A Classification Technique using Associative Classification

Prachitee B. Shekhawat Prof. Sheetal S. Dhande
Dept. of Computer science and Engineering
Sipna's College of Engineering and Technology, Amravati, Maharashtra, India.

ABSTRACT
Classification and association rule mining are two basic tasks of Data Mining. Classification rule mining is used to discover a small set of rules in the database to form an accurate classifier. Association rules mining has been used to reveal all interesting relationships in a potentially large database. An Apriori approach, which was used to generate the association rules from frequent patterns, turn out to generate a huge time-intensive query called as iceberg query. Various researches have been done under the Apriori-like approach to improve performance of the frequent pattern mining tasks but the results were not as much as expected due to many scans on the dataset. This project aims to propose a flexible way of mining frequent patterns by extending the idea of the Associative Classification methods. For better performance, the Neural Network Association Classification system is proposed here to be one of the approaches for building accurate and efficient classifiers. In this paper, the Neural Network Association Classification system is used in order to improve its accuracy. The structure of the network reflects the knowledge uncovered in the previous discovery phase. The trained network is then used to classify unseen data. The performance of the Neural Network Associative Classification system is compared with the previous Classification Based Association on four datasets from UCI machine learning repository.

Keywords
Data mining, Class Association Rules, classification, Backpropagation neural network.

1. INTRODUCTION
Momentous changes in the databases technology and the application of databases management system have been driven by the need of storing the large amount of information and at the same time have driven those changes. Information is collected almost everywhere in our everyday life and if the information can be extracted from the database they will create a lot of potential profit the companies. But this large amount of information makes human analysis intractable.

How this mountain of raw data can be explored? Most of it will never be seen by human eyes and if viewed could not be analyzed by hand. Computer provides the obvious answer. The computer we should use to process the data then becomes the issue. Although simple statistical methods were development long ago, they are not powerful enough. Data mining provides tools to reveal previously unknown information in large databases. Data Mining is categorized into different tasks depending on what the data mining algorithm under consideration is used for. A rough characterization of the different data mining tasks can be achieved by dividing them into descriptive and predictive tasks. A predictive approach tries to assign a value to a future or unknown value of other variables or database fields, whereas description tries to summarize the information in the database and to extract patterns. Association rule mining is a descriptive data mining task. A rule, out of the set of association rules, is one descriptive pattern—a compact description for a very small subset of the whole data. In other words, Association rules are used to analyze relationships between data in large databases. Typical predictive tasks are classification and regression. Classification involves learning a function which is capable of mapping instances into distinct classes. It is kind of supervised learning algorithm by inputting some data to training and another one that is different from training to test the data. The result from training data is classifier. Regression maps instances to a real-valued variable. Now both the association rule mining and classification rule mining can be integrated to form a framework called as Associative Classification and these rules are referred as Class Association Rules.

The use of association rules for classification is restricted to problems where the instances can only belong to a discrete number of classes. The reason is that association rule mining is only possible for categorical attributes. The head Y of an arbitrary association rule $X \rightarrow Y$ is a disjunction of items. However, association rules in their general form cannot be used directly. We have to restrict their definition. Every item which is not present in the rule body may occur in the head of the rule. When we want to use rules for classification, we are interested in rules that are capable of assigning a class membership. Therefore we restrict the head Y of a class association rule $X \rightarrow Y$ to one item. The attribute of this attribute-value-pair has to be the class attribute. A class association rule is obviously a predictive task. By using the discriminative power of the Class Association Rules we can also build a classifier.

Data mining in the proposed, Neural Network Associative Classification, system thus consists of three steps:
- If any, discretizing the continuous attributes,
- Generating all the Class Association Rules (CARs), and
- Building a classifier with the help of Backpropagation Neural Network based on the generated CARs set.

Here we have to analyze the data and mine all the accurate association rule that will be use to build an efficient classifier. This method will make the following contribution:
- It proposes a new way to build accurate classifiers. Experimental results show that classifiers built this way are, in general, more accurate than the previous classification system
- It makes association rule mining techniques applicable to classification tasks.

- It helps to solve a number of important problems with the existing classification systems.

The paper is organized as follows: section 2 contains the brief introduction to major previous work done about data mining. Section 3 describes the proposed system, Neural Network Associative Classification, architecture and section 4 presents our experimental setup and discussed the result. Section 5 concludes the paper.

2. LITERATURE SURVEY

The data analysis algorithms (or data mining algorithms, as they are more popularly known nowadays) can be divided into three major categories based on the nature of their information extraction [1]:

- Clustering (also called segmentation or unsupervised learning),
- Predictive modeling (also called classification or supervised learning), and
- Frequent pattern extraction.

Clustering is the major class of data mining algorithms. The goal of the search process used by these algorithms is to identify all sets of similar examples in the data, in some optimal fashion. One of the oldest algorithms for clustering is k-means [2]. The two disadvantages of this algorithm are initialization problem and that the cluster must be linearly separable. To deal with the initialization problem, the global k-means has been proposed [3], which is an incremental-deterministic algorithm that employs k-means as a local search procedure. Kernel k-means algorithm [4] avoids the limitation of linearly separable clusters and it mapped the data points from input space to a higher dimensional feature through a nonlinear transformation Ø and the k-means is applied in the feature space. Global kernel k-means [5] is an algorithm which mapped data points from input space to a higher dimensional feature space through the use of a kernel function and optimizes the clustering error in the feature space by locating near-optimal solution. Because of its deterministic nature, this makes it independent of the initialization problem, and the ability to identify nonlinearly separable cluster in input space. So global kernel k-means algorithm combines the advantages of both global k-means and kernel k-means. Another approach for clustering data is hierarchical clustering that is based on the Hungarian method [6] and the computational complexity of the proposed algorithm is $O(n^2)$.

The important classification algorithms are decision tree, Naive-Bayes classifier and statistics [2]. They use heuristic search and greedy search techniques to find the subsets of rules to find the classifiers. C4.5 and CART are the most well-known decision tree algorithms.

The final class of data mining algorithms is frequent pattern extraction. For a large databases, [7] describes an apriori algorithm that generate all significant association rules between items in the database. The algorithm makes the multiple passes over the database. The frontier set for a pass consists of those itemsets that are extended during the pass. In each pass, the support for candidate itemsets, which are derived from the tuples in the databases and the itemsets contain in frontier set, are measured. Initially the frontier set consists of only one element, which is an empty set. At the end of a pass, the support for a candidate itemset is compared with the minsupport. At the same time it is determined if an itemset should be added to the frontier set for the next pass. The algorithm terminates when the frontier set is empty. After finding all the itemsets that satisfy minsupport threshold, association rules are generated from that itemsets.

Bing Liu and et al.[8] had proposed an Classification Based on Associations (CBA) algorithm that discovers Class Association Rules (CARs). It consists of two parts, a rule generator, which is called CBA-RG, is based on Apriori algorithm for finding the association rules and a classifier builder, which is called CBA-CB. In Apriori Algorithm, itemset (a set of items) were used while in CBA-RG, ruleitem, which consists of a condset (a set of items) and a class. Class Association Rules that are used to create a classifier in [8] is more accurate than C4.5 [2] algorithm. But the Classification Based on Associations (CBA) algorithm needs the ranking rule before it can create a classifier. Ranking depends on the support and confidence of each rule. It makes the accuracy of CBA less precise than Classification based on Predictive Association Rules.

Neural network is a parallel processing network which generated with simulating the image intuitive thinking of human, on the basis of the research of biological neural network according to the features of biological neurons and neural network and by simplifying, summarizing and refining[9]. It uses the idea of non-linear mapping, the method of parallel processing and the structure of the neural network itself to express the associated knowledge of input and output. Initially, the application of the neural network in data mining was not optimistic, because neural networks may have complex structure, long training time, and uneasily understandable representation of results. But its advantages such as high affordability to the noise data and low error rate, the continuously advancing and optimization of various network training algorithms, especially the continuously advancing and improvement of various network pruning algorithms and rules extracting algorithm, make the application of the neural network in the data mining increasingly favored by the overwhelming majority of users. Xianjun Ni [10] describes Data mining process based on neural network. This process is composed of three main steps as data preparation, rule extraction and rules assessment.

3. NUERAL NETWORK ASSOCIATIVE CLASSIFICATION SYSTEM ARCHITECTURE

The proposed system, Neural network Associative Classification, architecture is shown in Figure 1. It consists of three phases: Pre-processing Phase, Class Association Rule Mining Phase and the Classification Phase.

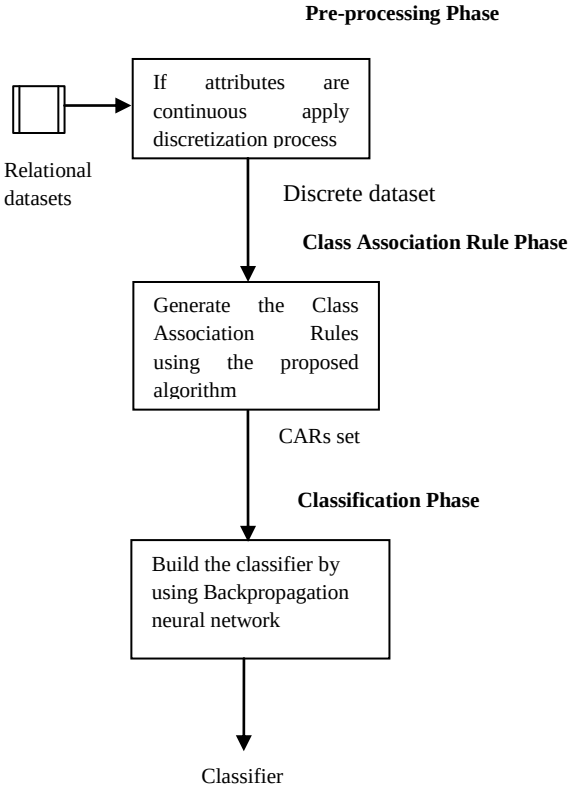

Figure 1: Neural network Associative Classification System Architecture.

3.1 Pre-processing phase

The main goal of this phase is to prepare the data for the next phase: Class Association Rule Mining Phase. The data mining based on the neural network can only handle the numeric data so it is needed to transform the character date into numeric data. Classification datasets often contain many continuous attributes. Mining of association rules with continuous attributes is still a research issue. So our system involves discretizing the continuous attributes based on the classification pre-determined class target. For discretization, we will use the algorithm from [11] which is implemented by Guangdi Li.

3.2 Class Association Rules Mining phase

This phase presents a way for finding the Class Association Rules (CARs) from the data. Class Association Rules have the following related theory.

3.2.1. Association Rules

The goal of this technique is to detect relationships or associations between specific values of categorical variables in large data sets.

Consider:

$D = \{d_1, d_2, \ldots d_n\}$ is a database that consist of set of n data and d ⊆ I

$I = \{i_1, i_2, \ldots i_m\}$ is a set of all items that appear in D

The association rule is an implication expression of the form $X \rightarrow Y$, where X and Y are disjoint itemsets, i.e., $X \cap Y = \emptyset$ and where $X, Y \subseteq I$. The strength of an association rule can be measured in terms of its support and confidence.

- Support value is frequency of number of data that consist of X and Y or P(X ∪ Y) and is given by

$$support, s(X \rightarrow Y) = \frac{\sigma(X \cup Y)}{N} \quad \ldots(1)$$

- Confidence is frequency of number of data that consist of X and Y or P(Y|X) and given by

$$confidence, c(X \rightarrow Y) = \frac{\sigma(X \cup Y)}{\sigma(X)} \quad \ldots(2)$$

We have to find all the rules that satisfy pre-specified minimum support count and minimum confidence.

3.2.2. Data Classification

Classification is a task of assigning objects to one of several predefined categories. It consists of predicting the value of a (categorical) attribute (the class) based on the values of other attributes (the predicting attributes).

3.2.3. Class Association Rules

The Class Association Rules are the special subset of the association rules whose head (right-hand-side) are restricted to the classification class attribute.

According to this, a class association rule is of the form

$$X \rightarrow a_i$$

where a_i is the class attribute and X is $\{a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n\}$.

Consider:

$D = \{d_1, d_2, \ldots d_i\}$ is a database that consist of set of data that have n attribute and class label by d = $\{a_1, a_2 \ldots a_n, y_k\}$ where k= 1, 2,...m.

$I = \{X_1, Xi_2, \ldots X_m\}$ is a set of all items that appear in D

$Y = \{y_1, y_2, \ldots y_m\}$ is a set of class label m class

A class association rule (CAR) is an implication of the form $X \rightarrow y$, where X ⊆ I, and y ⊆ Y. A rule $X \rightarrow y$ holds in D with support = s% and confidence = c% following these conditions:

- Support value is the frequency of number of all data that consist of item set X and class label y.
- Confidence value is the frequency of number of data that consist of class label y when data have item set X.

Before formulating the formulas for support and confidence, few notations are to be defining as:

For CAR, $condset \rightarrow y$, where condset is an itemset and y ⊆ Y.

- ruleitem has the format $< (condset, condsupCount), (y, rulesupCount) >$.
- k-ruleitem is ruleitem that condset consist of condset.
- condsupCount is the number of data that consist of condset.
- rulesupCount is the number of data that consist of condset and has the label y.

Then

$$support = \frac{rulesupCount}{|D|} * 100\% \quad \ldots(3)$$

$$confidence = \frac{rulesupCount}{condsupCount} * 100\% \quad \ldots(4)$$

The proposed algorithm, which is given below, will use this notations and formulae to find class association rules.

The algorithm 1 is shown below which is used to find the Class Association Rules:

F_1= {large 1-ruleitems};
CAR_1 = genRules(F_1);
for (k=2; F_{k-1}=Ø; k++)
C_k= candidateGen(F_{k-1});
    for each data case d ∈ D
        C_d = ruleSubset(C_k, d);
        for each candidate c ∈ C_d
            c.condsupCount++;
            if d.class=c.class then
                c.rulesupCount++;
            end
        end
    end
    F_k = { c ∈ C_k | c.rulesupCount ≥ minsup};
    CAR_k = genRules (F_k);
end
CARs= $\bigcup_k CAR_k$;

Algorithm 1: Proposed Algorithm

An example of searching the Class Association Rules following the above algorithm is given below. Define data to search Class Association Rules is detailed in Table 1.

- A and B are predicted attributes that A has possible values that are p, q, r and B has possible values are s, t, u.
- C is class label and possible values are Yes and No.

When search Class Association Rules by define:

- Minimum support(minsup) = 15%
- Minimum confidence(minconf) = 60%

Table 1. Example data to search Class Association Rules.

	A	B	C
1.	p	S	Yes
2.	p	T	Yes
3.	q	S	Yes
4.	r	T	No
5.	q	T	No
6.	p	S	No
7.	p	S	Yes
8.	r	U	Yes
9.	r	U	No
10.	q	U	No

In cases where a ruleitem is associated with multiple classes, only the class with the largest frequency is considered by current associative classification methods. We can express details as in Table 2. As we know the format for CAR $condset \rightarrow y$ is

$$< (condset, condsupCount), (y, rulesupCount) >$$

Consider for example the ruleitem $< \{(A,p)\}, (C, Yes) >$, means that the frequency of condset (A,p) in data is 4,it is also called as condsupCount, and the frequency of (A,p) and (C,Yes) occurring together is 3, also known as rulesupCount. So, the support and confidence are 30% and 75% respectively by using equations 3 and 4.

As it satisfy the minsup threshold ruleitem can be added to frequent 1-ruleitem (F_1). Also it satisfy the minconf threshold so it can form the Class Association Rules (CAR_1) as

$$(A,p) \rightarrow (C, Yes)$$

Repeat this procedure and generate all F_1 ruleitems. Using F_1 ruleitem form Candidate 2-ruleitms C_2 in the same way then draw out ruleitem that satisfy minsup and minconf threshold to form F_2 set.

From Table 2 we can express Class Association Rules as:

1. $(A,p) \rightarrow (C, Yes) \quad s = 30\%, c = 75\%$
2. $(A,q) \rightarrow (C, No) \quad s = 30\%, c = 66.67\%$
3. $(A,r) \rightarrow (C, No) \quad s = 20\%, c = 66.67\%$
4. $(B,s) \rightarrow (C, Yes) \quad s = 30\%, c = 757\%$
5. $(B,t) \rightarrow (C, No) \quad s = 20\%, c = 66.677\%$
6. $(B,u) \rightarrow (C, No) \quad s = 20\%, c = 66.67\%$
7. $\{(A,p),(B,s)\} \rightarrow (C, Yes) \quad s = 20\%, c = 66.677\%$

Where s and c are support and confidence value respectively.

3.3 Classification phase

Structure of Backpropagation Neural Network that input Associative Classification to learning consists of three parts:

- Input layer is the first layer of a network that input Association Classification has to learn. In this layer, the input node is equal to the number of predicted attributes.
- Hidden layer is a neighbor layer with inputs nodes connected, that is each hidden node will be fully connected with each predicted attribute or input node, then the number of nodes in this layer can be determine by try-and-error method.
- Output layer is a layer that expresses the result of the network, each output node will be fully connected with each hidden node and number of output node is equal to number of class of data that we need the network to learn.

Generally Backpropagation Neural Network use sigmoid activation function, the output at each node can be defined in this equation.

$$outputj = \sigma(netj) = \frac{1}{(1-e^{-netj})} \quad \ldots (5)$$

By

$$netj = \sum_{i=0}^{n} w_{ji}\, x_i \quad \ldots (6)$$

Where, w_{ji} is the connection strength from the node i to node j and x_i is the output at node i.

Data creation will consider predicted attribute values that learn the consistency of predicted attribute value in each input and output node. The structure of Backpropagation neural network that is created with input as Class Association Rules is shown in Figure 2.

Table 2. Generating Association classification rules following proposed algorithm

1st pass	F_1	$< (\{(A,p)\},4),((C,Yes),3) >$
		$< (\{(A,q)\},3),((C,No),2) >$
		$< (\{(A,r)\},3),((C,No),2) >$
		$< (\{(B,s)\},4),((C,Yes),3) >$
		$< (\{(B,t)\},3),((C,No),2) >$
		$< (\{(B,u)\},3),((C,No),2) >$
CAR_1	$(A,p) \rightarrow (C,Yes)$	
	$(A,q) \rightarrow (C,No)$	
	$(A,r) \rightarrow (C,No)$	
	$(B,s) \rightarrow (C,Yes)$	
	$(B,t) \rightarrow (C,No)$	
	$(B,u) \rightarrow (C,No)$	
2nd pass	C_2	$< \{(A,p),(B,s)\},(C,Yes) >$
		$< \{(A,p),(B,t)\},(C,Yes) >$
		$< \{(A,q),(B,s)\},(C,Yes) >$
		$< \{(A,q),(B,t)\},(C,No) >$
		$< \{(A,q),(B,u)\},(C,No) >$
		$< \{(A,r),(B,t)\},(C,No) >$
		$< \{(A,r),(B,u)\},(C,Yes) >$
		$< \{(A,r),(B,u)\},(C,No) >$
	F_2	$< (\{(A,p),(B,s)\},3),((C,Yes),2) >$
CAR_2	$\{(A,p),(B,s)\} \rightarrow (C,Yes)$	
CARs	$CAR_1 \cup CAR_2$	

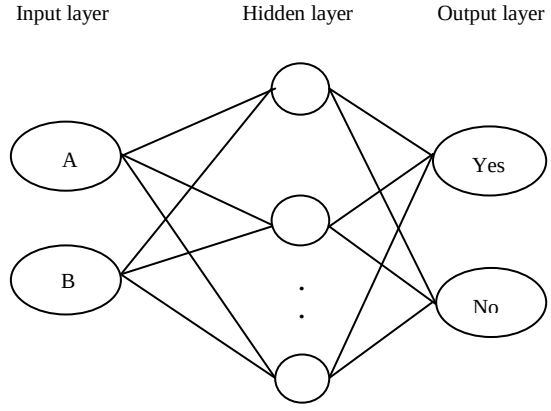

Figure 2: Structure of Backpropagation Neural Network from Class Association Rules

Suppose that data to learning is "A=p, B=s" and has class label "Yes". While inputting this data to a network, A with values p q, and r are encoded with numeric value as 1, 2, and 3 respectively. Similarly B with values s, t, and u are encoded with numeric value as 1, 2, and 3 respectively and now we will have the input is "1 1" and output will be "1 0" for this rule. This is shown in Figure 3.

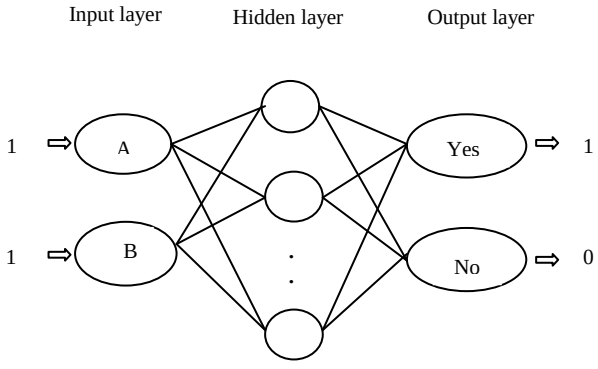

Figure 3: Structure of Backpropagation Neural Network for "$\{(A,p),(B,s)\} \rightarrow (C,Yes)$" as an input.

4. EXPRIMENTAL SETUP AND RESULTS

The Neural Network Associative Classification system is a user-friendly application developed in order to classify the relational dataset. This system is essentially a process consisting of three operations:

- Loading the relational (table) dataset,
- Let the user enters the two-threshold values support and confidence,

- Let the system perform the operations of calculating CAR_s (Class Association Rules),
- Let the system form a neural network, with the inputs as CARs set, by using Backpropagation algorithm,
- Let the system perform the testing on the network, and
- Let the user enter the unknown input and the system will provide the predicted class based on the training.

The discriminating feature of this system is that it uses the CARs to train the network to perform the classification task. So the system will render the efficient and accurate class based on the predicted attributes.

4.1 Data Description

For testing the proposed algorithm we will use data from UCI repository of Machine Learning Database with four sets. The details of the data are shown in Table 3.

The proposed Neural network Associative Classification system is implemented using MATLAB 7.5.0 (R2007b). The experiments were performed on the Intel® Core i3 CPU, 2.27 GHz system running Windows 7 Home Basic with 2.00 GB RAM.

Table 3: Dataset to testing

Dataset	No. of instances	No. of predicted attributes	Class	Attribute type
Lenses	24	4	3	Categorical, Integer
Iris	150	4	3	Categorical, real
Diabetes	768	8	2	Categorical, real
Glass	214	10	7	Categorical,real

4.2 The description of Neural network Associative Classification

First, finding the Class Association Rules (CAR) in relational datasets can be useful in many contexts. In general understanding the CARs relates the class attribute with the predicted attributes. Extraction of association rules was achieved by using the proposed algorithm in algorithm 1. The rules are extracted depending on the relations between the predicted and class attributes. The dataset in the Figure 4 is a segment of a lenses dataset from UCI repository of Machine Learning Database.

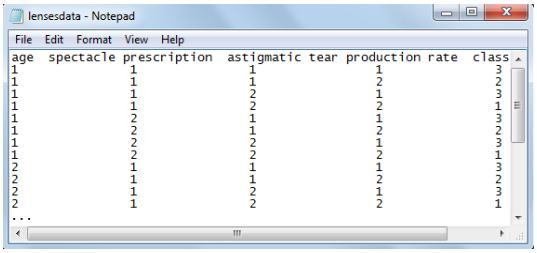

Figure 4. An excerpt from a lens dataset

There are four predicted attributes as age, spectacle prescription, astigmatic and tear production rate and a class attribute with three classes. The aim is to use the proposed algorithm to finds the relations between the features and represents them in CARs to form the input to the backpropagation neural network. The details of possible values at each attributes are given below:

- Age of the patient: (1) young, (2) pre-presbyopic, (3) presbyopic
- Spectacle prescription: (1) myope, (2) hypermetrope
- Astigmatic: (1) no, (2) yes
- Tear production rate: (1) reduced, (2) normal
- Class: (1 0 0) patient should be fitted with hard contact lenses, (0 1 0) the patient should be fitted with soft contact lenses, (0 0 1) patient should not be fitted with contact lenses.

Figure 5 shows the snapshot of the main screen of our project. The import button will load the dataset from the file into the MATLAB workspace. With the user define support 1% and confidence 50%, the extracted CARs are shown in Figure. 6. In Figure 6, s denotes support and c denotes confidence for that respective rule.

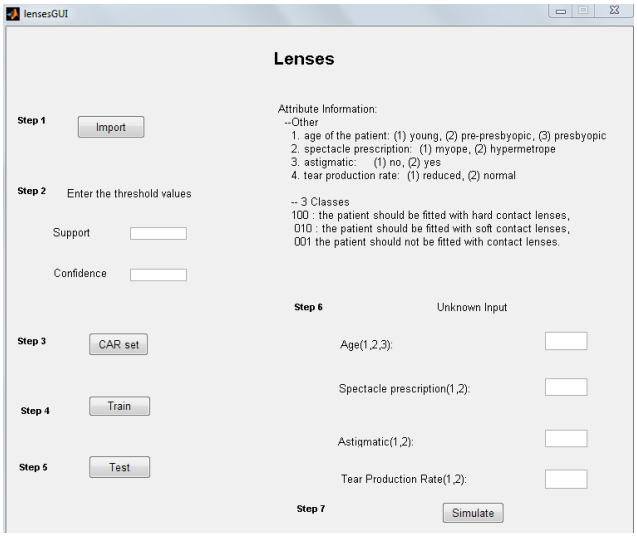

Figure 5: Neural Network Associative Classification System

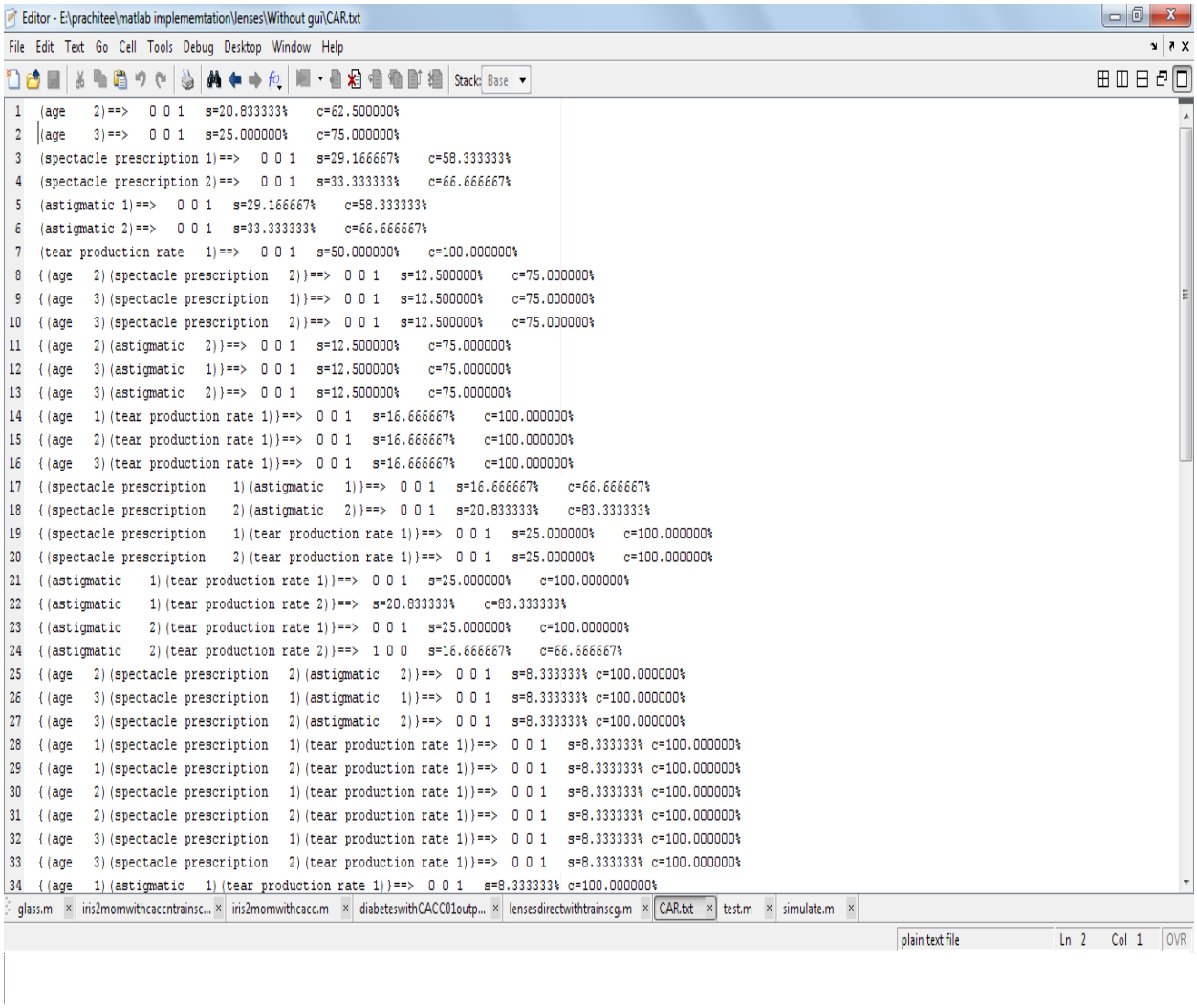

Figure 6: Class Association Rules generated.

For training the Backpropagation neural network CARs are used. To train the network , we have used the MATLAB function train(). In which 60% of data are used for training, 20% are used for testing and 20% for validation purpose. trainlm, learngdm and logsig are used as a training function, learning function and transfer function. Number of hidden layer nodes =74, learning rate = 0.001, momentum = 0.8 and learning will stop when error is less than 0.005 or epochs= 10,000 times. Figure 7 and 8 shows the snapshot of the training and testing respectively.

Now we have a trained network, which can more accurately predict the class based on the predicted attributes.

It can be observed that extracted CARs include the most important features and informative relations from the lenses dataset which will help neural network form the accurate classifier.

In this section, we have only shown the output of the Neural Network Associative Classification system related to the lenses dataset. Under the same mode we have build the simulation for diabetes, iris and glass datasets. The network complexities for these datasets are learning rate=0.001, training will stop when epochs= 15,000 or goal reaches 0.005 and the remaining are given in Table 4.

Table 4: Network complexities

Dataset	**Training function**	**Learning function**	**Transfer function**	**Number of hidden layer neurons**
Iris	traingdm	Learngdm	Tansig	301
Diabetes	trainscg	learngdm	logsig	2000
Glass	trainscg	Learngdm	Logsig	1000

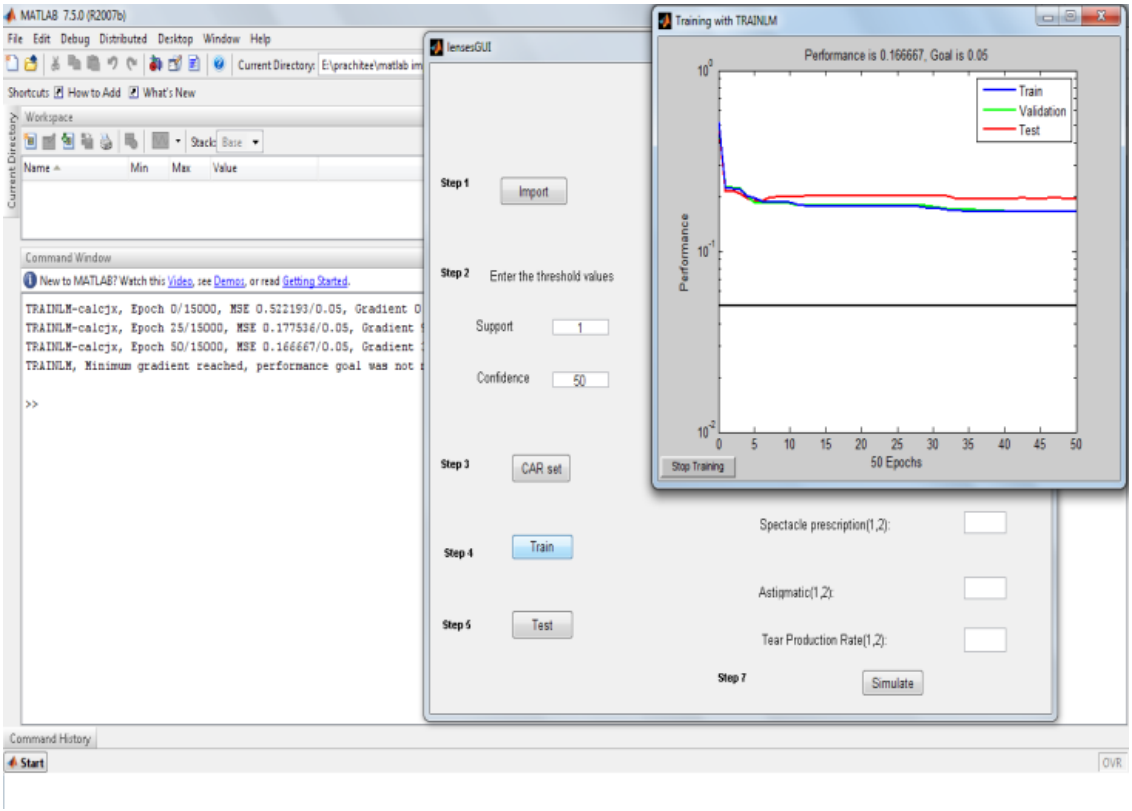

Figure 7: Training the network

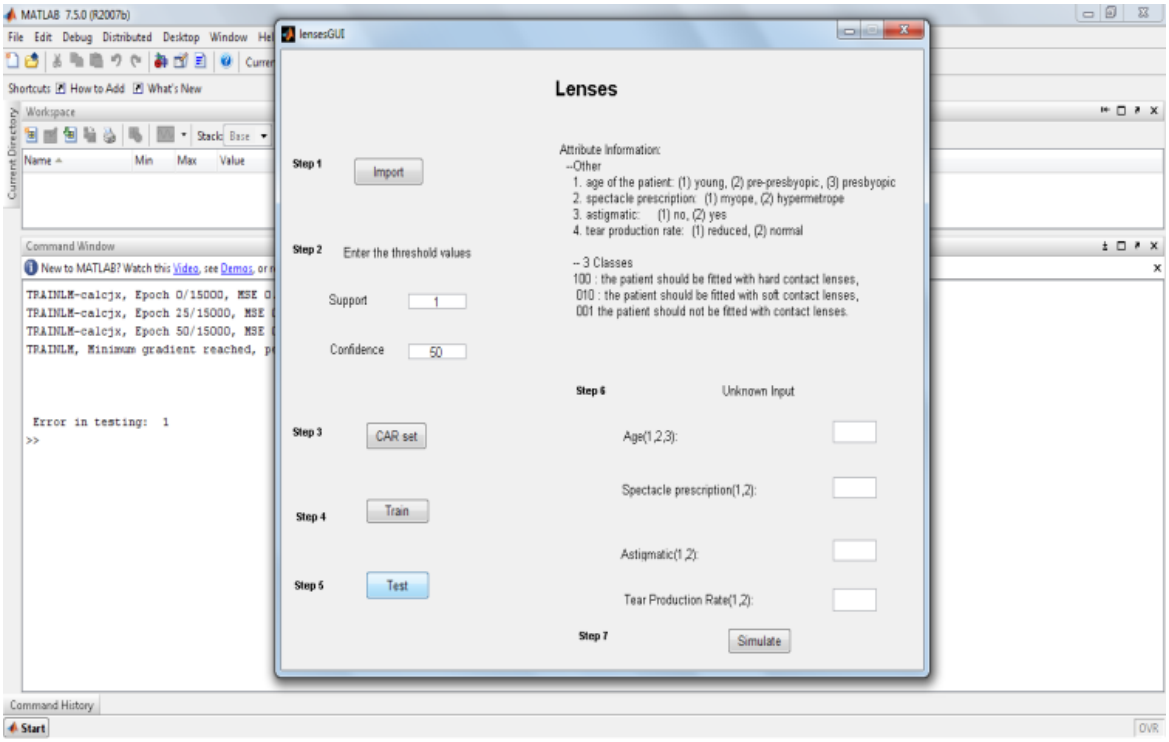

Figure 8: Testing the network

4.3 Result

To assay the performance of Neural Network Associative Classification system we will compare it with Classification Based Associations (CBA) on accuracy. Table 5 shows the accuracy of CBA and Neural network Associative Classification system on accuracy. Class association rules are obtained with minimum support= 1% and minimum confidence = 50%.

Table 5: Comparisons between the CBA system and the Neural Network Associative Classification system on accuracy

Dataset	**CBA**	**Neural network Associative Classification system**
Lenses	66.67	92.857
Iris	94.7	86.667
Diabetes	74.5	79.236
Glass	73.9	79.905

From the experimental results, Neural Network Associative Classification system will outperform than CBA in accuracy because neural network will learn to adjust weight from the input that is class association rules and will build a more accurate and efficient classifier.

5. CONCLUSION

Neural network Associative Classification system is for building accurate and efficient classifiers in order to improve its accuracy. The structure of networks reflects the knowledge uncovered in the previous discovery phase. The trained network is then used to classify unseen data.

Neural Network Associative Classification system presents a data classifier that more precisely match than old learning rules as shown in experimental result and adjusting weight according to the class association rules and best one is used to classify the data.

In future work, we intend to apply predictive apriori algorithm for finding the class association rules instead of applying apriori algorithm.

6. REFERENCES

[1] Apte, C. 1997. Data Mining:An industrial research perspective.IEEE Computational Science and Engineering April-June.

[2] Wu, X., Kumar, V., et.al. 2007.Top 10 algorithms in data mining. Knowledge Information System(2008) 14:1-37 DOI 10.1007/s10115-007-0114-2,Springer-Verlag London Limited.

[3] Likas, A., Vlassis, N., Verbeek, J.J. 2003.The global k-means clustering algorithm. Pattern Recognit., vol. 36, no. 2, pp. 451–461, Feb.

[4] Scholkopf, B., Smola, A.J., Muller, K. 1998. Nonlinear component analysis as a kernel eigenvalue problem. Neural Comput., vol. 10, no. 5, pp. 1299–1319, Jul.

[5] Tzortzis, G.F., Likas, A. 2009.The Global Kernel k-Means Algorithm for Clustering in Feature Space. IEEE transaction on neural networks, vol.20, no.7, July.

[6] Goldberger, J., Tassa, T. 2008. A hierarchical clustering algorithm based on the Hungarian method. Pattern Recognition Letters 29, 1632–1638, ScienceDirect.

[7] Agrawal, R., Imieilinski, T., Swami, A.1993. Mining Association Rules between sets of items in large databases. SIGMOD'93, pp. 207-216.

[8] Liu, B., Hsu, W., Ma, Y. 1998. Integrating Classification and Association rule mining. KDD'98, pp. 80-86.

[9] Schalkoff, R.J. Artificial Neural networks. McGraw Hill publication, ISE.

[10] Ni, X. 2008. Research of data mining based on neural networks. World academy of science, engineering and technology.

[11] Tsai, C.J., Lee, C., Yang, W. 2008. A discretization algorithm based on Class Attribute Contingency Coefficient. Inf. Sci., 178(3), 714-731.