Hybrid Approach for Prediction of Cardiovascular Disease Using Class Association Rules and MLP

K. Srinivas[1], B. Ramasubba Reddy[2], B. Kavitha Rani[1], Ravindar Mogili[3]
[1] Professor, Jyothishmathi Institute of Technology & Science, Karimnagar, TS, India
[2] Professor, SVEC, Tirupati, AP, India
[3] Associate Professor, Jyothishmathi Institute of Technology & Science, Karimnagar, TS, India

Article Info

Article history:

Received Dec 28, 2015
Revised Feb 26, 2016
Accepted Mar 10, 2016

Keyword:

Artificial neural networks
Associative classifier
Classification
CVD

ABSTRACT

In data mining classification techniques are used to predict group membership for data instances. These techniques are capable of processing a wider variety of data and the output can be easily interpreted. The aim of any classification algorithm is the design and conception of a standard model with reference to the given input. The model thus generated may be deployed to classify new examples or enable a better comprehension of available data. Medical data classification is the process of transforming descriptions of medical diagnoses and procedures used to find hidden information. Two experiments are performed to identify the prediction accuracy of Cardiovascular Disease (CVD).A hybrid approach for classification is proposed in this paper by combining the results of the associate classifier and artificial neural networks (MLP). The first experiment is performed using associative classifier to identify the key attributes which contribute more towards the decision by taking the 13 independent attributes as input. Subsequently classification using Multi Layer Perceptrons (MLP) also performed to generate the accuracy of prediction using all attributes. In the second experiment, identified key attributes using associative classifier are used as inputs for the feed forward neural networks for predicting the presence or absence of CVD.

Corresponding Author:

1. INTRODUCTION

With the ever-growing complexity in recent years, huge amounts of information in the area of medicine have been saved every day in different electronic forms such as Electronic Health Records (EHRs) and registers which is used for different purposes. Cardiovascular disease (heart disease) [1]-[3] referred as CVD is the class of diseases that involve the heart or blood vessels. It is essential to evaluate the presence or absence of cardiovascular disease (CVD) risk. Several methods are discussed by researchers to improve cardiovascular risk prediction. The data of the patients collected from different sources is stored in registers and mainly used for monitoring and analyzing health conditions. The existence of accurate epidemiological registers a basic prerequisite for monitoring and analyzing health and social conditions in the population. They are frequently used for research, evaluation, planning and other purposes by a variety of users in terms of analyzing and predicting the health status of individuals.

Data Mining aims at discovering knowledge out of data and presenting it in a form that is easily compressible to humans. It is a process that is developed to examine large amounts of data routinely collected. Data mining is most useful in an exploratory analysis scenario in which there are no predetermined notions about what will constitute an "interesting" outcome. Data mining is the search for new, valuable, and nontrivial information in large volumes of data. Best results are achieved by balancing the knowledge of human experts in describing problems and goals with the search capabilities of computers. In practice, the two primary goals of data mining tend to be classification and prediction. Prediction [4] involves using some variables or fields in the dataset to predict unknown or future values of other variables of interest. Classification [5],[6] refers to the task of analyzing a set of pre-classified data objects to learn a model (or a function) that can be used to classify unseen data object into one of several predefined classes. Description, on the other hand, focuses on finding patterns describing the data that can be interpreted by humans. For a given a collection of records (training set) each record contains a set of attributes, out of which one of the attribute is the class attribute or class variable. Other attributes are often called independent or predictor attributes (or variables). The set of examples used to learn the classification model is called the training data set. We need to find a model for class attribute as a function of the values of other attributes. Further previously unseen records should be assigned a class as accurately as possible. A test set is used to determine the accuracy of the model. Usually, the given data set is divided into training and test sets, training set used to build the model and test set used to validate it.

The basic organization of the paper is as follows: Section 2 presents the review of related works, Section 3 describes associative classifier and multilayer perceptrons (MLP), Section 4 describes the proposed hybrid classifier approach, Section 5 presents the results and discussion, and the conclusions are given in section 6.

2. RELATED WORK

When processing large databases, one faces two major obstacles such as numerous samples and high dimensionality of the feature space. For example, the documents are represented by several thousands of words; images are composed of millions of pixels, where each word or pixel is here understood as a feature. Currently, processing abilities are often not able to handle such high dimensional data, mostly due to numerical difficulties in processing, requirements in storage and transmission within a reasonable time. To reduce the computational time, it is common practice to project the data onto a smaller, latent space.

Moreover, such a space is often beneficial for further investigation due to noise reduction and desired feature extraction properties. Smaller dimensions are also advantageous when visualizing and analyzing the data. Thus, in order to extract desirable information, dimensionality reduction methods are often applied. The overall idea is to determine the coordinate system where the mapping will create low-dimensional compact representation of the data whilst maximizing the information contained within.

There are many solutions to this problem. Several techniques for dimensionality reduction have been developed which use both linear and non-linear mappings. Among them are, low-dimensional projections of the data, neural networks self-organizing maps. One can apply second order methods which use the covariance structure in determining directions. Principal Component Analysis that restricts directions to those is orthogonal. Factor Analysis which additionally allows the noise level to differ along the directions and Independent Component Analysis for which the directions are independent but not necessarily orthogonal.

Literature presents a lot of techniques for CVD using machine learning techniques. Here, we present some of the significant researches available in recent time.

Sellappan Palaniappan et al. [7] proposed a prototype Intelligent Heart Disease Prediction System (IHDPS) by means of data mining approaches such as Decision Trees, Naïve Bayes, and Neural Network. Results have revealed that each approach has its unique potency in realizing the objectives of the defined mining goals. The proposed system has the potential to forecast the possibility of patients getting a heart disease with the aid of medical profiles such as age, sex, blood pressure, and blood sugar. Also, the relevant knowledge has been established, for e.g. patterns, relationships between medical factors related to heart disease.

Carlos Ordonez et al. [8] proposed Evaluating association rules and decision trees to predict multiple target attributes and presented a detailed comparison between constrained association rules and decision trees to predict multiple target attributes. Important differences between both techniques are identified for such goal. Extensive experimental evaluation was done on a real medical data set to mine rules predicting disease on multiple heart arteries. The antecedent of association rules contains medical measurements and patient risk factors, whereas the consequent refers to the degree of disease on one artery or multiple arteries. Predictive rules found by constrained association rule mining are more abundant and have higher reliability than predictive rules induced by decision trees.

Mohammed Khalilia et al. [9] proposed predicting disease risks from highly imbalanced data using random forest, in which they presented an effective proactive approach requires an understanding of disease interdependencies and how they translate into a patient's future. Due to common genetic, molecular, environmental, and lifestyle-based individual risk factors, most diseases do not occur in isolation. Shared risk and environmental factors have similar consequences, prompting the co-occurrence of related diseases in the same patient. Therefore, a patient diagnosed for a combination of diseases and exposed to specific environmental, lifestyle and genetic risk factors may be at considerable risk of developing several other genetically and environmentally related diseases.

Anupriya et al. [10] proposed Enhanced Prediction of Heart Disease with Feature Subset Selection using Genetic Algorithm in which they used Genetic algorithms to determine the attributes which contribute more towards the diagnosis of heart ailments which indirectly reduces the number of tests which are needed to be taken by a patient. Thirteen attributes are reduced to 6 attributes using genetic search. Subsequently, three classifiers like Naive Bayes, Classification by clustering and Decision Tree are used to predict the diagnosis of patients with the same accuracy as obtained before the reduction of number of attributes.

This article is a significant extension of [11], where decision trees and MLP are combined for the first time, in the context of heart disease prediction. This work also showed that decision trees where numeric attributes are automatically split do not produce much better rules. In this new article we present a more comprehensive experimental evaluation comparing traditional classification techniques with proposed hybrid approach. We analyze how good predictive attributes are isolated by decision tree and used for further classification of CVD.

3. DEFINITIONS

Association rules bring a strong and inseparable bond between items and things that feature with regularity in a given dataset. The rules thus obtained to find deployment in real life events such as looking into the purchasing pattern of clients or preferences of shoppers for a particular product. Such observations come in useful when determining cross-marketing strategies, catalogue design and product placement. The point of departure for association rules stems from dataset mining done on a frequent basis. It is a point of common consent that strong associations correlate to frequent patterns and class-sets. Because association rules explore highly confident associations among multiple attributes, this approach may overcome some constraints introduced by decision-tree induction, which considers only one attribute at a time. The associative classification has been identified to be more precise than some traditional classifiers, like C4.5. In this paper we propose an approach for associative classification.

The overall process of the hybrid classifier is divided into two steps, such as 1) Identifying key attributes using associative classifier algorithm and 2) Prediction using MLP.

3.1. Associative Classifier

Association analysis, Classification and Clustering are three different Data Mining techniques. The aim of any classification algorithm is the design and conception of a standard model with reference to the given input. The model thus generated may then be deployed to classify new examples or enable a better comprehension of available data. Classification is a two step process consists of training phase and testing phase. The set of rules will be generated during the training phase from the training data. The test phase helps us to determine the accuracy of the classifier. Different approaches have been proposed to build accurate classifiers, such as, Support Vector Machines (SVM), naive Bayesian classification, Decision Trees based classification and so on.

The Associative classification is a new proposed classification technique [8]. It performs classification by using association rules. These rules are straight forward and simple to understand. In associative classification, the model consists of class association rules where each rule consequent is restricted to a class attribute. Recent studies show that the approach has achieved higher accuracy than traditional approaches.

The steps of associative classification technique are rule generation, classifier construction and classification and are shown in Figure 5.1. Rule generation phase will generate Class Association Rules (CARs) by using association rule mining techniques. Classifier is constructed from the rules obtained in the previous step. Classification phase assigns a class label to the given object.

We are proposing an approach to perform Classification based on Positive and Negative Association Rules which are known as Class Association Rules. Primarily Class Association Rules of the form X□ c are mined where X is a set of attributes and c is a category or a class of an object. Here X□ c need not be only positive association rule rather it may be a negative association rule. Then a classifier is constructed by considering strong rules, which is called an Associative Classifier. It takes an object as input then it attaches a class label for the given input.

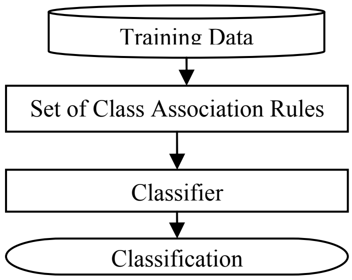

Figure 1. Associative Classifier

3.2. Multilayer Perceptrons (MLP)

A multilayer feed forward neural network is an interconnection of perceptrons in which data and calculations flow in a single direction, from the input data to the outputs. The number of layers in a neural network is the number of layers of perceptrons. MLP neural networks consist of units arranged in layers [12]-[14]. Each layer is composed of nodes and in the fully connected network considered here each node connects to every node in subsequent layers. Each MLP is composed of a minimum of three layers consisting of an input layer, one or more hidden layers and an output layer. The input layer distributes the inputs to subsequent layers. Input nodes have liner activation functions and no thresholds. Each hidden unit node and each output node have thresholds associated with them in addition to the weights. The hidden unit nodes have nonlinear activation functions and the outputs have linear activation functions. Hence, each signal feeding into a node in a subsequent layer has the original input multiplied by a weight with a threshold added and then is passed through an activation function that may be linear or nonlinear (hidden units). Neural networks allows flexibility in modeling real world complex relationships and able to estimate the posterior probability, which provides the basis for establishing classification rule and performing statistical analysis [15],[16].

4. PROPOSED HYBRID MODEL

The proposed hybrid model for the prediction of CVD comprises of two phases: (1) associative classifier and (2) prediction using MLP. The Associative classification performs classification by using association rules. Association rules are useful to identify the association exist among a given set of attributes. Generally, the number of attributes on the right side of association rules may be one or more. If consequent of an association rule contains single attribute such association rules are called class association rules. These rules are useful to identify key attributes for the prediction of CVD using MLP. The phase1 is a classification of attributes which are significant for prediction of CVD. The phase is a collection of five methods namely: PCR ( ), NCR1( ), NCR2( ), NCR3( ) and CNOPNAR( ). The method PCR( ) identifies positive association among the attributes , the methods NCR1( ), NCR2( ), NCR3( ) identifies association in the event of absence of attributes. Finally CNOPNAR ( ) method produces useful attributes for the prediction of CVD. The rules generated by the algorithms in the associative classification focus on the major contributing attributes. From this, we can identify the attributes with more contribution towards the disease identification and attribute association for the development of disease in human body. Out of all the generated rules the important attributes for prediction of disease are selected.

Further the selected attributes from the previous step will be considered as input to the multilayer perceptrons (MLP). MLP consists of 3 layers in the model. The first layer is an input layer, second is hidden layer and the third layer is an output layer which is used for predicting the presence or absence of the CVD. The training and testing of MLP is carried out using PASW18 (Predictive Analytics Software).

4.1. Algorithm for Association among the Attributes

In this section, we present algorithm for an Associative Classifier called CPNAR (Classification based on Positive and Negative Association Rules) to decide the classes to which the new objects belongs to. Database (DB), minimum support (ms) and minimum confidence (mc) are the inputs to the algorithm CPNAR. It consists of 5 procedures namely PCR ( ), NCR1 ( ), NCR2 ( ), NCR3 ( ) and CNOPNAR ( ).

PCR ( ) generates Positive Class Association Rules of the form $X \rightarrow c$,

NCR1 ( ) generates Negative Class Association Rules of the form $\neg X \rightarrow c$,

NCR2 ( ) generates Negative Class Association Rules of the form $\neg XY \rightarrow c$,

NCR3( )generates Negative Class Association Rules of the form $\neg X \neg Y \rightarrow c$, and CNOPNAR ( ) is the actual Associative Classifier.

Apriori-based implementation is used in this work as Apriori is simple and efficient to generate association rules. To determine the validity of ARs, the support and confidence measures have been used.

Algorithm: CPNAR ( )

```
{
   Call Procedure PCR ( )
   Call Procedure NCR1 ( )
   Call Procedure NCR2 ( )
   Call Procedure NCR3 ( )
   Call Procedure CNOPNAR ( )
}
```

Procedure PCR ( )

```
{
 Pcr = Φ
Find L1- Frequent 1-itemsets
L = L ∪ L1
 for (  K = 2; L K-1 ≠ Ø; K++)
{
     /* Generating PCK */
       for each l1,l2 ∈ LK-1
         {
              if(l1[1]=l2[1]^………………l1[k-2]=l2[k-2]^l1[k-1]<l2[k-1])
              PCK = PCK ∪ {{l1 [1]…….l1 [k-2],l1[k-1],l2[k-1]}
         }
     /* Pruning using Apriori property*/
     for each (K-1)- subsets s of I ∈ PCK
         {
             if s ∉  L K-1
                  PCK = PCK – {I}
         }
    /*Pruning using Support Count*/
    Scan the database and find  supp(I) for all I ∈ PCK
    for each I ∈ PCK
         {
             if supp (I) ≥ ms
             LK = LK  ∪ {I}
        }
        L =L ∪ LK
    }
 }
 /* Generating Positive Class Association Rules of the form I(=XY) → c*/
 for each I(=XY) ∈ L
 {
       for each  c ∈ C
      {
          if conf (I →c) ≥ mc
          Pcr = Pcr ∪ {I → c}
      }
 }
}
```

Initially the set PCR (Positive Classification Rules) is empty. Initially it finds L1-frequent 1-itemsets. In the above algorithm Line 4-25 generates all positive frequent itemsets. Line 6-10 generates positive candidate itemsets (PCK). The generated candidate itemsets are pruned using Apriori Principle (Line 12-16) and support count (Line 18-23). Line 29-34 generates positive class association rules by considering positive frequent itemset (I) and a class label (c). If the confidence of the aforementioned rule is more than minimum confidence (mc) then it is considered as a valid Positive Class Association Rule and will be included in PCR otherwise it will be discarded.

Procedure NCR1 ( )

```
{
Ncr1 = Φ /* Negative Class Association Rule set 1*/
NL = Φ /* Negative Frequent Itemset*/
for each I ∈ L
 {
     if 1-supp(I) ≥ ms
     NL=NL ∪ { I }
   }
/* Generating  Negative  Class Association Rules of the from  ¬I(=XY) → c*/
  for each I ∈ NL
   {
     for each  c ∈ C
       {
          if conf( ¬I→c ) ≥ mc
          Ncr1 = Ncr1 ∪ {¬I→c}
       }
   }
}
```

Initially Ncr1 and NL set to Φ. It generates negative class association rules of the form ¬I (=XY) → c. First it generates negative frequent itemsets from positive frequent itemsets generated in the previous procedure PCR by finding 1-supp (I). If it is more than minimum support (ms) then it is included in NL. Otherwise, it is not included in NL. It is shown in line 3 - 7. For each I in NL and a class label c it generates a rule ¬I→ c. If the confidence is more than mc then it is included in Ncr1. Otherwise it is discarded. It is shown in line 9 -15.

Procedure NCR2 ( )

```
{
Ncr2 = Φ  /* Negative Class Association Rule set 2*/
NNL = Φ   /*Negative Negative Frequent itemset*/
NNC2  = Set of negative canditate itemsets of the form ¬{i1 }¬{i2 } where |i1 , i2 ∈ L1 and i1 ≠ i2
for (K = 2; NNCK ≠ Ø; K + +)
{
   for all I = ¬X¬Y ∈  NNCK
    {
     if supp(I) ≥ ms
       NNLK = NNLK ∪ { I }
     else
     {
        for all i ∉  XY
          {
            /* Generating Candidates */
               Cand ={¬ ( X ∪ { i }) ¬Y,   ¬X ( ¬ Y  ∪ { i } )}
            /* Pruning Cand*/
             for each item in Cand
               {
        if (X{i} ∉ L or ¬X1 ¬Y1 ∈ NNL  where X1 ⊆ X{i} and Y1 ⊆ Y)
                Cand= Cand – {XY {i}}
                 NNCK+1=NNCK+1  ∪ C and
              }
        }
   }
}
/* Generating  Negative  class association  Rules of the from I(= ¬X¬Y) → c*/
for each I ∈ NNL
   {
         for each  c ∈ C
           {
```

```
                  if conf(I→c) ≥ mc
                  Ncr2 = Ncr2 ∪ { I→c }
              }
          }
      }
```

Initially NCR2 (Negative Class Association Rules of second type) and NNL (Negative Negative Frequent Itemset) set to Φ. Line 3 generates Negative Negative candidate2 itemset. It produces candidate 2-itemset by taking two positive frequent 1- itemsets from L1 and then applied negation to each item. Line 4-25 generates all valid Negative Negative Frequent itemsets. Line 8-9 generates negative negative frequent itemsets. Line 15 generates negative negative candidate itemsets for the next level. It generates negative negative candidate itemsets by adding a positive frequent 1-item, i.e., by adding frequent itemset i to an infrequent itemset ¬ X ¬Y. We will obtain two negative candidate itemsets ¬ X{i} ¬Y, ¬ X ¬Y{i}. Line 17-21 performs pruning on the generated candidate itemsets. Line 27-34 generates Negative Class Association Rules for Negative Negative Frequent itemsets obtained in the previous steps.

Procedure NCR3 ( )

```
{
 Ncr3 = Φ /* Negative Class Association Rule Set3*/
NPL= Φ  /* Negative and Positive Frequent Itemset*/
NPC1,1 = Set of negative itemsets of the form ¬{i1 }{i2 } where i1 , i2 ∈ L1   and  i1 ≠ i2
/* Negative and Positive Candidate itemset*/
 for (K = 1; NCK,1 ≠ Ø; K + +)
{
   for (P = 1; NCK,P ≠ Ø; P + +)
     {
         for all I ∈ NCK,P
           {
                 if supp(I) ≥ ms
                 NPLK,P = NPLK,P ∪ { I }
            }
         /*Generating Candidates*/
     for all  I1 , I2 ∈ NPLK,P /* I1and I2 are joinable itemsets*/
      {
       X = I1 .negative, Y = I1 .positive ∪ I2 .positive
      I = ¬XY
    if ((∄X1 ⊂ X)(supp(¬X1 Y ) ≥ ms) and  (∄Y1 ⊂ Y )(supp(¬XY1 ) < ms))
           NPCK,P+1 = NPCK,P+1 ∪ { I }
      }
  }
   for all X ∈ LK+1 , i ∈ L1
     {
           if (  ∄X1 ⊂ X)(¬X1 {i} ∈ NPL)
                NPCK+1,1 = NPCK+1,1 ∪ ¬X{i}
      }
  }
/*  Generating  Negative  Class Association Rules of the from  ¬XY→ c */
 for each I ∈ NPL
  {
      for each  c ∈ C
       {
          if conf ( I → c ) ≥  mc
            Ncr3 = Ncr3  ∪ {I→c}
       }
   }
}
```

/* I1 and I2 are joinable if I1≠ I2 , I1 .negative = I2 .negative, I1 .positive and I2.positive share the same k − 1 items, and I1 .positive ∪ I2 .positive ∈ L(P1)p+1 */

The association rules obtained by procedures PCR( ), NCR1( ), NCR2( ) and NCR3( ) are ordered based on the support and confidence values. It represents an actual classifier. The process of classification is looked into for the set of rules for those classes that are relevant to the object presented for classification.

Assignment of class label to the new object will be explained by the following procedure called CNOPNAR (Classification of New Object based on Positive and Negative Association Rules). The inputs to the CNOPNAR are Associative Classifier, Confidence Margin (CM) and a new object O to be classified. It produces a category attached to the new object.

Procedure CNOPNAR ( )

```
{
S = Φ   /* set of rules that match o*/
  for each rule R in  the sorted set of rules
  {
      if ( R ⊂ O)       /* O is an object which is to be Classified*/
       {count++}
      if(count = = 1)
         fr.conf  =  R.conf /* keep the first rule confidence*/
           S = S ∪ R
           else if( R.conf > fr.conf - CM)
                  S= S ∪ R
    }
   S= S1 ∪ S2 ∪ S3 ……………… Sn
  for each subset S1,S2, . . . . Sn
  {
```

$$ConfidenceScore_{\,i} = \frac{Sum\ of\ the\ Confidences\ of\ Rules\ in\ Si}{Total\ number\ of\ Rules}$$

```
  }
      O = cj, where   cj = max { ConfidenceScore1, ....... ConfidenceScoren }
  }
```

In the above algorithm CNOPNAR, the lines 2-11 selects a set of appropriate rules within a Confidence Margin (CM). The selected rules are in the interval [R.confidence – CM , fr.confidence]. The prediction of rules starts at line 11. In line 12, the set of selected rules is divided based on the classes. In line 12-14, the groups have been arranged based on the average confidence per class. In line 17 the classification is made by assigning a class that has the highest Confidence Score to the new object.

5. EXPERIMENTS AND RESULTS

To illustrate the introduced approach we use the dataset from the UCI Repository of Machine Learning Databases [17]. Cleveland heart foundation dataset with 14 attributes and 303 data items, Hungarian dataset with 13 attributes and 294 data items, Switzerland dataset with 13 attributes and 123 data items used in our experiments. This data consists of multivariate attributes out of which the last 1 attribute is a dependent attribute and the remaining attributes are independent attributes. The dependent attribute shows whether the CVD is present or absent in the patient data. The dependent attribute is transformed into binary data such as presence of disease with a value 1 or absence of disease with a value 0. Class distributions are 54% heart disease absent and 46% heart disease present.

We developed code in MATLAB and experiments are conducted. We compared our proposed method with the popular algorithms genfis2 [18], Fuzzy + decision tree classifier [19], clinical decision support system [20], a hybrid practical swarm optimization (PSO) based fuzzy expert system [21] and Prediction of Cardiovascular Disease-A Hybrid Approach [22]-[24] on the real datasets called Cleveland Dataset, Hungarian dataset and Switzerland dataset. The results are summarized as:

Table 2. Results summary

		Cleveland	Hungarian	Switzerland
Genfis2	**Sensitivity**	20.2	66.0	76.5
	Specificity	68.6	22.2	44.7
	Accuracy	39.2	39.5	62.3
Fuzzy + decision tree classifier	**Sensitivity**	50.549	55.974	59.375
	Specificity	62.386	62.386	55.737
	Accuracy	51.793	36.0	62.5
Clinical decision support system	**Sensitivity**	52.066	53.766	61.125
	Specificity	44.875	47.350	74.256
	Accuracy	46.483	42.417	58.183
A hybrid particle swarm optimization based fuzzy expert system	**Sensitivity**	62.00	60	95.0
	Specificity	61.09	60.369	56.37
	accuracy	53.93	45.09	63.45
proposed technique	**Sensitivity**	66.90	72.81	95.93
	Specificity	63.87	64.69	58.37
	Accuracy	84.967	71.88	77.59

The comparison of results over Cleveland Dataset, Hungarian dataset and Switzerland dataset using genfis2, Fuzzy + decision tree classifier, clinical decision support system, a hybrid practical swarm optimization based fuzzy expert system and proposed technique are show below as bar charts. From the graphs it is evident that the proposed method performs very well.

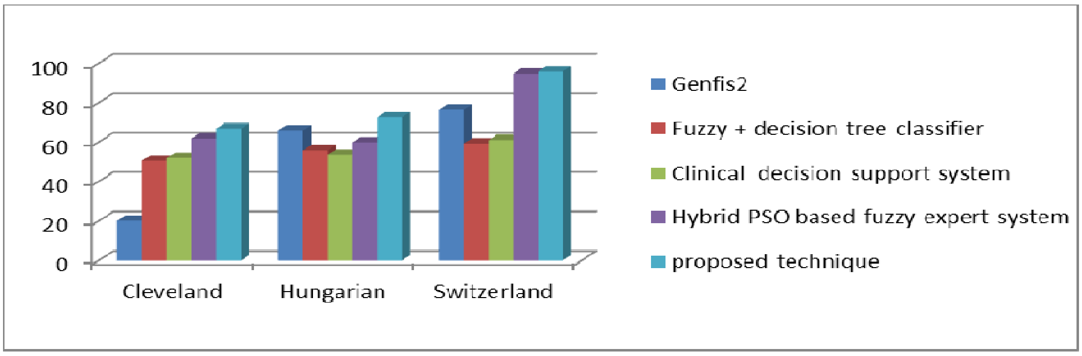

Figure 2. Comparison of Sensitivity

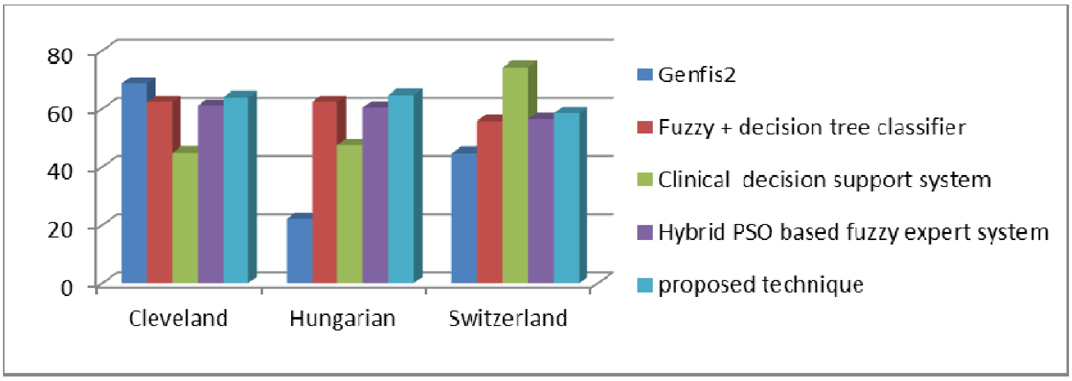

Figure 3. Comparison of Specificity

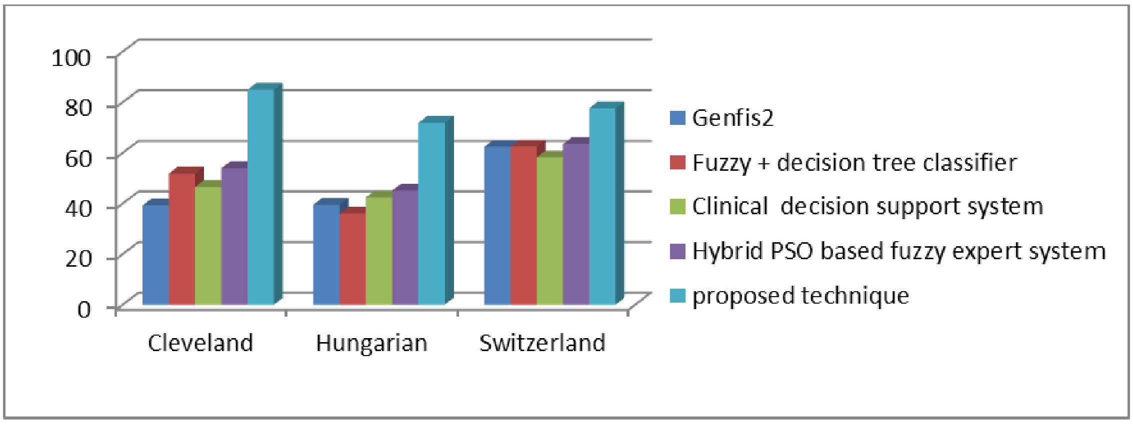

Figure 4. Comparison of Accuracy

6. CONCLUSION

The accuracy of proposed model is close to the results obtained by MLP with total attributes. We have introduced a new classifier using, associative classifier and MLP by blending the outputs of associative classifier as the inputs of MLP. The associative classifier does not provide a mathematical model for the classification of future objects, whereas the proposed method provides a mathematical model for the classification with limited number of variables. Two important steps utilized in rule generation process are selecting the important attributes using association rules and classification using MLP. Finally, the experimentation is carried out using the Cleveland, Hungarian and Switzerland datasets and the performance was analyzed with sensitivity, specificity and accuracy. From the above study it is observed that decision tree correctly classifies 77%, MLP with all 13 attributes correctly classifies 82% and the proposed method correctly classifies 85% with limited number of attributes. The proposed method outperforms than decision tree and other existing methods.

REFERENCES

[1] C. Becquet, *et al.*, "Strong associationrule mining for large scale gene expression data analysis: a case study on human SAGE data," *Genom Biol,* vol/issue: 3(12), 2002.
[2] "Cardiovascular diseases in India Challenges and way ahead," *International Heart Protection Summit*, 2011.
[3] B. Shah and P. Mathur, "Surveillance of cardiovascular disease risk factors in India: The need & scope," *Indian J Med Res*, vol. 132, pp. 634-642, 2010.
[4] F. L. Duff, *et al.*, "Predicting Survival Causes After Out of Hospital Cardiac Arrest using Data Mining Method," MEDINFO, Fieschi et al. (Eds)Amsterdam, IOS Press, 2004.
[5] A. Freitas, "Understanding the crucial differences between classification and association rules – a position paper," *SIGKDD Explorations,* vol/issue: 2(1), pp. 65–69, 2000.
[6] S. Özekes and A. Y. Çamurcu, "Classification and prediction in a data Mining application," *Journal of Marmara for Pure and Applied Sciences*, vol. 18, pp. 159-174, 2002.
[7] S. Palaniappan and R. Awang, "Intelligent Heart Disease Prediction System," *Computer Systems and Applications, AICCSA 2008. IEEE/ACS International Conference,* 2008.
[8] C. Ordonez and K. Zhao, "Evaluating association rules and decision trees to predict multiple target attributes," *Intelligent Data Analysis*, vol. 15, pp. 173–192, 2011.
[9] M. Khalilia, *et al*., "Predicting disease risks from highly imbalanced data using random forest," *BMC Medical Informatics and Decision Making,* vol/issue: 11(51), 2011. doi:10.1186/1472-6947-11-51.
[10] E. Anupriya, *et al.*, "Enhanced Prediction of Heart Disease with Feature Subset Selection using Genetic Algorithm," *International Journal of Engineering Science and Technology,* vol/issue: 2(10), pp. 5370-5376, 2010.
[11] K. Srinivas, *et al.*, "Analysis of Coronary Heart Disease and Prediction of Heart Attack in Coal Mining Regions Using Data Mining Techniques," *The 5th IEEE International Conference on Computer Science & Education, Hefei, China.*
[12] S. Haykin, "Neural Networks: A Comprehensive Foundation," 2nd Edn., New Jersey, Prentice-Hall, 1999.
[13] J. M. Nazzal, *et al.*, "Multilayer Perceptron Neural Network (MLPs)For Analyzing the Properties of Jordan Oil Shale," *World Applied Sciences Journal*, vol/issue: 5(5), pp. 546-552, 2008.
[14] B. K. Wong and Y. Selvi, "Neural network applications in finance: A review and analysis of literature (1990-1996)," *Information & Management*, vol. 34, pp. 129-139,1998.
[15] X. Wu, *et al*., "Top 10 algorithms in data mining," *Knowl Inf Syst*, vol. 14, pp. 1–37, 2008. DOI 10.1007/s10115-007-0114-2.

[16] J. Zirilli, "Financial prediction using neural networks," *International Thompson Computer Press,* 1997.
[17] http://archive.ics.uci.edu/ml/datasets/Heart+Disease
[18] A. Cruz and N. C. E. Mestrado, "ANFIS: Adaptive neuro-fuzzy inference systems," *IM, UFRJ, Mestrado NCE,* 2009.
[19] Y. Yuan and M. J. Shaw, "Induction of fuzzy decision trees," *Fuzzy Sets and systems,* vol/issue: 69(2), pp. 125-139, 1995.
[20] J. H. Eom, *et al.*, "AptaCDSS-E: A classifier ensemble-based clinical decision support system for cardiovascular disease level prediction," *Expert Systems with Applications,* vol/issue: 34(4), pp. 2465-2479, 2008.
[21] S. Muthukaruppan and M. J. Er, "A hybrid particle swarm optimization based fuzzy expert system for the diagnosis of coronary artery disease," *Expert Systems with Applications,* vol/issue: 39(14), pp. 11657-11665, 2012.
[22] P. K. Anooj, "Clinical decision support system: Risk level prediction of heart disease using weighted fuzzy rules," *Journal of King Saud University-Computer and Information Sciences,* vol/issue: 24(1), pp. 27-40, 2012.
[23] A. Draidi and D. Labed, "A Neuro-fuzzy Approach for Predicting Load Peak Profile," *International Journal of Electrical and Computer Engineering,* vol/issue: 5(6), 2015.
[24] M. Abdar, *et al.*, "Comparing Performance of Data Mining Algorithms in Prediction Heart Diseases," *International Journal of Electrical and Computer Engineering,* vol/issue: 5(6), 2015.