Volume 4, Issue 08, August 2019 ISSN (online): 2456-0006

International Journal of Science Technology Management and Research

Available online at: www.ijstmr.com

An Efficient Approach for Early Prediction of Cardiovascular Disease

Neha Makvana
M Tech(C.S.E.) 4 semester
Lord Krishna College of Technology
Indore M.P. India

Vijay Kumar Verma
Assistant Professor C.S.E. Department
Lord Krishna College of Technology
Indore M.P. India

Abstract: ***Diagnosis of Cardiovascular Disease is a significant task in medical science. The term Cardiovascular Disease includes the various diseases that involve the heart attack problem. The exposure of Cardiovascular Disease problem from different symptoms is an important issue for predicting problem. This research includes the study of various classification techniques like Classification by decision tree induction, Bayesian Classification, Rule-based classification, Classification by back propagation, Support Vector Machines (SVM) Neural Network as a Classifier The k-Nearest Neighbor Algorithm and Classification using Genetic Algorithms (G.A.)***

In this paper we have taken 9 attribute which are responsible for the Cardiovascular Disease problem. We used normal condition suggested by the physician for Cardiovascular Disease and the condition which are responsible for Cardiovascular Disease. Possible conditions for Cardiovascular Diseas Age>45, BP>120, Cholesterol range>240, Resting ECG>1, Thalach value>100 Beats/Minute, Old peak>0,Slope>=2,ThalValue>3. We used an efficient approach to predict Cardiovascular Disease. We find the distance between the training data set with the given data set for Cardiovascular Disease. We find those records which has minimum distance with the given tuple. We used Euclidian distance to find the distance value. Consider only those tuple which have minimum distance and found the class value which is yes or no. We found the percentage value for class yes and for class No. To compare the performance of the proposed approach we used Bayesian classifiers

Keywords: Disease ; Classification; Distance; Attribute, Cardiovascular

I. INTRODUCTION

Classification is a classic data mining technique based on machine learning. Basically classification is used to classify each item in a set of data into one of predefined set of classes or groups. Classification method makes use of mathematical techniques such as decision trees, linear programming, neural network and statistics. Classification divides data samples into target classes. The classification technique predicts the target class for each data points. For example, patient can be classified as “high risk” or “low risk” patient on the basis of their disease pattern using data classification approach. It is a supervised learning approach having known class categories. Binary and multilevel are the two methods of classification. In binary classification, only two possible classes such as, “high” or “low” risk patient may be considered while the multiclass approach has more than two targets for example, “high”, “medium” and “low” risk patient. Data set is partitioned as training and testing dataset. Using training dataset we trained the classifier. Correctness of the classifier could be tested using test dataset. Classification is one of the most widely used methods of Data Mining in Healthcare organization. Different classification method such as decision tree, SVM and ensemble approach is used for analyzing data. Classification techniques are also used for predicting the treatment cost of healthcare services which is increases with rapid growth every year and is becoming a main concern for everyone.

II. CLASSIFICATION TECHNIQUES

Classification is the task of generalizing known structure to apply to new data. The classification task can be seen as a supervised technique where each instance belongs to a class, which is indicated by the value of a special goal attribute or simply the class attribute. The goal attribute can take on categorical values, each of them corresponding to a class. One of the major goals of a Classification algorithm is to maximize the predictive accuracy obtained by the classification model when classifying examples in the test set unseen during training Three are several techniques are used for classification some of them are.

1. Decision Tree,
2. K-Nearest Neighbor,
3. Support Vector Machines,
4. Naive Bayesian Classifiers,

5. Neural Networks.

1. Decision Trees

A Decision Tree Classifier consists of a decision tree generated on the basis of instances. A decision tree is a classifier expressed as a recursive partition of the instance space. The decision tree consists of nodes that form a rooted tree, meaning it is a directed tree with a node called "root" that has no incoming edges. All other nodes have exactly one incoming edge. A node with outgoing edges is called an internal or test node. All other nodes are called leaves (also known as terminal or decision nodes). In a decision tree, each internal node splits the instance space into two or more sub-spaces a certain discrete function of the input attributes values.The outcome of the test determines the branch traversed, and the next node visited. The class for the instance is the class of the final leaf node. The estimation criterion in the decision tree algorithm is the selection of an attribute to test a teach decision node in the tree. The goal is to select the attribute that is most useful for classifying examples. A good quantitative measure of the worth of an attribute is a statistical property called information gain that measures how well a given attribute separates the training examples according to their target classification. This measure is used to select among the candidate attributes at each step while growing the tree.

2.K-Nearest Neighbour Classifiers (KNN)

K-Nearest neighbor classifiers are based on learning by analogy. The training samples are described by n dimensional numeric attributes. Each sample represents a point in an n-dimensional space. All of the training samples are stored in an n-dimensional pattern space. When given an unknown sample, a k-nearest neighbor classifier searches the pattern space for the k training samples that are closest to the unknowns ample. Unlike decision tree induction and back propagation, nearest neighbor classifiers assign equal weight to each attribute. This may cause confusion when there are many irrelevant attributes in the data. Nearest neighbor classifiers can also be used for prediction, that is, to return areal-valued prediction for a given unknown sample. In this case, the classifier returns the average value of the real-valued associated with the k nearest neighbors of the unknown sample. The k-nearest neighbors' algorithm is amongst the simplest of all machine learning algorithms. An object is classified by a majority vote of its neighbors, with the object being assigned to the class most common amongst its k nearest neighbors. k is a positive integer, typically small. If k = 1, then the object is simply assigned to the class of its nearest neighbor. In binary (two class) classification problems, it is helpful to choose k to be an odd number as this avoids tied votes.

3. Support Vector Machine (SVM)

SVM is a very effective method for regression, classification and general pattern recognition. It is considered a good classifier because of its high generalization performance without the need to add a priori knowledge, even when the dimension of the input space is very high. It is considered a good classifier because of its high generalization performance without the need to add a priori knowledge, even when the dimension of the input space is very high. The aim of SVM is to find the best classification function to distinguish between members of the two classes in the training data. The metric for the concept of the "best" classification function can be realized geometrically. For a linearly separable dataset, a linear classification function corresponds to a separating hyper plane f(x) that passes through the middle of the two classes, separating the two. SVMs were initially developed for binary classification but it could be efficiently extended for multiclass problems.

4 Naive Bayes Classifier

Bayesian classifiers are statistical classifiers. They can predict class membership probabilities, such as the probability that a given tuple belongs to a particular class. The Naive Bayes Classifier technique is particularly suited when the dimensionality of the inputs is high. Despite its simplicity, Naive Bayes can often outperform more sophisticated classification methods. Naïve Bayes model identifies the characteristics of patients with Cardiovascular Disease . It shows the probability of each input attribute for the predictable state.

Let D be a training set of tuples and their associated class labels. As usual, each tuple is represented by an n-dimensional attribute vector, X=(x1, x2,…, xn), depicting n measurements made on the tuple from n attributes, respectively, A1, A2,.., An.. Suppose that there are m classes, C1, C2,…, Cm. Given a tuple, X, the classifier will predict that X belongs to the class having the highest posterior probability, conditioned on X. That is, the naïve Bayesian classifier predicts that tuple x belongs to the class Ci if and only if P (Ci/X)>P (Cj/X) for $1<= j<= m, j \neq i$

5. Neural Networks.

Neural Network used for classification that uses gradient descent method and based on biological nervous system having multiple interrelated processing elements known as neurons, functioning in unity to solve specific problem. Rules are extracted from the trained Neural Network (NN) help to improve interoperability of the learned network. To solve a particular problem NN used neurons which are organized processing elements. Neural Network is used for classification and pattern recognition. An NN is adaptive in nature because it changes its structure and adjusts its weight in order to minimize the error. Adjustment of weight is based on the information that flows internally and externally through network during learning phase. In NN multiclass, problem may be addressed by using multilayer feed forward technique, in which Neurons have been employed in the output layer rather using one neuron.

A main concern of the training phase is to focus on the interior weights of the neural network, which adjusted according to the transactions used in the learning process. For each training transaction, the neural network receives in addition the expected output. This concept drives us to modify the interior weights while trained neural network used to classify new images.

III. LITERATURE SURVEY

In 2010 N. Suneetha, V. Sunil Kumar proposed **"Modified Gini Index Classification: A Case Study of Heart Disease Dataset Classification for Predicting Medical Diagnosis".** When the numbers of classes are large, and the biases are increased, the Gini-based decision tree method is modified to overcome the known problems, by normalizing the Gini indexes by taking into account information about the splitting status of all attributes. Instead of using the Gini index for attribute selection ratios of Gini indexes are used and their splitting values in order to reduce the biases. They propose a tree – base approaches to analyzing multiple responses using classification algorithms comparing with a modified decision tree method for classification to overcome the known problems for the Gini-based decision tree method, normalizing the Gini indexes by taking into account information about the splitting status of all attributes. Instead of using the Gini index for attribute selection as usual, we use ratios of Gini indexes and their splitting values in order to reduce the biases [1].

In 2011 Mrs. G. Subba lakshmi and Mr. K. Ramesh proposed **"Decision Support in Heart Disease Prediction System using Naive Bayes"**. They proposed a Decision Support in Heart Disease Prediction System (DSHDPS) using data mining modeling technique, namely, Naïve Bayes. Using medical profiles such as age, sex, blood pressure and blood sugar it can predict the likelihood of patients getting a heart disease. They implement the system by using web based questionnaire application. This system helps to train nurses and medical students to diagnose patients with heart disease. Decision Support in Heart Disease Prediction System is developed using Naive Bayesian Classification technique. The system extracts hidden knowledge from a historical heart disease database. This is the most effective model to predict patients with heart disease. This model could answer complex queries, each with its own strength with respect to ease of model interpretation, access to detailed information and accuracy. DSHDPS can be further enhanced and expanded [2].

In 2012 M. Akhil jabbar, Dr. Priti Chandra, Dr. B. L. Deekshatuluc **"Heart Disease Prediction System using Associative Classification and Genetic Algorithm".** Associative classification is a recent and rewarding technique which integrates association rule mining and classification to a model for prediction and achieves maximum accuracy. Associative classifiers are especially fit to applications where maximum accuracy is desired to a model for prediction. There are many domains such as medical where the maximum accuracy of the model is desired. Heart disease is a single largest cause of death in. They proposed an efficient associative classification algorithm using genetic approach for heart disease prediction. The main motivation for using genetic algorithm in the discovery of high level prediction rules is that the discovered rules are highly comprehensible, having high predictive accuracy and of high interestingness values. They proposed a system for heart disease prediction using data mining techniques. In our feature work we plan to reduce no. of attributes and to determine the attribute which contribute towards the diagnosis of disease using genetic algorithm [3].

In 2012 Sunita Soni and O. P. Vyas proposed "**Fuzzy Weighted Associative Classifier: A Predictive Technique For Health Care Data Mining"**. They extend the problem of classification using Fuzzy Association. They proposed a new Fuzzy Weighted Associative Classifier (FWAC) that generates classification rules using Fuzzy Weighted Support and Confidence framework. They propose a theoretical model to introduce new associative classifier that takes advantage of Fuzzy Weighted Association rule mining. This work presents a new foundational approach to Fuzzy Weighted Associative Classifiers where quantitative attributes are discretized to get transformed binary database. Fuzzy WARM (FWARM) Algorithm has been proposed and redefines the weighted support and weighted confidence to adapt in Fuzzy environment. Each Fuzzy attribute is allowed to have weight depending upon their importance in predicting the class labels. In future work the proposed concept needs to be implemented to find out how much accuracy is improved by adapting the above concept. One of existing associative classifiers is to be chosen or new algorithm needs to be developed that can be integrated with Fuzzy weighted association rule miner [4].

In 2013 V. Krishnaiah , Dr. G . Narsimha, Dr. N. Subhash Chandra **"Diagnosis of Lung Cancer Prediction System using Data Mining Classification Technique".** Cancer is the most important cause of death for both men and women. They briefly examine the potential use of classification based data mining techniques such as Rule based, Decision tree, Naïve Bayes and Artificial Neural Network to massive volume of healthcare data. This is an extension of Naive Bayes to imprecise probabilities that aims at delivering robust classifications also when dealing with small or incomplete datasets. Discovery of hidden patterns and relationships often goes unexploited. Diagnosis of Lung Cancer Disease can answer complex "what if" queries which traditional decision support systems cannot. Using generic lung cancer symptoms such as age, sex, Wheezing, Shortness of breath, Pain in shoulder, chest, arm, it can predict the likelihood of patients getting alung cancer disease. Aim of the paper is to propose model for early detection and correct diagnosis of the disease which will help the doctor in saving the life of the patient. Prototype lung cancer disease prediction system is developed using data mining classification techniques. The system extracts hidden knowledge from a historical lung cancer disease database [5].

In 2013 M. Akhil Jabbar, B.L Deekshatulu & Priti Chandra **"Classification of Heart Disease using Artificial Neural Network and Feature Subset Selection".** They introduce a classification approach which uses ANN and feature subset selection for the classification of heart disease. PCA is used for preprocessing and to reduce no. Of attributes which indirectly reduces the no. of diagnosis tests which are needed to be taken by a patient. We applied our approach on Andhra Pradesh heart disease data base. By the experimental analysis they show that accuracy improved over traditional classification techniques. This system is feasible and faster and more accurate for diagnosis of heart disease. They proposed a new feature selection method for heart disease classification using ANN and various feature selection methods for Andhra Pradesh Population. They applied different feature selection methods to rank the attributes which contribute more towards classification of heart disease, which indirectly reduces the no. of diagnosis tests to be taken by a patient[6].

In 2014 Mariammal D., Jayanthi S., Dr. P. S. K. Patra **"Major Disease Diagnosis and Treatment Suggestion System using Data Mining Techniques".** They proposes a model to systematically close those gaps to discover if applying single and multiple data mining techniques to all disease treatment data can provide as reliable performance as that achieved in diagnosing disease. Using multiple data mining techniques the accuracy also improved. Disease prediction is a major challenge in the health care industry. Instead of going for a number of tests, predicting the major disease with less number of attributes is a challenging task in Data Mining. Decision Support in Disease Prediction System is developed using all the five data mining techniques. The Disease diagnosis system extracts hidden knowledge from a historical disease database. This is the most effective model to predict patients with disease[7].

In 2015 Dr. G. Rasitha Banu J. H. Bousal Jamala **"Heart Attack Prediction Using Data mining Technique"** Data mining techniques are used to analyze this rich collection of data from different perspectives and deriving useful information. They design and develop diagnosis and prediction system for heart diseases based on predictive mining. Heart disease is a term that assigns to a large number of medical conditions related to heart. These medical conditions describe the abnormal health conditions that directly influence the heart and all its parts. Heart disease is a major health problem in to days time. They analyzing the various data mining techniques introduced in recent years for heart disease prediction. Cardiovascular disease remains the biggest cause of deaths worldwide. They proposed a new unsupervised classification system is adopted for heart attack prediction at the early stage using the patient's medical record. The information in the patient record are preprocessed initially using data mining techniques and then the attributes are classified using a Fuzzy C means classifier. In the classification stage 13 attributes are given as input to the Fuzzy C Means (FCM) classifier to determine the risk of heart attack. FCM is an unsupervised clustering algorithm, which allows one piece of data to belong to two or more clusters [8].

In 2016 R. Subha K. Ananda kumar **" Study on Cardiovascular Disease Classification Using Machine Learning Approaches"**. The diagnosis of heart disease which depends in most cases on complex grouping of clinical and pathological data. Due to this complexity, the interest increased in a significant amount between the researchers and clinical professionals about the efficient and accurate heart disease prediction. In case of heart disease, the correct diagnosis in early stage is important as time is very crucial. Numeral number of tests must be requisite from the patient for detecting a disease. Machine learning based method is used to classify between healthy people and people with disease. Cardiovascular disease is the principal source of deaths widespread and the prediction of Heart Disease is significant at an untimely phase. In order to reduce number of deaths from heart diseases there has to be a quick and efficient detection technique. They present a comprehensive review for the prediction of cardiovascular disease by using machine learning based approaches [9].

In 2017 Sanjay Kumar Sen **"Predicting and Diagnosing of Heart Disease Using Machine Learning Algorithms".** In order to reduce the large scale of deaths from heart diseases, a quick and efficient detection technique is to be discovered. Data mining techniques and machine learning algorithms play a very important role in this area. The researchers accelerating their research works to develop a software with the help machine learning algorithm which can help doctors to take decision regarding both prediction and diagnosing of heart disease. The main objective of this research paper is predicting the heart disease of a patient using machine learning algorithms. Comparative study of the various performances of machine learning algorithms is done through graphical representation of the results. They carried out an experiment to find the predictive performance of different classifiers. We select four popular classifiers considering their qualitative performance for the experiment. They choose one dataset from heart available at UCI machine learning repository. Naïve base classifier is the best in performance. In order to compare the classification performance of four machine learning algorithms, classifiers are applied on same data and results are compared on the basis of misclassification and correct classification rate and according to experimental results and concluded that Naïve base classifier is the best as compared to Support Vector Machine, Decision Tree and K-Nearest Neighbor[10].

In 2018 Poornima V, Gladis D "**A novel approach for diagnosing heart disease with hybrid classifier."** They proposed an Orthogonal Local Preserving Projection (OLPP) method to reduce the function dimension of the input high-dimensional data. The dimension reduction improves the prediction rate with the help of hybrid classifier i.e. Group Search Optimization Algorithm (GSO) combine with the Leven berg-Marquardt (LM) training algorithm in the neural network. The LM training algorithm is used to solve the optimization problem and it determines the best network parameters such as weights and bias that minimizes the error. The final output of the optimization technique is combined with the performance metrics as accuracy, sensitivity, and specificity. From the result, it is observed that hybrid optimization techniques increase the accuracy of the heart disease prediction system. The proposed LM algorithm based neural network was prepared under n number of iterations and another best weight was acquired which improved the classification accuracy. The result concluded that the performance measures such as accuracy, sensitivity, and specificity achieve optimal value in GSO-LM algorithm while compared to existing approaches. The proposed method minimizes the error and predicts the heart disease with high accuracy[11].

IV. PROBLEM STATEMENT

There are several algorithms and methods have been developed to solve the problem of classification. there are some important problem that need to consider are

1. ***Speed*:** This refers to the computational costs involved in generating and using the given classifier.
2. ***Robustness*:** This is the ability of the classifier to make correct predictions given noisy data or data with missing values. Robustness is typically assessed with a series of synthetic data sets representing increasing degrees of noise and missing values.
3. ***Scalability:*** This refers to the ability to construct the classifier efficiently given large amounts of data. Scalability is typically assessed with a series of data sets of increasing size.

4. ***Interpretability:*** This refers to the level of understanding and insight that is provided y the classifier or predictor. Interpretability is subjective and therefore more difficult to assess. Decision trees and classification rules can be easy to interpret, yet their interpretability may diminish the more they become complex.

OBJECTIVES

There are several algorithms and methods have been developed to solve the problem of classification. But problem are always arises for finding a new algorithm and process for extracting knowledge for improving accuracy and efficiency.
Our major objective are-

1) Design a classification based algorithm which classify the given data set efficiently and accurately.
2) Design a classification based algorithm which classify the given data set using simple calculation and also reduce complexity.
3) Design a classification based algorithm which consider all the attributes related to the given dataset.
4) Design a classification based algorithm which work for both categorical as well as numerical of the attributes.

V. PROPOSED METHOD

In our everyday life there are several example exit where we have to analyze the historical data for example a bank loans officer needs analysis of her data in order to learn which loan applicants are “safe” and which are “risky” for the bank. Similarly for amedical researcherit is necessary to analyze breast cancer data in order to predict specific treatmentsfor a patient. These are some examples where the data analysis task required before taking any decision. Classification is a data analysis process, where a classifier is constructed to predict class, for bank loan exampleprediction class is “yes” or “no” Similarly for medical researcher prediction class is “treatment A”, “treatment B” or “treatment C” for the medical data. Proposed approach is Easy to understand because of 2 of the key reasons. Depending on the distance value be quite accurate..Features with a larger range of values can dominate the distance metric relative to features that have a smaller range, so feature scaling is important.
Let **(X_i, C_i) where i = 1, 2……., n** be data points. X_i denotes feature values & C_i denotes labels for **X_i** for each i. assuming the number of classes as ‘**C**’. **C ∈ {1, 2, 3, ……, C}** for all values of I Let X be a point for which label is not known, and we would like to find the label class using proposed approach.
Input:
Given Cardiovascular Disease database Given tuple with condition for Cardiovascular Disease
Output: Cardiovascular Disease yes/no
Method:

1. Calculate “d(X, X_i)” i =1, 2, ….., n; where d denotes the Euclidean distance between the points.
2. Arrange the calculated n Euclidean distances in non-decreasing order.
3. Select minimum distances from this sorted list.
4. Find those distance value which have minimum distance with points corresponding to this point.
5. Let m denotes the number of points belonging to the class X points i.e. $m \geq 0$
6. Now find all class level for m point. Calculate percentage values for Yes and No condition.

Illustrate With Example

Consider a simple Cardiovascular Disease patients data set with 10 records

Table 1 simple dataset of Cardiovascular Disease patients

S. No	1 (Age) 45	2 (Sex)	3 (Blood pressure) 120	4 (Cholesterol) 240	5 (ECG)	6 (Thalach value) 100	7 (Old peak) 0	8 (Slope) 0	9 (Thal) 3	CD
1	63	1	142	242	2	112	2	3	5	Y
2	62	1	150	256	2	118	2	2	4	N
3	62	1	140	241	2	115	2	3	6	Y
4	60	1	143	245	2	109	2	2	5	Y
5	64	1	148	254	2	122	2	2	5	N
6	62	0	143	276	2	121	3	3	4	Y
7	61	1	140	245	2	108	2	3	6	Y
8	63	1	141	243	2	119	1	3	5	N
9	60	1	144	242	2	115	2	2	6	N
10	62	0	145	275	2	121	3	3	4	Y

Table 2 given data of Cardiovascular Disease patient

.No	1 (Age) 45	2 (Sex)	3 (Blood pressure) 120	4 (Cholesterol) 240	5 (ECG)	6 (Thalach value) 100	7 (Old peak) 0	8 (Slope) 0	9 (Thal) 3	CD
1	63	1	142	250	2	115	2	3	4	

Euclidean distance between two objects

$$D(obj1, obj2) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + (x_4 - y_4)^2 + (x_5 - y_5)^2 + (x_6 - y_6)^2 + (x_7 - y_7)^2 + (x_8 - y_8)^2 + (x_9 - y_9)^2}$$

$$\sqrt{(63-63)^2 + (1-1)^2 + (142-142)^2 + (242-250)^2 + (2-2)^2 + (112-115)^2 + (2-2)^2 + (3-3)^2 + (5-4)^2}$$

$$\sqrt{(0)^2 + (0)^2 + (0)^2 + (8)^2 + (0)^2 + (3)^2 + (0)^2 + (0)^2 + (1)^2}$$

$$\sqrt{0+0+0+64+0+9+0+0+1}$$

$$\sqrt{74}$$

=8.6

VI. IMPLEMENTATION RESULT AND ANALYSIS

We evaluate the performance of proposed algorithm and compare it with Bayesian classification. The experiments were performed on Intel Core i3processor1GB main memory and RAM: 4GB Inbuilt HDD: 400GB OS:Windows7.The algorithms are implemented in using Dot Net Framework language version 4.0.1. Synthetic datasets are used to evaluate the performance of the algorithms. We have used SQL Server 2008 R2 for storing patient's database.

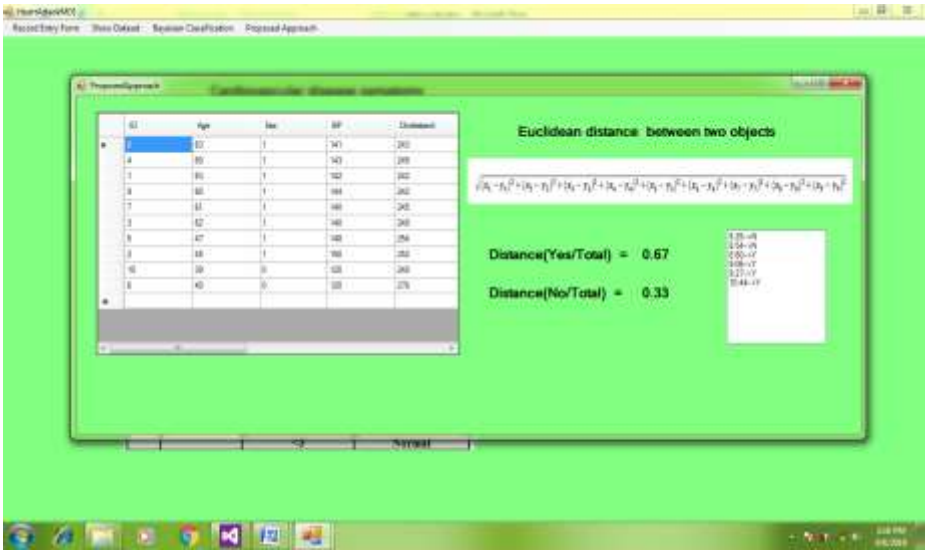

Figure 1. Implementation of proposed approach

Comparison on the basis of number of records and accuracy

For comparing the performance of the proposed approach we implement the Bayesian Classifiers and proposed approach. Our comparison is based on accuracy and number of tuples.

Table 3 Number of Record and accuracy in percentages

Number of Records	Bayesian Classifiers	Proposed Approach
100	0.4337	0.6682
200	0.4664	0.7622
500	0.4182	0.7556

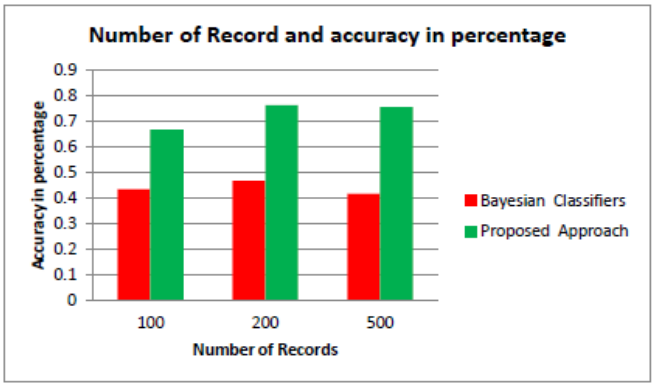

Figure 2 comparison graph Number of Record and accuracy in percentages

CONCLUSION

There are several algorithms and methods have been developed for classify Cardiovascular Disease problem accurately. But problem are always arises for finding a new algorithm and process for extracting knowledge for improving accuracy and efficiency The most popular classification methods are Artificial neural networks, Decision Tree and Support Vector Machine and Naïve Bayes Classifier. From the experiment it clear that proposed method is more accurately classify the recodes as compared to previous method. Proposed method considers all attribute given to Cardiovascular Disease condition. Proposed method is also simple to understand and calculation is easy.

LIMITATION

Proposed method has the following limitations

1. We consider only 9 attribute but there are several other like attribute induced angina, smoking, Number of major vessels coloured by fluoroscopy etc. are also responsible for Cardiovascular Disease.
2. Proposed classifiers works only for those person which have age greater than 45.
3. The proposed classifier considers only those values which satisfy the given condition for 1 and 0 but there are lots of other possibilities.

REFERENCE

1. N. Suneetha ,Ch.V.M.K.Hari and Sunil Kumar " Modified Gini Index Classification: A Case Study Of Heart Disease Dataset" (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 06, 2010, 1959-1965.
2. Mrs. G. Subbalakshmi , Mr. K. Ramesh Mr. M. Chinna Rao "Decision Support in Heart Disease Prediction System using Naive Bayes " et al. Indian Journal of Computer Science and Engineering (IJCSE) ISSN : 0976-5166 Vol. 2 No. 2 Apr-May 2011.
3. M. Akhil jabbar, Dr. Priti Chandrab , Dr. B. L Deekshatuluc "Heart Disease Prediction System using Associative Classification and Genetic Algorithm" International Conference on Emerging Trends in Electrical, Electronics and Communication Technologies-ICECIT, 2012
4. Sunita Soni and O.P.Vyas "Fuzzy Weighted Associative Classifier: A Predictive Technique For Health Care Data Mining" International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), Vol.2, No.1, February 2012.
5. V. Krishnaiah, Dr. G. Narsimha and Dr. N. Subhash Chandra" Diagnosis of Lung Cancer Prediction System Using Data Mining Classification Techniques" (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 4 (1) 2013, 39 – 45
6. M. Akhil jabbar, Dr B.L Deekshatulu and Dr Priti Chandra" Heart Disease Classification Using Nearest Neighbor Classifier With Feature Subset Selection" Computer Science and Telecommunications 2012|No.3(39).
7. Mariammal. D, Jayanthi. S, Dr. P. S. K. Patra Major Disease Diagnosis and Treatment Suggestion System Using Data Mining Techniques International Journal of Advanced Research in Computer Science & Technology IJARCST All Rights Reserved 338 Vol. 2 Issue Special 1 Jan-March 2014 ISSN: 2347 - 8446 (Online) ISSN: 2347 – 9817
8. Dr. G. Rasitha Banu, J.H.Bousal Jamala "Heart Attack Prediction Using Data mining Technique" International Journal of Modern Trends in Engineering and Research (IJMTER) Volume 02, Issue 05, [May – 2015] ISSN (Online):2349–9745 ; ISSN (Print):2393-8161
9. R. Subha K. Ananda kumar Study on Cardiovascular Disease Classification Using Machine Learning Approaches International Journal of Applied Engineering Research ISSN 0973-4562 Volume 11, Number 6 (2016) pp 4377-4380 © Research India Publications. http://www.ripublication.com
10. Sanjay Kumar Sen Predicting and Diagnosing of Heart Disease Using Machine Learning Algorithms International Journal Of Engineering And Computer Science ISSN:2319-7242 Volume 6 Issue 6 June 2017, Page No. 21623-21631 Index Copernicus value (2015): 58.10 DOI: 10.18535/ijecs/v6i6.14.
11. Poornima V, Gladis D "A novel approach for diagnosing heart disease with hybrid classifier". Biomedical Research 2018; 29 (11): 2274-2280 ISSN 0970-938X www.biomedres.info.