+(1)	-(0)
+(1)	True Positive	False Negative
-(0)	False Positive	True Negative

Accuracy is utilized for judging all effect of classification that is regarded to whole ratio accurately the samples which are predicted to the whole examples. This is able to be computed as it is mentioned in equation (6).

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{6}$$

- **Precision**: Precision illustrate that what positive identifications proportion are really accurate. We are able to describe the precision with utilizing Equation (7).

$$\text{Precision} = \frac{TP}{TP+FP} \tag{7}$$

where TP is amounts of True-positive, FP is amounts of False positive.

- **Recall**: Recall illustrates what real positives proportion are accurately recognized. We are able to describe the with utilizing equation (8).

$$\text{Recall} = \frac{TP}{TP+FN} \tag{8}$$

where TP is amounts of True-positive, FN is amounts of False negative.

4.3. Result and discussion

The part includes basic study outcomes or findings as well as the related arguments about the related input parameters performance and performance of model (in comparison with last techniques).

We have compared also tested whole algorithms according to accuracy. To assess the accuracy, we have utilized famous assessing metrics like precision, recall, accuracy. Five algorithms' outcomes are described in Table 3 in precision, recall, accuracy case. As you can see in Table 3, we are able to supervise that the proposed method has outperformed and whole other algorithms in accuracy case as well as other metrics of assessing. Also, as it is illustrated in Figure 3, Figure 4 and Figure 5 whole algorithms have been graphically compared in accuracy as well as the other metrics case, respectively.

Therefore, based on the achieved outcomes that are illustrated in Table 3 and which are shown graphically in Figure 3, Figure 4 and Figure 5 we have concluded that algorithm of proposed method predicts the best outcome. Now, accuracy is 89.66 percent that is the highest rate between the other algorithms. We know that as the precision is higher, recall better is the result, therefore; in proposed method naïve bayes is the algorithm that has the highest recall and precision. Whole algorithms Comparison is shown in Table 2.

Table 2. Comparison of all the algorithms

Algorithms	Accuracy (%)	Precision	Recall
Paper [6]	85	86	85
Paper [7]	85.25	-	-
Paper [8]	80.22	81.25	81.25
Paper [10]	88.47	87.5	92.8
Proposed method	89.66	87.81	97.30

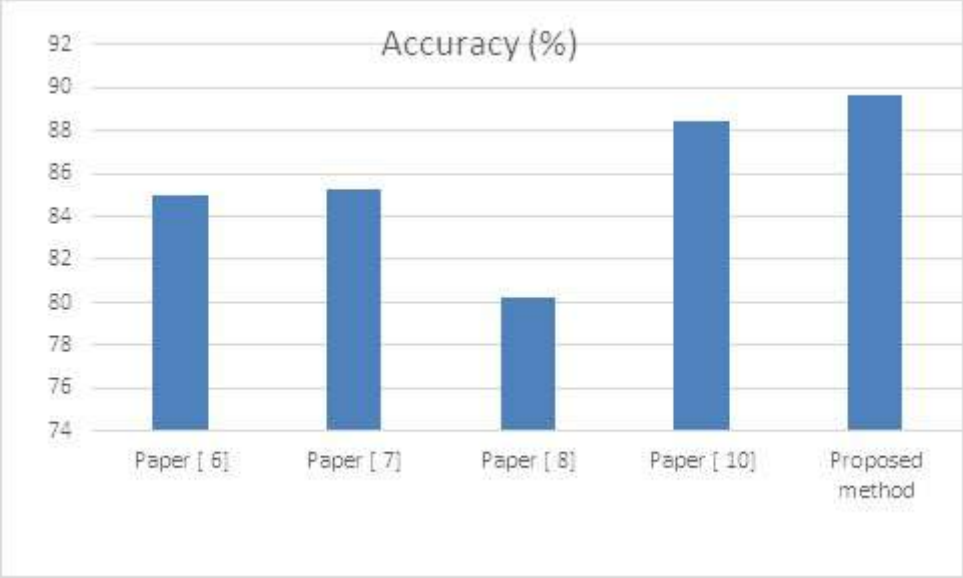

Figure.5 Comparison of algorithms in terms of accuracy

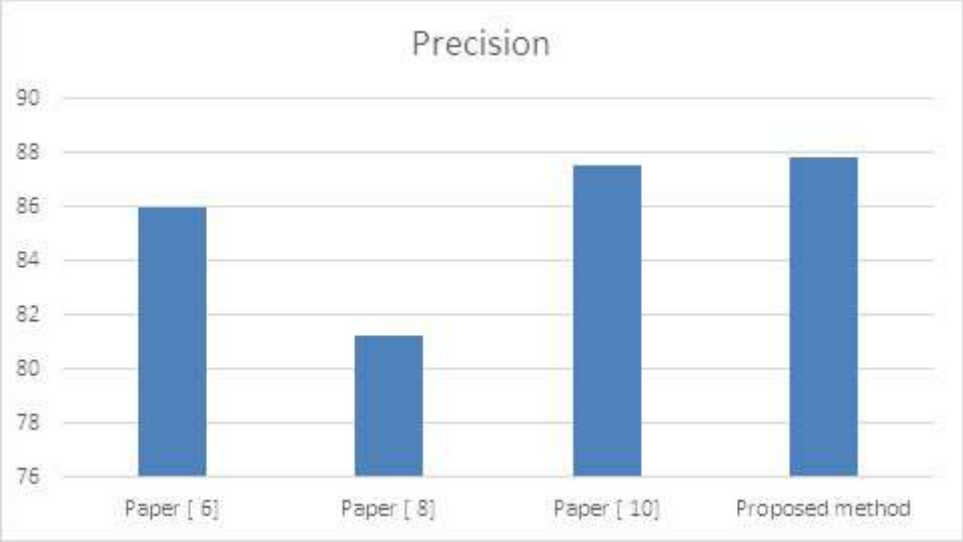

Figure.6 Comparison in terms of evaluating metrics (Precision)

Figure.7 Comparison in terms of evaluating metrics (Recall)

Conclusion and Future Work

We have recognized that having heart disease risk is able to be a serious issue Via the present research. We have learned various agents which are able to raise having a heart disease chances. Also, the research defines various heart disease kinds that an individual is able to suffer from as well as what is able to be symptoms. Here, we present the Tree-NB that shares big classification in small ones with utilizing structure of tree on set of data of UCI for the prediction of heart disease. We attempted in order to know which algorithm is able to aid in correct results as well as better prediction for heart disease also realizing that random forest is carried out on the taken dataset the best. Naïve bayes presented the best outcomes accuracy and the other metrics of assessing case. Todays, for the better results Deep learning is being utilized in each field. In future, by utilizing Deep learning for performing the results of naïve bayes better we have aim for implementing several some other algorithms. In the future, several other sensors will be added, such as the temperature and humidity sensor, to be sent also through the application, and the project can also be extended from a fixed object to a mobile one that is sent to places containing fires to measure the level of gas leakage or the level of fire occurring in a specific location[16].

Although the two topics in this Section title don't necessarily go hand in hand, they do represent basic constructs essential for any program. In this Section, you closely observe both procedures and conditions in order to establish some basic tools with which to work in VBA. Specifically, in this Section I will discuss [17].

References

1. Katarya, R., & Meena, S. K. (2021). Machine learning techniques for heart disease prediction: a comparative study and analysis. Health and Technology, 11(1), 87-97.
2. Mary, M. M. A., & Beena, T. L. A. (2020). Heart disease prediction using machine learning techniques: A survey. *Int. J. Res. Appl. Sci. Eng. Technol.*, *8*(10), pp. 441-447.
3. Repaka, A. N., Ravikanti, S. D., & Franklin, R. G. (2019). Design and implementing heart disease prediction using naives Bayesian. In 2019 3rd International conference on trends in electronics and informatics (ICOEI), pp. 292-297.
4. Bajaj, P., & Gupta, P. (2014). Review on heart disease diagnosis based on data mining techniques. International Journal of Science and Research (IJSR), 3(5).
5. Alsheref, F. K., & Gomaa, W. H. (2019). Blood diseases detection using classical machine learning algorithms. International Journal of Advanced Computer Science and Applications (IJACSA). Blood, 10(7).
6. Muibideen, M., & Prasad, R. (2020). A Fast Algorithm for Heart Disease Prediction using Bayesian Network Model. arXiv preprint arXiv:2012.09429.
7. Porkodi, K., Aishwarya,, A., Divya, R., Indhuja, K., & Manasa, S. (2020). Efficent classification of heart disease using machine learning algorithm. Journal of Xi'an Shiyou University, Natural Science Edition, 17(7), pp. 59-66.
8. Vanitha Guda, S. K., & Shivani, C. (2020). Heart Disease Prediction Using Hybrid Technique. Journal of Interdisciplinary Cycle Research. 6(5)920-927.
9. Kavitha, M., Gnaneswar, G., Dinesh, R., Sai, Y. R., & Suraj, R. S. (2021). Heart Disease Prediction using Hybrid machine Learning Model. In *2021 6th International Conference on Inventive Computation Technologies (ICICT)*, pp. 1329-1333.
10. Mohan, S., Thirumalai, C., & Srivastava, G. (2019). Effective heart disease prediction using hybrid machine learning techniques. *IEEE access*, *7*, pp. 81542-81554.

11. Khourdifi, Y., & Bahaj, M. (2019). Heart disease prediction and classification using machine learning algorithms optimized by particle swarm optimization and ant colony optimization. *International Journal of Intelligent Engineering and Systems*, *12*(1), pp. 242-252.

12. Singh, A., & Kumar, R. (2020). Heart disease prediction using machine learning algorithms. In *2020 international conference on electrical and electronics engineering (ICE3)*, pp. 452-457.

13. Sajjadnia, Zeinab, Raof Khayami, and Mohammad Reza Moosavi. "Preprocessing Breast Cancer Data to Improve the Data Quality, Diagnosis Procedure, and Medical Care Services." Cancer Informatics 19 (2020): 1176935120917955.

14. Ren, X., Gu, H., & Wei, W. (2021). Tree-RNN: Tree structural recurrent neural network for network traffic classification. *Expert Systems with Applications*, *167*, 114363.

15. Durairaj, M., & Sivagowry, S. (2014). A pragmatic approach of preprocessing the data set for heart disease prediction. *international journal of Innovative Research in computer and communication Engineering*, *2*(11), 6457-6465.

16. Ali Hasan Ali 2023. Smart Fire System using IOT. CENTRAL ASIAN JOURNAL OF MATHEMATICAL THEORY AND COMPUTER SCIENCES. 4, 3 (Apr. 2023), 88-113.

17. Hasan Ali , A., M Jebur, H., & Alzamili, Z. marid J. (2023). DESIGN OF A VIRTUAL REALITY SIMULATOR OF A DORMITORY BY USING EXCEL VBA. CENTRAL ASIAN JOURNAL OF MATHEMATICAL THEORY AND COMPUTER SCIENCES, 4(5), 99-119. https://doi.org/10.17605/OSF.IO/CMY8X.