INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact factor: 4.295
(Volume3, Issue2)

Automated Diagnosis of Heart Disease using Random Forest Algorithm

Prof. Priya R. Patil
Marathwada Institute of Technology

Prof. S. A. Kinariwala
Marathwada Institute of Technology

Abstract: ***The accurate diagnosis of a heart diseases, is one of the most important biomedical problems whose administration is imperative. In the proposed work, decision support system is made by three data mining techniques namely Classical Random Forest, Modified Random Forest and Weighted Random Forest. The classical random forests constructs a collection of trees. In Modified Random Forest, the tree is constructed dynamically with online fitting procedure. A random forest is a substantial modification of bagging. Forest construction is based on three step process.***

1. ***Forest construction***
2. ***The polynomial fitting procedure***
3. ***The termination criterion***

In Weighted Random Forest, The Attribute Weighting Method is used for improving Accuracy of Modified Random Forest. There are Two Techniques are used in Attribute Weighting:

1. ***Averaged One-Dependence Estimators (AODE)***
2. ***Decision Tree-based Attribute Weighted Averaged One-dependence Estimator (DTWAODE).***

Keywords: Data Mining, Heart Disease.

1. INTRODUCTION

An important task of any diagnostic system is the process of attempting to determine and/or identify a possible disease or disorder and the decision reached by this process. For this purpose, machine learning algorithms are widely employed [1][2][3]. For these machine learning techniques to be useful

in medical diagnostic problems, they must be characterized by high performance, the ability to deal with missing data and with noisy data, the transparency of diagnostic knowledge, and the ability to explain decisions.

As people are generating more data everyday so there is a need for such a classifier which can classify those newly generated data accurately and efficiently. This System mainly focuses on the supervised learning technique called the Random forests for classification of data by changing the values of different hyper parameters in Random Forests Classifier

to get accurate classification results.

In the proposed system, the improvement of the random forests classification algorithm, which meets the aforementioned characteristics, is addressed. This is achieved by determining automatically the only tuning parameter of the algorithm, which is the number of base classifiers that compose the ensemble and affects its performance. The proposed method has some advantages over the aforementioned methods since it does not include any tuning parameter, which can be related to the number of base classifiers, such as

the pre selection methods, and it does not contain an overproduction phase, such as the post selection methods; thus, it does not construct base classifiers in advance that may not be needed. The proposed method determines the members of the ensemble dynamically taking into account the combination performance of the base classifiers, in contrast to the ranking methods. It does not differentiate the members of the ensemble depending on the instance being classified and on how the neighbors of this instance were classified by the initial pool, like weighted voting methods, but it creates an ensemble that works well for all the instances. the proposed system aims to construct an ensemble with optimal accuracy and correlation. The proposed system incorporates into the termination criterion both the features that an ensemble classifier should fulfill: high accuracy and low correlation. More specifically, the construction of the forest is initiated by adding a tree. As a new tree is added each time, the new accuracy and the new correlation of the forest are computed, and an online fitting procedure is applied on the curves expressing the variation of accuracy and correlation, respectively. The procedure is terminated when the differences 1) between the curve of the accuracy and the fitted curve and 2) between the curve of the correlation and the fitted one meet a specific criterion. The aforementioned characteristics permit the proposed method to be fully integrated into any diagnostic or therapeutic system since it improves random forests algorithm, thus, providing a classification algorithm of high performance, time, and computational effective that works independently of the medical problem and the nature of data, it can handle noisy or missing data, a common characteristic of medical datasets, and it does not require any human intervention since the only tuning parameter of the algorithm is determined automatically[1].

2. LITERATURE SURVEY

Data mining is an important and interesting field in Computer Science and has received a lot of attention from the research community particularly over the past decade. Data mining is the practice of automatically searching large stores of data to discover patterns and trends that go beyond simple analysis [4]. Data mining uses sophisticated mathematical algorithm to segment the data and evaluate the probability of future events. The key properties of data mining are:

- Automatic discovery of patterns
- Prediction of likely outcomes
- Creation of actionable information
- Focus on large data sets and databases

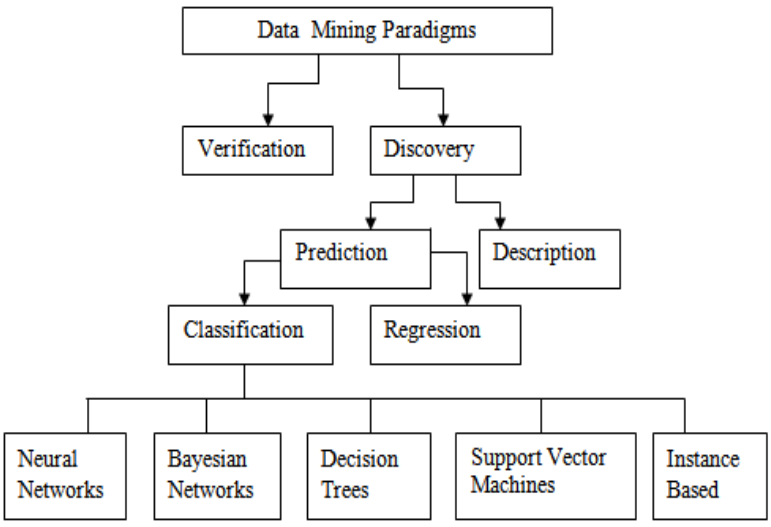


Figure 2.1: Taxonomy of Data Mining Methods

It is useful to distinguish between two main types of data mining: verification-oriented discovery-oriented. Fig 2.1 illustrates this taxonomy. Each type has its own methodology [5].Discovery methods, which automatically identify patterns in the data, involve both prediction and description methods. Description methods focus on understanding the way the underlying data operates while prediction-oriented methods aim to build a behavioral model for obtaining new and unseen samples and for predicting values of one or more variables related to the sample. Most of the discovery-oriented techniques are based on inductive learning, where a model is constructed explicitly or implicitly by generalizing from a sufficient number of training examples. Verification methods, on the other hand, evaluate

a hypothesis proposed by an external source. These methods include the most common methods of traditional statistics, like the goodness-of-fit test, the t-test of means, and analysis of variance.

2.1 Data mining Algorithm and Techniques used for Heart Disease

In Data Mining two techniques are available for the data analysis: Data Classification and Data Prediction. The Classification techniques are mainly used to predict the discrete class labels for the new observation or new data on the basis of training data set provided to the classifier algorithm, and prediction techniques generally works with the continuous valued functions [6].

2.1.1 Classification

Classification is the most commonly applied data mining technique, which employs a set of pre-classified example to develop a model that can classify the population of records at large. The data classification process involves learning and classification. In Learning the training data is analyzed by classification algorithm. In classification test data is used to estimate the accuracy of the classification rules. . Classification method makes use of mathematical techniques such as decision trees, linear programming, neural network and statistics.

A two step process is involved in classification.

• Model construction

• Model usage

Model construction describes a set of predetermined classes. Each sample is assumed to belong to a predefined class as determined by the class label attribute. The set of samples used for model construction: training set. The model is represented as classification rules, decision trees or mathematical formula. Model usage is used for classifying future and unknown objects. Estimate accuracy of the model. Accuracy rate is used to show the percentage of test set samples that are correctly classified by the prescribed model. Test set should be an independent of training set, otherwise over-fitting will occur [7].

2.1.2 Prediction

Regression technique can be adapted for prediction. Regression analysis can be used to model the relationship between one or more independent variables and dependent variables. In data mining independent variables are attributes already known and response variables are to be predicted. Unfortunately many real-world problems are not simply prediction.

2.2 Types of Classification model:

- Classification by decision tree induction [8]
- Bayesian classification
- Neural networks
- Support vector machines [9]
- Classification based on Association Rule
- Classification based on ensemble techniques [10][11]
-

2.3 Ensemble Methodology

Fig 2.6 shows Three widely used ensemble approaches, namely, boosting, bagging, and Random Forest.

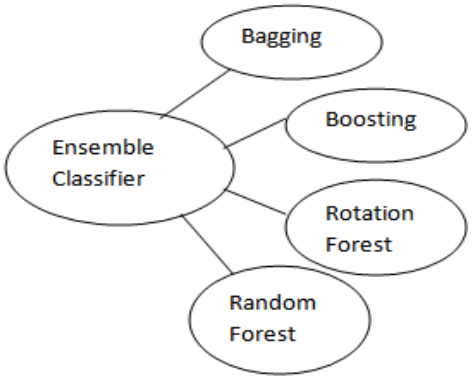


Figure 2.2: Variants of Ensemble Classifiers

2.3.1 Boosting:

Boosting is an incremental process of building a sequence of classifiers, where each classifier works on the incorrectly classified instances of the previous one in the sequence.

Bagging is based on allowing each base classifier to be trained with different random subset of training set with the goal of bringing diversity in the base classifier. One of the major problem come across in implementing this concept of boosting is improper distribution of data and method may require a large training data set. Fig 2.7 shows Boosting technique for combining multiple base classifiers whose combined performance is significantly better than that of any of the base classifiers.

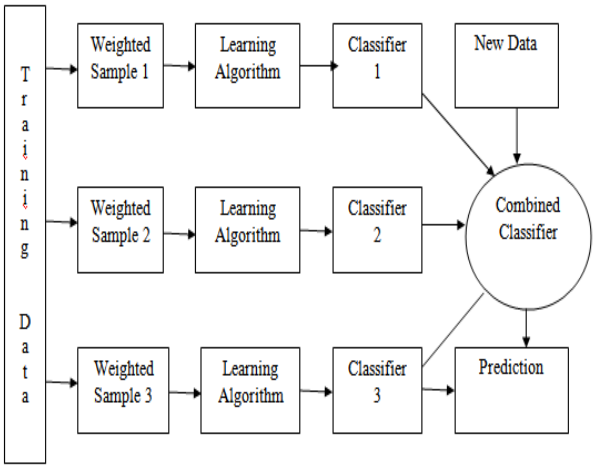


Figure 2.3: Boosting

2.3.2 Bagging

The other class of ensemble approaches is the Bootstrap Aggregating (Bagging) (Breiman, 1996a). Bagging involves building each classifier in the ensemble using a randomly drawn sample of the data, having each classifier giving an equal vote when labeling unlabeled instances. Bagging is known to be more robust than boosting against model overfitting. RF is the main representative of bagging (Breiman, 2001). Fig 2.8 shows Bagging technique which combines predictions of independent base classifiers for arriving at final prediction. Bagging works because if a learning algorithm (i.e. decision tress) is unstable a small change in training set causes large changes in the learned classifier and Bagging always.

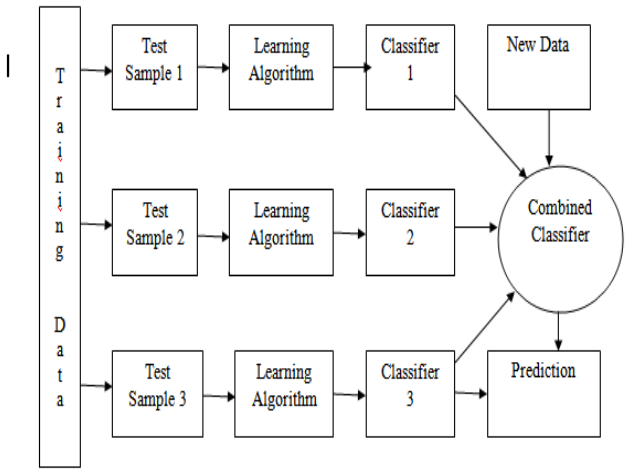


Figure 2.4: Bagging

2.3.3 Random forest:

Random Forest is essentially an ensemble of unpruned classification trees. It gives excellent performance on a number of practical problems, largely because it is not sensitive to noise in the data set, and it is not subject to overfitting. It works fast, and generally exhibits a substantial performance improvement over many other tree-based algorithms. Random forests are built by

combining the predictions of several trees, each of which is trained in isolation. There are three main choices to be made when constructing a random tree. These are

- The method for splitting the leafs.
- The type of predictor to use in each leaf.
- The method for injecting randomness into the trees.

Specifying a method for splitting leafs requires selecting the shapes of candidate splits as well as a method for evaluating the quality of each candidate. In Brieman's early work each individual tree is given an equal vote and later version of Random Forest allows weighted and unweighted voting [12]. The technique on which Random Forest ensemble is formed can be considered over following parameters:

i) Base Classifier: It describes the base classifier used in the Random Forest ensemble. Base classifier can be decision tree, Random tree, or extremely randomized tree.

ii) Split Measure: If base classifier of Random Forest is decision tree, then which split measure is found at each node of the tree to perform the splitting. To perform splitting Gini index, Info gain etc are used.

iii) Number of Passes: For building Random Forest classifier, if single pass is sufficient or multiple passes through data are needed.

iv) Combine Strategy: In Random Forest ensemble, all the base classifiers generated are used for classification. At the time of classification, how the results of individual base classifiers are combined is decided by the combine strategy.

v) Number of attributes used for base classifier generation: This parameter gives the number of how many attributes are to be used which are randomly selected from the original set of attributes at each node of the base decision tree [13]. Filter and Wrapper these are main techniques used for feature selection and extraction.

2.4 Summary of developments to RFs

Ho (1995) proposed a method to overcome a fundamental limitation on the complexity of decision tree classifiers derived with traditional methods. The proposed method uses oblique decision trees which are convenient for optimizing training set accuracy.

Amit and Geman (1997) proposed a shape recognition approach based on the joint induction of shape features and tree classifiers.

Ho (1998) proposed a method to solve the dilemma between overfitting and achieving maximum accuracy.

Breiman (2001) Proposed a RF ensemble learning method used for classification and regression.

Latinne et al. (2001) Used McNemar non-parametric test of significance to a priori limit the number of trees that will participate in majority voting and without loss in accuracy.

Robnik-Šikonja (2004) Decreased correlation between trees by using several attribute evaluation measures. Used weighted voting instead of majority voting. Tsymbal et al. (2006) Replaced majority voting with more sophisticated dynamic integration techniques: DS, DV, and DVS

Amaratunga et al. (2008) Improved the declining performance when the number of features is large and the number of truly informative features is small by using weighted random sampling instead of simple random sampling when picking features to split each node.Saffari et al. (2009) Introduced a novel online RF algorithm to remedy the limitations of the off-line algorithm. Bader-El-Den and Gaber (2012) Used genetic algorithms to boost the performance of RF.

Xu et al. (n.d.) Proposed a hybrid RF approach for classifying very high-dimensional data that outperformed the traditional RF [14][15].

2.5 Conclusion From Literature Survey

Supervised learning algorithms are commonly described as performing the task of searching through a hypothesis space to fine a suitable hypothesis that will make good predictions with a particular problem. An ensemble is a technique for combining many weak learners in an attempt to produce a strong learner. Evaluating the prediction of an ensemble typically requires more computation than evaluating the prediction of a single model, so ensemble may be thought of as a way to compensate for poor learning algorithms by performing a lot of extra computation. For example fast algorithms such as decision tree sometime have relatively poor accuracy compared to other knowledge models like neural networks. In order to overcome this problem, a large number of decision trees are generated for the same data set, and used simultaneously for prediction. Random forest is one such ensemble based method which is commonly used with decision trees.

3. SYSTEM DEVELOPMENT

Initially Machine Learning (ML) systems were developed to analyze the medical data sets. The ML system is more useful to solve medical diagnosis problems because of its good performance, the ability to deal with missing data, the ability to explain the decision and transparency of knowledge.

3.1 Random Forest:

In decision tree algorithm of Random Forest, the tree is constructed dynamically with online fitting procedure. A random forest is a substantial modification of bagging.

Each tree of Random Forest is grown can be explained as follows: Suppose training data size containing N number of records, then N records are sampled at random but with replacement, from the original data, this is known as bootstrap sample along with M number of attributes. This sample will be used for the training set for growing the tree. If there are N input variables, a number n << N is selected such that at each node, n variables are selected at random out of N and the best split on these m attributes is used to split the node. The value of m is held constant during forest growing. The decision tree is grown to the largest extent possible. A tree forms "inbag" dataset by sampling with replacement member from the training set. It is checked whether sample data is correctly classified or not using out of bag error with the help of out of bag data which is normally one third of the "inbag" data. Prediction is made by aggregating (majority vote for classification or averaging for regression) the predictions of the ensemble. Random forests include 3 main tuning parameters.

Node Size: unlike in decision trees, the number of observations in the terminal nodes of each tree of the forest can be very small. The goal is to grow trees with as little bias as possible.

Number of Trees: in practice, 500 trees is often a good choice.

Number of Predictors Sampled: the number of predictors sampled at each split would seem to be a key tuning parameter that should affect how well random forests perform. Sampling 2-5 each time is often adequate.

3.2 Random forest Algorithm

Input: Dataset

Output: Predicted class label

Step 1 : Set Number of classes = N, Number of features =M

Step 2 : Let 'm' determine the number of features at a node of decision tree, (m < M)

Step 3 : For each decision tree do

Select randomly: a subset (with replacement) of training data that represents the N classes and use the rest of data to measure the error of the tree

Step 4 : For Each node of this tree do

Select randomly: m features to determine the decision at this node and calculate the best split accordingly.

Step 5: End For

Step 6 : End For

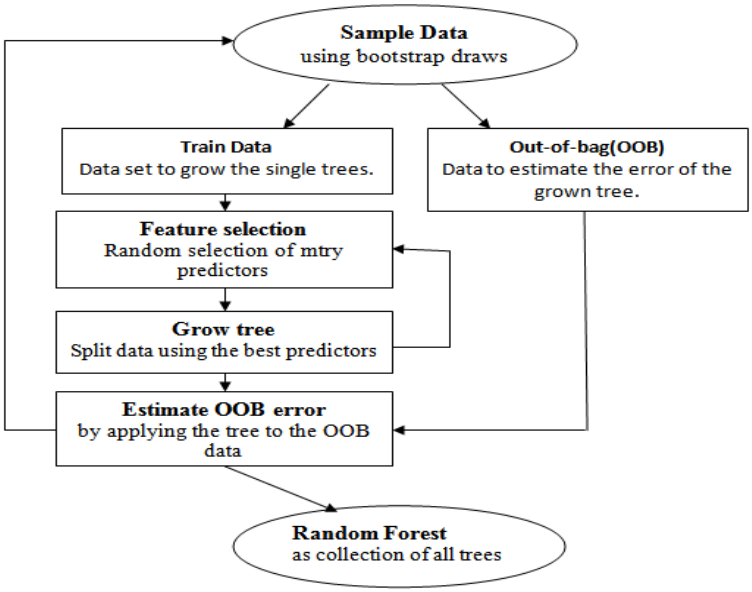


Figure 3.1: Random Forest Algorithm

3.3 Proposed Method:

3.3.1 Schematic representation of the proposed method

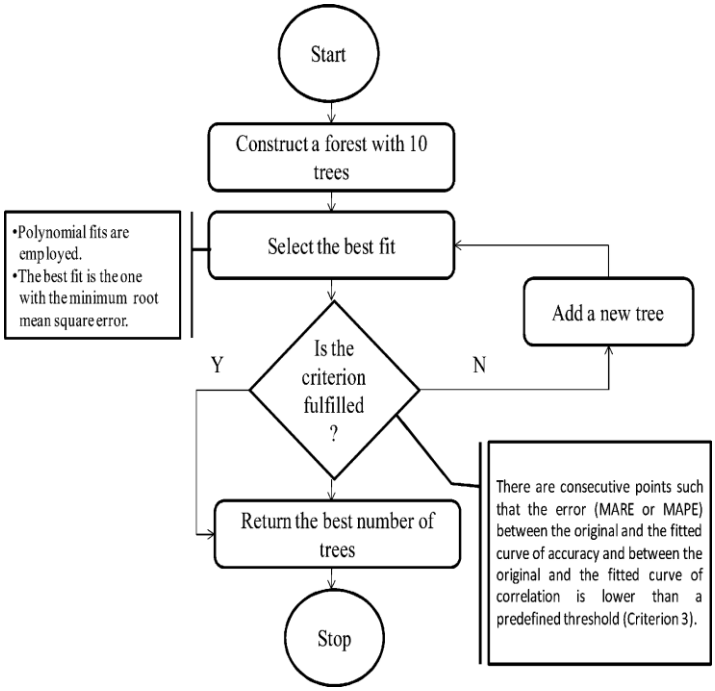


Figure 3.2: Schematic representation of the proposed method

3.4 Forest construction: Fig 3.3 shows Schematic representation of the proposed method. Random Forest is a collection of decision trees. From the training data the Random forest is constructed. In each step the tree is constructed with other data which has been selected as a best split. The forest is constructed without pruning. Forest construction is based on three step process.

i) Forest construction
ii) the polynomial fitting procedure
ii) the termination criteria.

3.4.1 Construction of the Forest: In the first step, the method constructs a forest with ten trees. For the construction of the forest, the classical random forests and some modifications of it are used.

- Random forests with ReliefF
- random forests with multiple estimators
- RK Random Forests
- RK Random Forests with multiple estimators
-

3.4.2 Fitting Procedure: After the construction of the initial forest, an iterative procedure is used. The procedure consists of three basic stages:

- Add a new tree
- Apply polynomial fits
- Select the best fit.

The polynomial fits that are employed are given by:

$$f_{n-1}(x) = p_n x^n + p_{n-1}x^{n-1} + ... + p_o, ...n = 2,...9.$$

where *x* is the data to be fitted and *pn* are the coefficients of the polynomial. The best fit is the one with the minimum rms error (the root of the average of the squares of the differences between the predicted and actual values).

3.4.3 Examination of the Termination Criterion*:* In this step, the method examines if the stopping criterion is fulfilled. For this purpose, three different criteria were tested in order to conclude to the best one.

criterion 1 : The first criterion searches for consecutive points in the fitted curve, where the difference between the fitted curve and the curve of the accuracy is greater than a predefined threshold.

criterion 2: The second criterion is based on the comparison of the two curves (original and fitted one) using measures of similarity or dissimilarity.

criterion 3: The third criterion combines accuracy and correlation since we are interested to have members of the ensemble to be not only accurate, but diverse too.

3.5 Weighed Voting Method

There are various kinds of voting systems. Two main voting systems are generally utilized, namely weighted voting and un-weighted voting. In the weighted voting system, each base classifier holds different voting power. On the other hand, in the unweight system, individual base classifier has equal weight, and the winner is the one with most number of votes. The simplest kind of ensemble is the way of aggregating a collection of prediction values base level giving different voting power for its prediction.

3.5.1 Weighted Random Forests

Assume that the data consist of a binary response variable coded 0 or 1, and *p* predictor variables collected on *N* samples. The traditional Random Forests (RF) method would build an ensemble of *ntree* classification trees to predict the outcome from the predictors, with each tree trained on a different bootstrap sample of *N* subjects, and a random subset of*mtry* predictors considered at each node of the tree. The original implementation of RF then aggregates tree-level results equally across trees. We implement the usual RF algorithm to build the trees of the forest; however, we utilize performance-based weights for tree aggregation.

3.5.2 Weighed Random Forest

To Improve the Accuracy of Modified Random Forest algorithm we use weighted voting method. All Data mining algorithm assumes attributes are independent of each other, the same may affect the accuracy of the system.

To solve this attribute independence issue and to increase accuracy, another data mining algorithm - Averaged One Dependence Estimator i. e. AODE and Decision Tree-based Attribute Weighted AODE (DTWAODE) is used.

Averaged One Dependence Estimator i. e. AODE:

It performs classification by aggregating the predictions of multiple one-dependence classifiers in which all attributes depend on the same single parent attribute as well as the class.

$$v_{AODE} = \arg\max_{v_i \in V} [\sum_{i:1 \le i \le n \wedge F(v_i) \ge m} P(x_i, v_i)] \prod_{j=1, j \ne i}^{n} P(x_j \mid x_i, v_j)]$$

Two methods for learning the weights-(attribute weighed methods)

- Gain Ratio
- Correlation –based Feature Selection

Gain Ratio: Attributes are individually ranked on the basis of separation of classes of the training data elements i.e. individual rows of the data set.

The attribute ranks, with respect to the class, are calculated using the information gain.

$$W_i = \frac{GainRatio(A_i) \times m}{\sum_{i=1}^{m} GainRatio(A_i)}$$

Correlation –based Feature Selection

Correlation-based Feature Selection (CFS) which uses the attribute quality measures described above in its heuristic evaluation function. The heart of CFS algorithm is a heuristic process for evaluating the worth or merit of a subset of features. The feature subset heuristic evaluation function mentioned in CFS is defined as follows

$$Merit_s = \frac{t\overline{r_{cf}}}{\sqrt{t + t(t-1)\overline{r_{ff}}}}$$

Where *Merits* is the heuristic "merit"of a feature subset *ssub* containing *t* features. *rcf* is the average feature-class correlation, and *rff* is the average feature-intercorrelation. Hall (2006) used this method to evaluate the importance of attribute according to the heuristic "merit". The weight of attribute *Ai* can be defined as

$$W_i = \frac{1}{\sqrt{indexrank(A_i)} + 1}$$

Where *indexrank*(*Ai*) is the index in the ranked attributes by the order that they are added to the subset during a forward selection search in CFS.

Feature selection is necessary either because it is computationally infeasible to use all available features, or because of problems of estimation when limited data samples (but a large number of features) are present.

Decision Tree-based Attribute Weighted AODE

Decision Tree-based Attribute Weighted AODE is improved AODE algorithm by eagerly assigning each attribute different weight according to the depth of the attribute in the tree.

$$v_{DTWAODE} = \arg\max_{v_i \in V}\Big[\sum_{i:1\le i\le n \wedge F(v_i)\ge m} P(x_i, v_i)\Big]\prod_{j=1, j\ne i}^{n} P(x_j \mid x_i, v_j)^{w_j}$$

In DTWAODE, to learn the weight w_j two major approaches have been widely used.

- One is to conduct a search process to find the weights that maximize the performance of the resulting model. The other way is to directly compute the weight.
- In second way weights are calculated according to the degree to which they depend on the values of other attributes.
- If the attribute does not appear in the tree, the weight is set to zero.
-

4. PERFORMANCE ANALYSIS

In this chapter implementation of systems, comparison of systems with the use of various factors, and experimental setup is performed.

4.1 Experimental setup

System is tested on University of California Irvine(UCI) machine learning dataset. UCI dataset consist of patient history of 303 cases. The Random Forest algorithm and modifications are verified with Processed clevel and datasets.

4.2 Performance Evaluation

This chapter evaluates performance of proposed algorithm and performance of proposed algorithm is analyzed with Modified Random Forest algorithm on the parameters of Precision, Recall and Accuracy.

Precision is the ratio of the number of relevant records retrieved to the total number of relevant and irrelevant records retrieved in the database.

Recall is the ratio of the number of relevant records retrieved to the total number of relevant items in the database. It is usually expressed as a percentage.

Accuracy is the portion of all relevant and irrelevant features against all features. An accuracy of 100% means that the features are exactly the same as the actual features. Precision is the portion relevant items against all features in the corpus or dataset. Recall is the portion irrelevant features against all actual features. F1 is a harmonic average of precision and recall.

Comparison among various parameters between existing and proposed method.

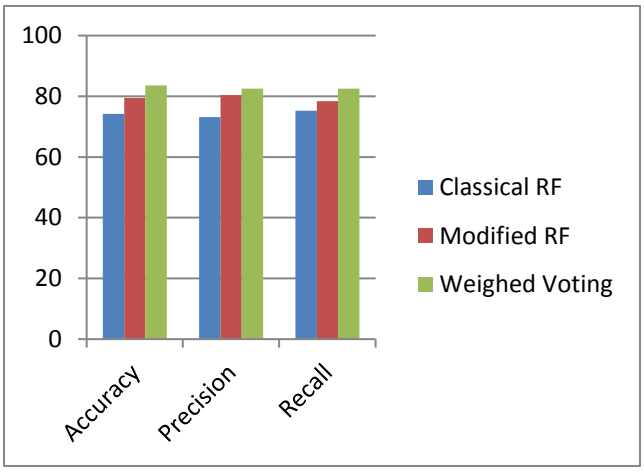
4.3 Experimental Results

Results are based on 3 measurements. Accuracy, Precision and Recall is the basic measures used in evaluating search strategies. Comparison among various parameters between different methods given below, which is based on random forest.

Table 4.1: Comparison among Various Parameters between Different Methods

	Classical RF	**Modified RF**	**Weighted RF**
Accuracy	74.19%	79.42%	83.6%
Precision	73.15%	80.46%	82.55%
Recall	75.24%	78.37%	82.55%

4.4 Graphs (Precision vs. Recall)



CONCLUSION

The dynamic construction of an ensemble of classifiers, In the random forests, was addressed in this Proposed System. We propose an automated method for the determination of the number of base classifiers in the random forests classification algorithm using an online fitting procedure.

A prototype heart disease prediction system is developed using data mining techniques with 14 input attributes .Following are the conclusion drawn from proposed system.

- The system extracts hidden Knowledge from a historical heart disease database; models are trained and validated against a test dataset.
- Classification matrix methods are used to evaluate the effectiveness of the models.
- Modified Random Forest and Weighted Random Forest both the two methods could answer complex queries, each with its qwn strength with respect to ease of model interpretation, access to detailed information and accuracy.
- when 14 attributes of UCI data set is used for Classical Random Forest then Accuracy gain is 74.19% and for Modified Random Forest it is 79.42% similarly In Weighed Random Forest it is 83.6%
- When 14 attributes of UCI data is used then precision for Classical Random Forest is 73.15%, Modified Random Forest is 80.46%and Weighed Random Forest is 82.22%.similarly Recall for Classical Random Forest is 75.24%,Modified Random Forest is 78.37%and Weighed Random Forest is 82.55% .
- Modified Random Forest and Weighed Random Forest both methods Results are easier to read and interpret. Weighed Random Forest better than Modified Random Forest as it could identify all the significant medical predictors.

5.1 Future Work

- The system can be further enhanced and expanded for the implementation of different fitting procedures, the refinement of the termination criterion in order to stop the method earlier.

- It can also incorporate other data mining techniques, such as Neural Networks and Support Vector Machines.
- The size of the data set is also need to be expanded for better accuracy results as result is dependent of quantity of data set.

REFERENCES

[1] Evanthia E. Tripoliti, "Automated Diagnosis of Diseases Based on Classification:Dynamic Determination of the Number of Trees in Random Forests Algorithm", IEEE, 2012

[2] Jehad Ali1, Rehanullah Khan2, Nasir Ahmad3, Imran Maqsood4 "Random Forests and Decision Trees" IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 5, No 3, 2012

[3] I. Kononenko, "Machine learning for medical diagnosis: History, state of the art and perspective", Artif. Intell. Med., vol. 23, no. 1, pp. 89–109, 2001.

[4] G. D. Magoulas and A. Prentza, "Machine learning in medica lapplications", Mach. Learning Appl. Berlin/Heidelberg, Germany: Springer, vol. 2049, pp. 300–307, 2001.

[5] Jiawai Han, Micheline Kamber ,"Data Mining:Concepts and Techniques",2nd Edition, Elsevier, 2006

[6] Mr. Hitesh H. Parmar, Prof GloryH. Shah, "Experimental and Comparative Analysis of Machine Learning Classifiers",ISSN:22 77 128X,volume 3.Issue 10, 2013

[7] P. Deepika, P. Vinothini, "Heart Disease Analysis And Prediction Using Various Classification Models-A survey" , ISSN-2250-1991,Volume:4,Issue:3,Mar2015

[8] Simon Bernard, Laurent Heutte, Sehastien Adam, "Forest-RK: A New Random Forest Induction Method", ICIC (2), Springer, pp.430-437, Lecture Notes in Computer Science, vol. 5227, 2009

[9] Shashikant Ghumbre, Chetan Patil, Ashok Ghatol, "Heart Disease Diagnosis using Support Vector Machine", ICCSIT, 2011

[10] Robert E. Banfield, Lawrence O. Hall, Kevin W. Bowyer, Divya Bhadoria, "A Comparison of Ensemble Creation Techniques", Fifth International Conference on Multiple Classifier System, 2004

[11] Abdelhmid Salih Mohmed Salih1 and Ajith Abraham, "Novel Ensemble Decision Support and Health Care Monitoring System", Journal of Network and Innovative Computing ISSN 2160-2174, Volume 2, 2014

[12] Sarika Pachange, Bela Joglekar, "Random Forest approach for characterizing Ensemble Classifiers", ISSN 2348-4853, 2014

[13] A Sheik Abdullah, R.R. Rajalaxmi, "A Data Mining Model for Predicting The Coronary Heart Disease using Random Forest Classifier", ICON3C, IJCA, 2012

[14] Khaled Fawagreh, Mohamed Medhat Gaber & Eyad Elyan, " Random forests: from early developments to recent advancements", ISSN, 2014

[15] Gayathri P. N. Jaisankar, "Comprehensive Study of Heart Disease Diagnosis using Data Mining and Soft Computing Techniques", IJET, vol 5 No 3, 2013

