Research Journal of Applied Sciences, Engineering and Technology 11(9): 994-1002, 2015
DOI: 10.19026/rjaset.11.2140
ISSN: 2040-7459; e-ISSN: 2040-7467

Submitted: May 25, 2015 Accepted: June 22, 2015 Published: November 25, 2015

Research Article

A Review on the Performance of Neural Network Classifier in Health Care Diagnostic System

[1]M. Nandhini, [2]S.N. Sivanandam and [3]M. Rajalakshmi
[1]Computer Science and Engineering, PSG College of Technology,
[2]Computer Science and Engineering, Karpagam College of Engineering,
[3]Computer Science and Engineering and Information Technology, Coimbatore Institute of Technology, Coimbatore, India

Abstract: In recent times, the classification systems for diagnosing the patient's disease have received its attention. Neural network is well known classification technique widely applied to health care systems. Health care data diagnosis is a significant task that needs to be accomplished accurately and efficiently. Disease prediction based on patient's symptoms may lead to wrong assumptions. This study aims in implementing a neural network based Health Care Diagnostic System (HCDS) to predict the likelihood of patient getting a disease based on medical factors. In this study, Multi Layer Perceptron Neural Network (MLPNN) with Back Propagation algorithm (BP) is used to build the HCDS. To improve the accuracy of the diagnosis, MLPNN is constructed using the reduced set of significant Class Association Rules (CAR's) as training instances instead of datasets. Genetic Algorithm (GA) is employed to generate the reduced set of CAR's from the health care datasets. Experiments were conducted using six health care datasets from UCI machine learning repository. Based on the experiments, the combination of MLPNN with BP using significant CAR's as training instances yields promising results in terms of classifier accuracy and training time.

Keywords: Back Propagation (BP), classification, Class Association Rule (CAR), Genetic Algorithm (GA), Multi Layer Perceptron Neural Network (MLPNN), neural network

INTRODUCTION

In health care systems, correct classification and disease diagnosis are significant parameters which require much experience and domain knowledge. Data mining tasks such as Association Rule Mining (ARM) and classification are commonly used for discovering hidden knowledge that is useful for patient's disease diagnosis. An association rule discovers the inter-relationship among the patient's symptoms from health care datasets. A CAR is a kind of association rule which have a predetermined class label as consequent, used to classify unknown tuples. In literature, Apriori (Agrawal and Srikant, 1994) and Frequent Pattern (FP) Growth (Han *et al*., 2000) are the popular association rule mining algorithms used to extract association rules. The major drawback of ARM is generation of large volume of rules corresponding to a specific attribute. Extracting useful knowledge from large volume of rules is a difficult task, because the relevant information for mining may be hidden within the rules. Support and Confidence are primary objective measures commonly used to generate interesting rules. In this study, GA is used to generate the reduced set of significant CAR's from the voluminous rules obtained using basic Apriori algorithm.

Classification is a predictive data mining task used to determine a model using historical data to predict some response of interest. In this study, a classification system for health care datasets is designed using MLPNN with BP (MLPNN-BP). Generally, the original dataset is taken as training instances for training the MLPNN whereas in this study, significant CAR's are taken as training instances. Mostly health care systems deal with high numerosity and dimensional datasets. Existence of high dimensions and tuples influences the accuracy and training time. To increase the classifier accuracy and decrease training time, the reduced set of significant CAR's obtained using GA is considered as training instances for MLPNN-BP. This study reviews the impact of training the neural network using the CAR's over the MLPNN-BP. From the experiments, it is found that the MLPNN-BP using reduced set of significant CAR's yields better results in terms of classifier accuracy and training time.

Corresponding Author: M. Nandhini, Computer Science and Engineering, PSG College of Technology, Coimbatore, India

LITERATURE REVIEW

Neural network: Health care data diagnosis is a potential application that exploits classification techniques. Artificial Neural Network (ANN) or simply Neural Network (NN) is the prominent classifier for health care data diagnosis. It is one of the promising alternatives to traditional classifiers for solving realistic classification problem in radiology, urology, cardiology and oncology diagnosis and also provides better results in disease classification. NN classifiers assist the physicians to recommend medicine for the patients at early stage and they are ideal in predicting the diseases. Health care data mining has great potential for exploring hidden patterns in the health care data sets. It has been applied in variety of real time applications which accomplish tasks such as pattern recognition, image processing, language processing, control systems etc. NN offers a lot of advantages over other classifiers, as it requires very minimal training to identify all possible relationships between class and non-class attributes and the availability of many algorithms for training the network to detect complex nonlinear relationships exists between attributes. Because of its advantages, it is one of the widely used classifier for health data diagnosis. Though neural network have been successfully applied for many classification problems such as insolvency prediction, script recognition, language recognition, error detection, this study discusses the issues related to NN in health care data diagnosis.

McCulloch and Pitts (1943) pioneered the concept of neural computing. Hebb (1949) proposed the first learning rule for neural network classification. Later, Rosenblatt (1958) introduced perceptron consisting of a single neuron with adjustable synaptic weights and a threshold activation function. Rumelhart *et al*. (1986) Proposed BP algorithm to train the MLPNN. BP was built on high mathematical foundation; it increases the learning ability of the NN to discriminate non-linearly separable classes. In literature, there are several neural network architectures proposed, one which is most popular is MLPNN with BP. MLPs are feed-forward networks of simple processing units with at least one "hidden" layer, where each processing unit is similar to perceptron. Baxt (1990) applied a non linear neural network using back propagation to diagnose the acute myocardial infarction (coronary occlusion). In his study, NN resulted in 80% accuracy. Wu *et al*. (1993) designed a three-layer; feed-forward neural network with a back-propagation algorithm to interpret mammograms using the features extracted by experienced radiologists. It was concluded that NN found to be useful in mammographic decision-making task of distinguishing between benign and malignant lesions. Tu (1996) compares the performance of NN and logistic regression, the commonly used method for developing predictive models for dichotomous outcomes in medicine. Setiono (1996) proposed a new algorithm for pruning neural network to extract rules for breast cancer diagnosis. The rules extracted from the network achieve an accuracy of more than 95% on the training dataset as well as test dataset. Khan *et al*. (2001) proposed a method for classifying cancers into four diagnostic cancers based on their gene expression signatures using Artificial Neural Networks (ANNs). Yan *et al*. (2006) proposed feed forward neural network, the simplest type of artificial neural network to support the diagnosis of heart diseases. The three different assessment methods such as cross validation, holdout and bootstrapping were applied to assess the neural network system. In their work, MLP had obtained a very high diagnosis accuracy of more than 90%. Lisboa and Taktak (2006) were conducted a systematic review to assess the benefit of ANNs as decision making tools in the field of cancer. Kumar and Abhishek (2012) found that MLPNN- BP using two hidden layers to be best model for kidney stone diagnosis in urology.

Predictive accuracy of various classifiers was compared in Ture *et al*. (2005). In that study, based on their analysis, it was concluded that neural network classifiers such as MLP and RBF yields better results than other classifiers in predicting hypertension. Thus, the contribution of neural networks plays a vital role in the domain of health care.

Genetic algorithm: Genetic Algorithm (GA) is an inherent method motivated by nature used to provide solutions to optimization problems. It is one of the evolutionary algorithms used to produce solutions to optimization problems using techniques inspired by biological evolution process, such as selection, mutation, crossover and inheritance. In the late 1980's, combination of GA and NN has emerged as a powerful field of research. Combination of GA and NN demonstrates powerful classification problem solving capability. Generally, GA's can be used to improve the accuracy of the classification ability of the NN. Though the success of the NN depends on the number and quality of parameters set in each layer, there are no proper rules formally defined to set these parameters. Hence, GA can be used as an effective tool to determine NN parameters for optimizing the NN performance. In literature, GA's were combined with NN's in numerous ways for performance optimization. Vinterbo and Ohno-Machado (2000) defined GA to determine minimal sets of disorders covering all symptoms for diagnosing multi disorders of a patient. Karegowda *et al*. (2011) used GA to initialize and optimize the connection weights of BP network. It was found that GA-optimized BP network outperformed the BP network without GA optimization. Mantzaris *et al*. (2011), used GA to determine the number of diagnostic factors required for training NN. Use of GA minimizes the number of nodes in input and hidden layer thus minimizing the Mean Square Error (MSE) of the NN. Elveren and Yumuşak (2011), trained MLPN using GA for tuberculosis diagnosis. It was presented that combination of MLPN and GA yields 94.88%

accuracy. Ahmad *et al*. (2013) proposed a novel segmented multi-chromosome crossover operation for gene offspring's to inherit gene segments from multiple parent chromosomes. It was concluded that, enhanced GA with NN yields better results in terms of average accuracy.

METHODOLOGY

Proposed methodology: The proposed methodology incorporates four major phases: In first phase, all possible class rules are extracted from the preprocessed datasets using Apriori algorithm. Second phase involves generation of reduced set of significant CAR's using GA. In third, MLPNN-BP is trained using significant CAR's. Finally, the performance of the trained MLPNN-BP classifier is analyzed based on the classification of test tuples. The activities involved in this methodology are outlined in the Fig. 1.

Data preprocessing: Initially the given dataset is pre-processed by categorizing the attribute values based on the domain. Health care dataset consists of continuous valued attributes which cannot be directly taken for processing. With Weka 3.7 the continuous valued attributes are discretized using Discretize filter. Discretization is performed by simple binning with findNumBins set as False and number of bins as 10. Even the missing values for nominal and numeric attributes in a dataset are also replaced with the modes and means from the training data. Specialty of health care data lies in the fact that the attribute values can only be within certain ranges. As per the requirements of the algorithm, the entire non class attribute values are transformed to nominal attributes. All the six health care datasets are separately preprocessed using Weka 3.7 according to the requirements of the designed system.

Extraction of CAR's from the preprocessed dataset: An association rule is an implication of the form, X→Y where X and Y are disjoint item sets. A Class Association Rule (CAR) is the special subset of association rules whose consequent (right-hand-side) is restricted to a class attribute. Apriori is one of the well known standard algorithms for discovering interesting rules between attributes of the dataset. This phase presents a way for finding the CAR's from the dataset. Apriori algorithm identifies all possible associations between the non-class attributes towards class attribute in the dataset by generating the set of CAR's satisfying the minimum support count and confidence.

Generation of reduced set of significant rules using GA: The major drawback in association rule mining is the generation of large volume of uninteresting rules. Mining useful knowledge from the large volume of rules is tedious task. Extracting interesting CAR's encourages the efficiency of the classification system. Hence all CAR's extracted from the dataset are taken into GA for optimizing the class rules. A CAR represents a chromosome in an initial population. The fitness value of each chromosome is evaluated. The set of chromosomes satisfying the fitness function is selected for the next generation. In this study, the size of the population has been set as 10 and single point crossover, mutations are chosen as GA operators for producing next generation chromosomes. Single point crossover combines two individuals to create new individuals (i.e., it exchanges features of two individuals to produce new individuals). A single point is fixed at a particular position in the rule pairs selected for mating thereby results in generating new set of class rules. This generational process is repeated until either of the termination conditions such as maximum number of iteration is 100 or fitness value of the chromosome is 0.5 is reached. Finally, after termination, reduced set of CAR's are retained as interesting rules. The detailed steps followed in GA for generating the reduced set of significant CAR's is shown in Fig. 2.

Fitness function: In this study, the fitness value is used to evaluate the significance of each chromosome. The fitness value of each chromosome is evaluated using the maximization function given in Eq. (1). The fitness value of a CAR X→Y, where X is a set of non-class attributes and Y is a class attribute is as follows:

$$Max\ f(X \to Y) = \alpha * Cov(X \to Y) + \beta * Conf(X \to Y) + \gamma * Supp(X \to Y) \quad (1)$$

Subject to:

$$\alpha + \beta + \gamma = 1,\ 0 \le \alpha \le 1,\ 0 \le \beta \le 1,\ 0 \le \gamma \le 1$$

where,

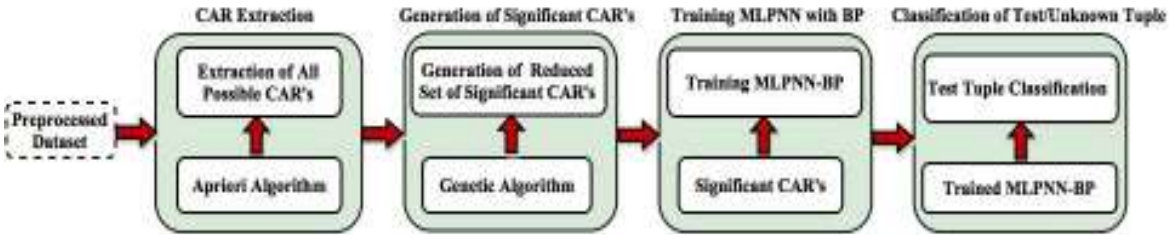

Fig. 1: Work flow of the proposed methodology

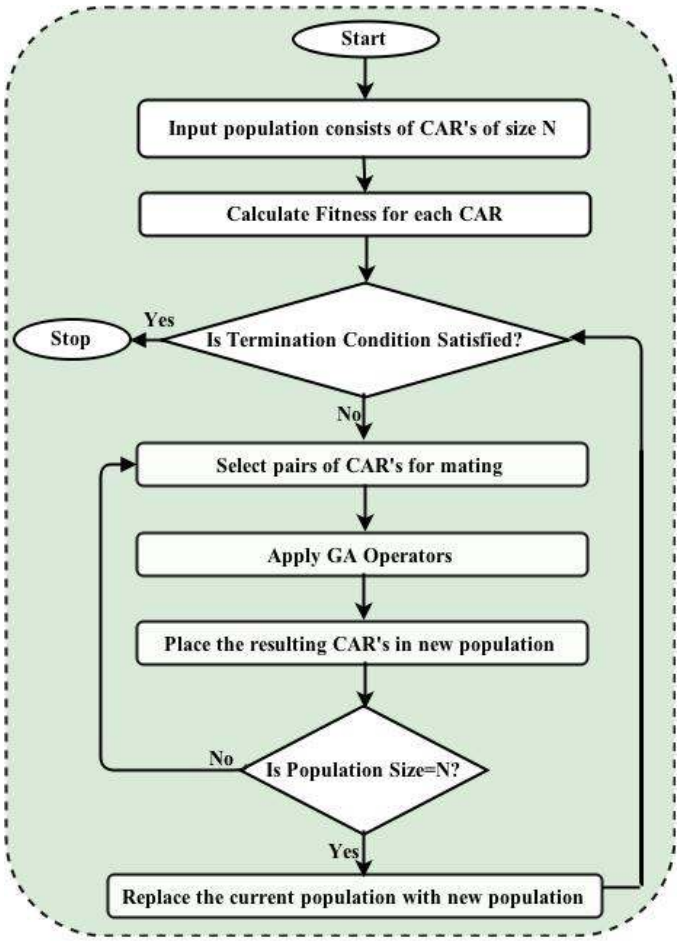

Fig. 2: Steps followed in generating the reduced set of significant CAR's using GA

$$Cov(X \rightarrow Y) = \frac{No.\ of\ training\ tuples\ containing\ X}{Total\ No.\ of\ training\ tuples} \quad (2)$$

$$Conf(X \rightarrow Y) = \frac{No.\ of\ training\ tuples\ containing\ X\ and\ Y}{No.\ of\ training\ tuples\ containing\ X} \quad (3)$$

$$Supp(X \rightarrow Y) = \frac{No.\ of\ training\ tuples\ containing\ X\ and\ Y}{Total\ No.\ of\ training\ tuples} \quad (4)$$

where, $\acute{a}$, β and γ are user specified significance value for Coverage (Cov), Confidence (Conf) and Support (Supp) measures.

Training the MLPNN with BP using significant CAR's: In this study, the MLPNN-BP is trained using reduced set of significant CAR's instead of preprocessed dataset. Figure 3 illustrates the architecture of MLPNN-BP. Accomplishing successful NN training depends on the number and quality of parameters set in each layer whereas in the literature, there were no proper guidelines to determine these parameters. Generally an interesting CAR consists of an antecedent and a class attribute as consequent, where antecedent holds a set of significant attribute (s) associated towards class attribute. Use of reduced set of significant CAR's as training instances makes MLPNN elegant. i.e., MLPNN can be designed using few significant nodes and links. On the other hand, using preprocessed dataset as training instances requires many nodes in the input layer. Hence training a NN using few significant CAR's always yields better results than training using the original dataset.

The working of MLPNN using BP algorithm is outlined in Fig. 4. Initially the random weights are assigned to the links between the nodes. Output of each neuron is calculated using Eq. (5). Equation (7) is used to calculate the squared error i.e., difference between the expected output and the actual output. The calculated squared error is back propagated and link weights are updated until the Total Mean Squared error (TMS) is less than minimum threshold value. In this study, minimum threshold value is set as 0.01:

$$O_j = \frac{1}{1 + e^{-I_j}} \quad (5)$$

where, O_j is the output of j^{th} neuron, I_j is the input of the j^{th} neuron:

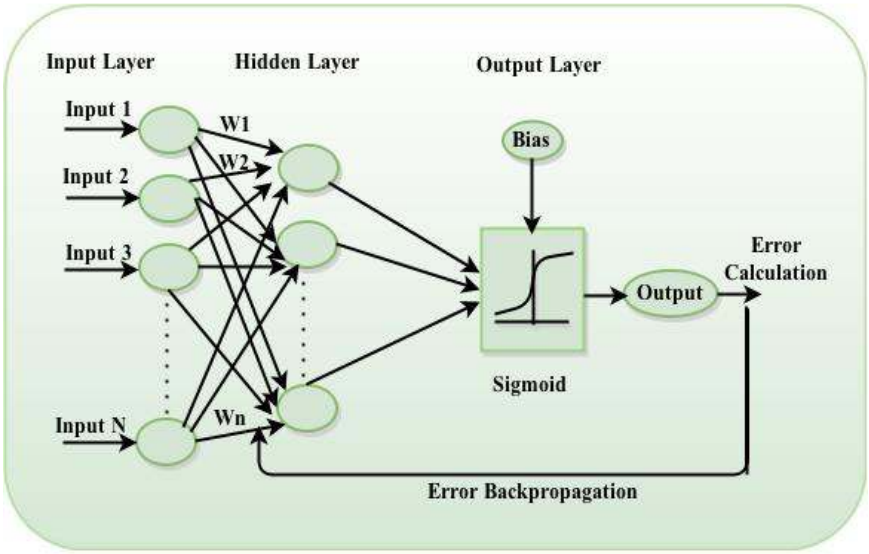

Fig. 3: The architecture of MLPNN-BP

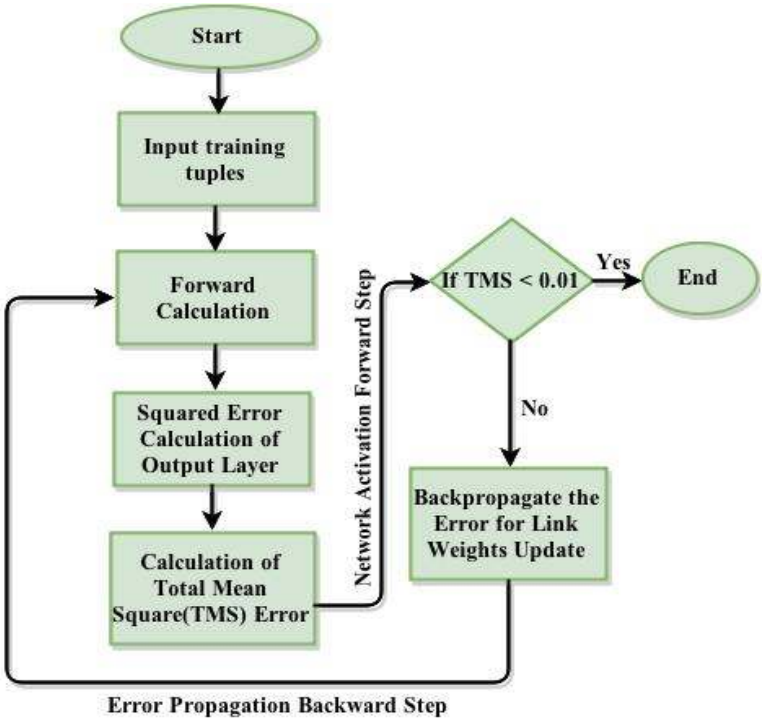

Fig. 4: Working procedure of MLPNN using BP algorithm

$$I_j = \sum O_i . W_{ij} - \theta_j \quad (6)$$

where, W_{ij} represents the weight of the link (i, j), θ_j represents bias for node 'j':

$$E = (t - y)^2 \quad (7)$$

where, 't' represents the desired output of the output neuron and 'y' represents the actual output in the training dataset.

Test tuple classification: The classifier accuracy of the MLPNN is evaluated using 10-fold cross validation (i.e., the entire dataset is split into 10 sets of equal size. 9 sets are taken for training and 1 set is taken for testing. Once the classifier is built using training sets; its accuracy is evaluated using test set. The mean accuracy for 10 iterations is considered as final classifier accuracy). The accuracy of the final classifier is calculated using Eq. (8) as the ratio between the numbers of correctly classified tuples to the total number of tuples in the dataset:

Accuracy = No. of correctly classified tuples/Total no. of tuples (8)

EXPERIMENTS AND RESULTS

The experiments were conducted using six health care datasets namely Heart, Hepatitis, Breast Cancer, Cleve, Pima and Sick from the UCI machine learning repository. As a sample, details of the Heart disease dataset are discussed in this study. Heart disease dataset is binary class dataset commonly used to predict the presence of heart disease for a given set of patient's signs and symptoms. It consists of 270 tuples and 14 attributes, where the 14th attribute is a class attribute used to predict the presence/absence of the heart disease. Table 1 depicts the attributes and the domain values of the Heart disease dataset. Initially the given dataset is pre-processed by categorizing the attribute values based on the domain. As per the requirements of the algorithm, the entire non class attribute values are transformed to nominal attributes. The data preprocessing tasks that are carried out in this study is explained in above section. After pre-processing, the basic Apriori algorithm in Weka 3.7 is employed for CAR's extraction. The number of rules extracted from the Heart disease dataset is 90. The number of class rules generated for the class label 0 (i.e., absence of heart disease) is 77, whereas for the class label 1 (i.e., presence of heart disease) is 13. CAR's extracted from the Heart disease dataset using Apriori algorithm is shown as screenshot in Fig. 5.

In order to improve the accuracy of the classifier merely the efficient class rules with high quality are required for the classifier construction. GA is used to uncover the significant rules from the voluminous rules. Each extracted class rule is taken as a chromosome which is represented in string format. The parameters

Fig. 5: CAR's extracted using basic apriori algorithm

such as the total population size, minimum threshold fitness value, maximum number of iterations, á, β and γ are set with default values 10, 0.5, 100, 0.5, 0.34 and 0.16 respectively. Initial population contains CAR's extracted from the dataset using basic Apriori algorithm. The fitness value of each CAR is evaluated using Eq. (1). For each iteration, the termination condition is verified. Selection of the rule pairs suitable for mating is of random type. The single point cross over and mutation are performed among them to

Table 1: Attributes of the heart disease dataset

Attribute	Domain
Age	Numeric, values ranges from 1-10
Sex	Numeric, values ranges from 1-10
Chest pain	Numeric, values ranges from 1-10
Resting BP	Numeric, values ranges from 1-10
Cholesterol	Numeric, values ranges from 1-10
Blood sugar	Numeric, values ranges from 1-10
ECG	Numeric, values ranges from 1-10
Max heart rate	Numeric, values ranges from 1-10
Angina	Numeric, values ranges from 1-10
Old peak	Numeric, values ranges from 1-10
STSlope	Numeric, values ranges from 1-10
Vessels	Numeric, values ranges from 1-10
Thal	Numeric, values ranges from 1-10
Class	Diagnosis of heart disease (Disease status, Value 0: <50% diameter narrowing, Value 1: >50% diameter narrowing

Table 2: Sample tuples

Attr-1	Attr-2	Attr-3	Attr-4	Attr-5	Attr-6	Attr-7	Attr-8	Attr-9	Class
A1	A11	A21	A31	A42	A51	A62	A71	A81	C0
A5	A17	A27	A31	A45	A58	A63	A74	A81	C0
A5	A13	A21	A32	A42	A51	A62	A71	A81	C0
A3	A11	A21	A31	A42	A51	A62	A71	A81	C1
A3	A11	A21	A31	A42	A51	A62	A71	A81	C1

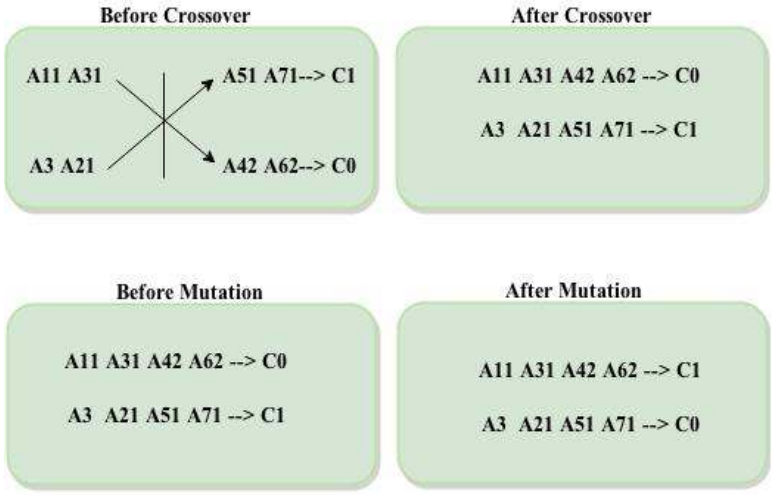

Fig. 6: Single point crossover and mutation

```
A12 A22==> C0 0.7816667
A12 A82 A114==> C0 0.7404487
A2 A82 A114==> C0 0.7404487
A2 A82 A123==> C0 0.7404487
A1 A114 A123==> C0 0.7292222
A2 A114 A123==> C0 0.7292222
A12 A82 A114 A123==> C0 0.6920889
A1 A82 A114 A123==> C0 0.6920889
A12 A52 A103==> C0 0.6896243
A12 A114 A123  ==> C0 0.6772951
A11 A23 A111==> C1 0.67529684
A1 A52 A114 A123==> C0 0.66813594
A12 A82 A95 A123==> C0 0.6592886
A12 A82 A95 A114==> C0 0.6592886
A12 A52 A82 A103==> C0 0.6529223
A12 A52 A82 A114==> C0 0.6529223
A12 A52 A82 A123==> C0 0.6529223
A1 A95 A114 A123==> C0 0.65182495
A1 A52 A114 A123  ==> C0 0.63384557
A12 A95 A82 A123  ==> C0 0.6325872
A12 A114 A82 A123  ==> C0 0.63042545
A12 A52 A82 A114 A123==> C0 0.6300271
A12 A103 A123  ==> C0 0.62064767
A12 A52 A95 A123==> C0 0.6151007
A12 A52 A95 A103==> C0 0.6151007
A12 A82 A95 A114 A123==> C0 0.60999477
A1 A82 A95 A114 A123==> C0 0.60999477
A11 A23 A52 A111==> C1 0.60958135
A1 A52 A95 A114 A123==> C0 0.59738517
A12 A52 A82 A95 A114==> C0 0.58223796
```

Fig. 7: The reduced set of significant CAR's obtained using GA

generate the new set of CAR's. The computation of fitness value for CAR's using GA is explained using the sample tuples given in Table 2.
For example:

A42 A51→C0, Coverage = 0.8, Confidence = 0.5, Support = 0.4, Fitness = 0.634

In the above example, to calculate the fitness value of the class rule (A42 A51→ C0), the total number of tuples in Table 2 satisfied by both the rule antecedent and consequent is 2.0, hence the total rule count is set as 2.0. The number of tuples satisfied by the rule antecedent is 4.0, hence LHS count is 4.0. By using this, the coverage, confidence and support values for this class rule are found to be 0.8, 0.5 and 0.4 respectively. Using Eq. (1), the fitness value of this class rule is obtained as 0.634. Similarly the fitness value for the CAR's extracted from the dataset is calculated. The set of CAR's satisfying the minimum threshold fitness value are then taken for cross over and mutation in GA to generate the new set of significant CAR's suitable for an efficient classifier construction. The mutation has also been performed on the class labels to form new set of class rules after cross over. i.e., the rule with the class label C0 is mutated as C1 and vice versa. Figure 6 describes the single point crossover and mutation carried on CAR's.

After performing the crossover and mutation, if the number of significant CAR's (i.e., population size) thus obtained using GA is 10, then the current CAR's in the population should be replaced with the new set of rules. If the size is not 10, then the random selection, single point cross over and mutation are need to be performed to generate the new set of significant rules. The iteration proceeds until the either of the termination condition (i.e., maximum iteration of 100 or to the average minimum fitness threshold value of 0.5) is satisfied. Use of GA greatly reduces the size of class rules from 90 to 30 for the Heart disease dataset. Figure 7 shows the reduced set of significant CAR's obtained using GA along with their fitness value for Heart disease dataset. Table 3 represents the facts of different types of training instances such as the dataset tuples, all possible CAR's extracted using Apriori algorithm and the reduced set of CAR's after GA considered in this study.

It is inferred from Table 3, for the Breast cancer dataset, use of GA reduces the CAR's from 56 to 20. Similarly for the other datasets such as Cleve, Hepatitis, Heart, Pima and Sick, GA greatly reduces the number of CAR's from 110, 35, 90, 121 and 91 to 62, 13, 30, 58 and 49 respectively. Thus the reduced set of significant CAR's obtained using GA is taken as input for the MLPNN-BP construction. Each attribute

Table 3: Training instances

Dataset	No. of tuples in the dataset after preprocessing	No. of CAR's extracted using Apriori	No. of significant CAR's generated using GA
Breast cancer	699	56	20
Cleve	303	110	62
Hepatitis	155	35	13
Heart	270	90	30
Pima	768	121	58
Sick	2800	91	49

Table 4: Comparison of MLPNN-BP in terms of classifier accuracy (%) and training time (ms)

	Classifier accuracy (%)			Training time (ms)		
Dataset	PDS	All-CAR's	R-CAR's	PDS	All-CAR's	R-CAR's
Breast cancer	95.99	92.85	90	489	21	12
Cleve	83.16	94.37	95.45	212	41	37
Hepatitis	83.87	62.85	70	109	13	8
Heart	79.62	95.03	96.67	386	37	18
Pima	73.30	95.34	94.82	538	45	32
Sick	93.60	92.30	91.02	967	34	29
Average	84.92	88.79	89.66	450.17	31.83	22.67

PDS (Preprocessed DataSets) i.e., Pre-processed datasets with significant attributes and tuples; All-CAR's (All possible CAR's) i.e., All possible class rules extracted using Apriori algorithm; R-CAR's (Reduced set of significant CAR's) i.e. Reduced set of CAR's generated using GA

available in the CAR is provided as input to the nodes in the input layer. The number of hidden layer is set as 3. Initially the weight of the each link between the nodes is assigned as 1.0. The number of the output nodes depends on the class label available in the dataset. Since the health care datasets considered in this study are binary class. The number of output node is 1. The sigmoid function given in Eq. (5) is used to compute the output of each neuron. The squared error is also calculated using Eq. (7) for identifying the difference between the actual and the predicted output. Finally the accuracy and training time of the MLPNN-BP is computed.

Table 4, shows the classifier accuracy and training time of MLPNN-BP using preprocessed datasets, complete set of CAR's and reduced set of CAR's using GA as training instances using 10 fold cross validation test option. From the Table 4, it is found that, the MLPNN-BP has achieved the average accuracy of 84.92% over the six health care datasets. Construction of MLPNN-BP using CAR's greatly improves the accuracy of MLPNN-BP classifier by 4% nearly i.e., 84.92 to 88.79%. Results prove that, use of CAR's as training instance for MLPNN-BP shows improvement in classifier accuracy for almost all datasets except Breast Cancer, Hepatitis and Sick. Further, in this study, GA is employed to reduce the number of CAR's without compromising the interestingness factor. MLPNN-BP using GA achieved a classification accuracy of 89.66% (i.e., averaged over all datasets) approximately 1% higher the accuracy achieved without GA. Results show that time required to train the MLPNN-BP is drastically reduced from 450.17 ms (i.e., averaged for six datasets) to 22.67 ms which is approximately 1% of its original time. As part of this work, an experimental study on the performance of MLPNN-BP has been done. Employing CAR's as training instances for MLPNN-BP encourages

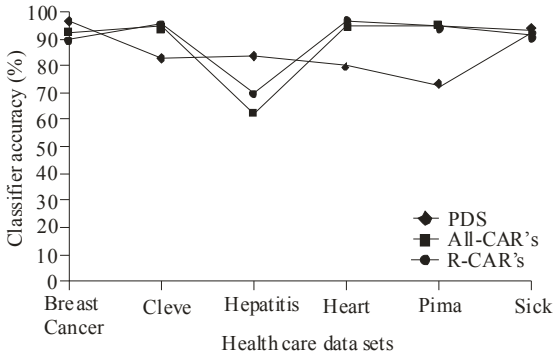

Fig. 8: Performance comparison of MLPNN-BP algorithm in terms of classifier accuracy (%) over six health care datasets with and without CAR's as training instances

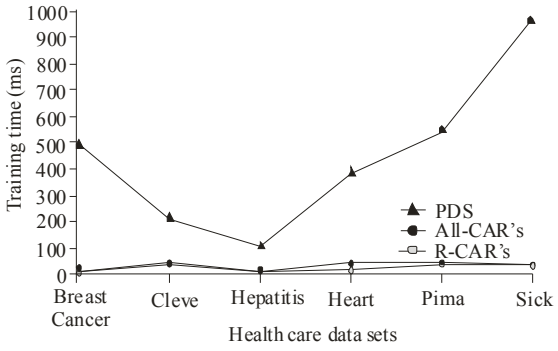

Fig. 9: Performance comparison of MLPNN-BP algorithm in terms of training time (ms) over six health care datasets with and without CAR's as training instances

accuracy by 12, 17, 21%, respectively than MLPNN-BP using preprocessed datasets as training instances for Cleve, Heart and Pima datasets. Use of reduced set of significant CAR's discourages accuracy by 6, 14 and 3%, respectively for Breast Cancer, Hepatitis and Sick

datasets. Since most of the attributes of these datasets are naturally significant and cannot be ignored, hence the generation of few significant CAR's from these datasets representing only few attributes and tuples which does not promote the classifier accuracy. The accuracy and training time of the MLPNN-BP over the six health care datasets with and without CAR's are represented as line graphs in Fig. 8 and 9.

CONCLUSION AND FUTURE SCOPE

MLPNN-BP is used to construct a classifier with significant CAR's as training instances for HCDS. It achieves an average accuracy of 88.79% which is approximately 4% higher than the average accuracy achieved using preprocessed datasets. Further the accuracy of the classifier has been enhanced by generating the reduced set of significant CAR's using GA. Hence the MLPNN-BP built using the reduced set of significant CAR's as training instances achieves an average accuracy of 89.6% which is approximately 1% higher than the average accuarcy obtained using all CAR's. Combination of MLPNN-BP and GA achieves better classifier accuracy compared to other combinations. From the results, it is found that the computation of reduced set of significant CAR's using GA decreases the training time by 96.24% i.e., from 450.17 ms to 22.67 ms. It is infered that the reduced set of significant CAR's requires fewer nodes and links in the construction of MLPNN-BP which makes training process simple. Binary class datasets such as Breast Cancer, Cleve, Pima, Heart, Sick and Hepatitis are considered for experimentation. In future, multi-class datasets can be focused; other optimal class rule extraction techniques can be employed to get better accuracy and training time of the classifier.

REFERENCES

Agrawal, R. and R. Srikant, 1994. Fast algorithms for mining association rules in large databases. Proceeding of the 20th International Conference on Very Large Databases (VLDB, 1994). Santiago, Chile, September, pp: 487-499.

Ahmad, F., N.A.M. Isa, Z. Hussain and M.K. Osman, 2013. Intelligent medical disease diagnosis using improved hybrid genetic algorithm-multilayer perceptron network. J. Med. Syst., 37(2): 1-8.

Baxt, W.G., 1990. Use of an artificial neural network for data analysis in clinical decision-making: The diagnosis of acute coronary occlusion. Neural. Comput., 2(4): 480-489.

Elveren, E. and N. Yumuşak, 2011. Tuberculosis disease diagnosis using artificial neural network trained with genetic algorithm. J. Med. Syst., 35(3): 329-332.

Han, J., J. Pei and Y. Yin, 2000. Mining frequent patterns without candidate generation. SIGMOD Rec., 29(2): 1-12.

Hebb, D., 1949. The Organization of Behavior. John Wiley and Sons, New York.

Karegowda, A.G., A.S. Manjunath and M.A. Jayaram, 2011. Application of genetic algorithm optimized neural network connection weights for medical diagnosis of pima indians diabetes. Int. J. Soft Comput., 2(2): 15-23.

Khan, J., J.S. Wei, M. Ringner, L.H. Saal, M. Ladanyi, F. Westermann and P.S. Meltzer, 2001. Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. Nat. Med., 7(6): 673-679.

Kumar, K. and B. Abhishek, 2012. Artificial neural networks for diagnosis of kidney stones disease. Int. J. Inf. Tech. Comput. Sci., 4(7): 20-25.

Lisboa, P.J. and A.F. Taktak, 2006. The use of artificial neural networks in decision support in cancer: A systematic review. Neural Networks, 19(4): 408-415.

Mantzaris, D., G. Anastassopoulos and A. Adamopoulos, 2011. Genetic algorithm pruning of probabilistic neural networks in medical disease estimation. Neural Networks, 24(8): 831-835.

McCulloch, W.S. and W. Pitts, 1943. A logical calculus of the ideas immanent in nervous activity. B. Math. Biophys., 4(5): 115-133.

Rosenblatt, F., 1958. The perceptron: A probabilistic model for information storage and organization in the brain. Psychol. Rev., 65(6): 386-408.

Rumelhart, D.E., G.E. Hinton and R.J. Williams, 1986. Learning representations by back-propagating errors. Nature, 323(6088): 533-536.

Setiono, R., 1996. Extracting rules from pruned neural networks for breast cancer diagnosis. Artif. Intell. Med., 8(1): 37-51.

Tu, J.V., 1996. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. J. Clin. Epidemiol., 49(11): 1225-1231.

Ture, M., I. Kurt, A.T. Kurum and K. Ozdamar, 2005. Comparing classification techniques for predicting essential hypertension. Expert Syst. Appl., 29(3): 583-588.

Vinterbo, S. and L. Ohno-Machado, 2000. A genetic algorithm approach to multi-disorder diagnosis. Artif. Intell. Med., 18(2): 117-132.

Wu, Y., M.L. Giger, K. Doi, C.J. Vyborny, R.A. Schmidt and C.E. Metz, 1993. Artificial neural networks in mammography: Application to decision making in the diagnosis of breast cancer. Radiology, 187(1): 81-87.

Yan, H., Y. Jiang, J. Zheng, C. Peng and Q. Li, 2006. A multilayer perceptron-based medical decision support system for heart disease diagnosis. Expert Syst. Appl., 30(2): 272-281.