A Fuzzy ART2 Model for Finding Association Rules in Medical Data

Yo-Ping Huang, Vu Thi Thanh Hoa, Jung-Shian Jau and Frode Eika Sandnes

***Abstract*—This paper describes a model that discovers association rules from a medical database to help doctors treat and diagnose a group of patients who show similar prehistoric medical symptoms. The proposed data mining procedure consists of two modules. The first is a clustering module that is based on a neural network, Adaptive Resonance Theory 2 (ART2), which performs affinity grouping tasks on a large amount of medical records. The other module employs fuzzy set theory to extract fuzzy association rules for each homogeneous cluster of data records. In addition, an example is given to illustrate this model. Simulation results show that the proposed algorithm can be used to obtain the desired results with a reduced processing time.**

I. Introduction

DATA mining is popularly referred to as knowledge discovery from data (KDD). It is the process of extracting desirable knowledge or interesting patterns from existing databases for specific purposes. Many types of knowledge and technology have been proposed for data mining. Among them, finding association rules from transaction data is the most commonly studied whelm.

An association rule is represented by $X \rightarrow Y$ where X and Y are a set of items. The rule means that the transaction records in a database that contain X also tend to contain Y. Many effective algorithms for mining association rules from large databases have been proposed [1], [2].

Over the last two decades, artificial neural networks (ANNs) have been developed for solving pattern classification problems and finding association rules [3]. In the medical domain, neural networks have been used as a diagnostic decision support system. For example, a supervised learning neural network was developed for leukemia diagnosis [4]. Other fault detection models based on abdicative network model, and combined fuzzy logic and neural network, were proposed in [5], [6]. In [7] a hybrid model of Adaptive Resonance Theory (ART) and fuzzy c-mean clustering for medical classification and diagnosis with missing features was developed.

Manuscript received Jan. 31, 2010. This work was supported by National Science Council, Taiwan under Grant NSC97-2221-E-027-034-MY3.

Y.-P. Huang is with the Department of Electrical Engineering, National Taipei University of Technology, Taipei, Taiwan 10608 R.O.C. (Tel: +886-2-27712171ext 1002 , e-mail: yphuang@ntut.edu.tw).

T. T. H. Vu is with the Department of Electrical Engineering, National Taipei University of Technology, Taipei, Taiwan 10608 R.O.C. (e-mail: t7318111@ ntut.edu.tw).

J.-S. Jau is with the Department of Electrical Engineering, National Taipei University of Technology, Taipei, Taiwan 10608 R.O.C. (e-mail: pikachu1982124@yahoo.com.tw).

F.E. Sandnes is with the Faculty of Engineering, Oslo University College, Oslo, Norway. (e-mail: frodes@hio.no).

The efficiency of combining fuzzy logic and neural network has been shown in many applications [8], [9], [10] but the problem of discovering association rules in data mining is an open research question. Thus, this study proposes a methodology using fuzzy ART2 to increase the efficiency when discovering association rules. The input patterns are first fuzzified using fuzzy logic. The fuzzified patterns are then clustered into groups by the ART2 neural network. The patterns in each group have similar properties that in turn allow us to find the association rules among them. This will significantly reduce the computational cost in finding the interesting rules.

We choose ART2 for clustering the patterns, because it is computationally effective and allows the user to easily control the degree of similarity of patterns placed on the same cluster. The inter- and intra-cluster differences among data indicate that ART2 clusters data according to Euclidean distance approximately.

In this paper, one case study is presented involving the grouping of patients who have undergone surgery for breast cancer. Those with similar symptoms are used to discover group association rules to assist the doctors in treatment and diagnosis.

The remainder of this paper is organized as follows. Section II gives an overview of related work, while the proposed approach is presented in Section III. Section IV summarizes the experimental results and discussions. Finally, the conclusion remarks are made in Section V.

II. Related Work

A. Fuzzy Neural Networks

The learning algorithms of ANNs can be divided into two categories: supervised and unsupervised. Supervised learning involves training instances with labeled outputs, which give feedback about how learning is progressing. This is akin to having a supervisor who can tell the agent whether or not it was correct. In unsupervised learning, the goal is unknown as there are no pre-determined categorizations.

ANN technology offers a decisive job in terms of summarizing, organizing, and classifying data. Requiring a few assumptions, it also helps identify patterns among input data and achieves a high degree of predictive accuracy [11]. The learning and recall procedures allow ANN to mimic the thinking models of human beings: through memorization and recall. In general, an ANN includes the following features: parallel processing, error tolerance, recall memory, and optimization solutions.

Both ANNs and fuzzy models have been applied in many areas [8], [9], [10], [12], [13]; each with its own advantages and disadvantages. Therefore, how to successfully combine these two approaches, ANNs and fuzzy modeling, has become an active area of research. Due to its lack of explanation ability, ANNs are unable to offer easily understandable explanations while the outputs are being generated [15]. Nevertheless, fuzzy models with its ability to articulately express knowledge and technology could compensate for the shortcomings of ANNs. A fuzzy neural network (FNN) system uses the ANN learning algorithm to produce parameters. It then adapts these parameters for optimization.

B. *Association Rules Mining*

Mining association rules from large databases is a core topic of data mining. It detects hidden linkages of otherwise seemingly unrelated data. These linkages are rules that overpass a preset threshold and are deemed interesting. Interesting rules allow actions to be taken based upon data patterns. They can also help make and justify decisions.

Association rules are defined in the form $\{X_1, X_2, \ldots, X_n\} \to Y$, in which Y may present in the transaction if $X_1, X_2, \ldots, X_n$ are all in the transaction. Notice the use of *may* to imply that the rule is only probable, not deterministic. The probability of finding Y in transactions with all $X_1, X_2, \ldots, X_n$ is called confidence. The threshold that a rule holds in all transactions is called support. The level of confidence that a rule must exceed is called interestingness.

There are different forms of association rules. The simplest type, Boolean association rules, only shows valid or invalid association. In our medical example, "Patients who have old age and large number of positive auxiliary nodes detected will die within 5 years" is a Boolean association rule.

The problem of discovering association rules can be generalized into two steps:

(1) Find all large (frequent) itemsets – A large itemset is a set of items that exceeds the minimum support.

(2) Generate rules from the large itemsets.

For step 1, the Apriori algorithm has been the mostly mentioned algorithm. Many modifications [16]-[17], e.g., speeding up and scaling up, of step 1 are about improving the Apriori algorithm by addressing its fallacy of generating too many itemsets. There are also algorithms that are not based on Apriori [18]-[20] but aim at addressing the issues of data mining efficiency.

Step 2 is mostly characterized by confidence and interestingness. Research has addressed different ways of generating rules [21] and alternative measures to interestingness [22].

III. Methodology

A. *Data Transformation*

The original raw data are collected from experiments and investigations. To use these data in each application, we have to transform them into a suitable form. Transformations are widely used in statistics to normalize data to standard form. Some common methods of re-expressing data are centering, standardizing and normalizing.

In this study the input patterns for ART2 are fuzzified and the clustered results are used to find the association rules among data. Assume that each input pattern contains n attributes ($X_1, X_2, \ldots, X_n$). Generally, such a function maps each attribute value to a real number in interval [0,1], $\mu_A(x): X \to [0,1]$, where μ_A is called the membership function of set A. For convenience, we denote the fuzzy membership function by μ_f , where the subscript f indicates the corresponding fuzzy set. The total number of fuzzy sets is denoted by S ($f \in [0,S]$). The number of membership functions for each attribute belongs to the interval of the attribute and the distribution of these attribute values.

The input pattern for ART2 can be represented by new patterns in form of membership degrees ($\mu_{1f_1}, \mu_{2f_2}, \ldots, \mu_{nf_n}$) (f_1 from 0 to S_1, f_2 from 0 to S_2, …, f_n from 0 to S_n). Otherwise, for describing the input patterns of the algorithm to make efficient in discovering the association rules, we use linguistic terms instead of membership degrees.

For example, for attribute X_1 in a pattern, X_1 is expressed by 3 fuzzy sets with linguistic terms: Small, Medium, Large (S, M, L) as shown in Fig. 1. If the value of X_1 is larger than c, X_1 will be represented by Large (L).

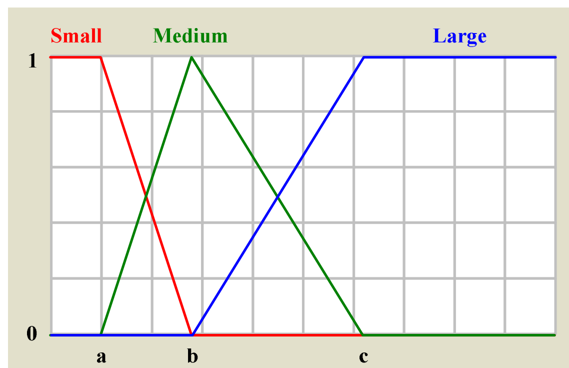

Fig. 1. Linguistic terms of X_1.

B. *Fuzzy ART2 Neural Network*

1) Network Structure: After the fuzzy inputs have been extracted, the proposed fuzzy neural network (FNN) called the fuzzy ATR2 is employed to automatically cluster the pattern data. The input and output relation of the proposed fuzzy ART2 can be described as follows:

Input layer: The input layer consists of units that are called short term memories (STM).

Weight layer: The weight layer consists of two kind of weights, i.e., bottom-up and top-down weights (b_{ij} and t_{ij}), that are called long term memories (LTM).

Output layer: The output layer is used to express the clustering results for the given data.

2) Learning Procedures:

The fuzzy ART2 learning procedure contains four steps as follows:

Step 1: Input the fuzzy vector into input layer.

Step 2: Calculate the distance between the bottom-up weights and fuzzy inputs. Find the shortest distance.

Step 3: If the shortest distance fails to the vigilance test, a new node is created with its weights equal to the fuzzy inputs. If a cluster wins the vigilance test, the centroid of the cluster is adjusted to adopt the new input.

Step 4: The process is repeated until all given data have been clustered into suitable clusters.

If all winners do not pass the vigilance parameter test, it is necessary to create a new cluster and add the corresponding weights. If the state is "resonance", the current fuzzy input is assigned to this cluster by modifying the corresponding weights.

The purposes of using the fuzzy ART2 before discovering the association rules include:

- Decentralize the volume of data from the originally large database to some subsets that contain the smaller number of patterns.
- Because patterns clustered to the same cluster possess the similar characteristics, the time taken to discover the association rules will be shorter than to the originally large data.

C. Algorithm for Finding Association Rules

After grouping all patterns of the original data into some clusters, the data mining algorithms are used to find the association rules from each cluster. As mentioned above, we need two 2 steps to find association rules. First, to improve the efficiency of the Apriori algorithm, a direct hashing and pruning (DHP) table [17] is used to reduce the size of candidate set by filtering any *k*-itemset out of the hash table if the hash entry does not reach a minimum support.

Then, rules are determined by interesting measures:

$$confidence\,(A \Rightarrow B) = \frac{\#\_tuples\_containing\_both\_A\_and\_B}{\#\_tuples\_containing\_A}$$

Rules become association rules when they have confidence larger than or equal to the minimum confidence.

IV. Experimental Results and Discussions

In this section, we use one benchmark study, Haberman's Survival problem, to illustrate the effectiveness of the proposed model. The data set is available from the UCI machine learning repository [23]. The Harbeman's Survival datasets consist of 306 samples. Each data sample constituted 4 attributes: age of patient, patient's year of operation, number of positive auxiliary nodes detected and survival status. Survival status attribute contains only two values that are one and two to indicate the patient survived 5 years or longer and the patient died within 5 years, respectively. Table I gives all linguistic terms for 4 attributes in each pattern for this study.

The vigilance value in ART2 model will dominate the clustering results that in turn affect the number of patterns clustered to each cluster. Depending on the database size and the properties of the data, it is hard to decide the number of suitable clusters before executing the clustering job. In this study, mid-size clusters, such as 3 to 7, are tested for the medical data. In case only few data were found in a cluster, a check is performed to see whether they are outliers or not. For example, when analyzing datasets we found that one pattern (83, 58, 2, 2) is an outlier numerically distant from the remaining data. Therefore, only 305 patterns are used in the simulations to avoid the overhead of processing this outlier pattern.

First, we carry out the methodology by fuzzifying the input patterns using the labels of membership functions of 3, 3, 3 and 1 for four attributes, respectively (case 1, the vigilance value is set to 3.6). These membership functions are illustrated in Fig. 2 for details. The cover range and the shape of membership functions may affect the results of association rules. Although those parameters can be further optimized by the genetic algorithm framework [14] they are not the goal of this research. In this example, all membership functions are empirically determined.

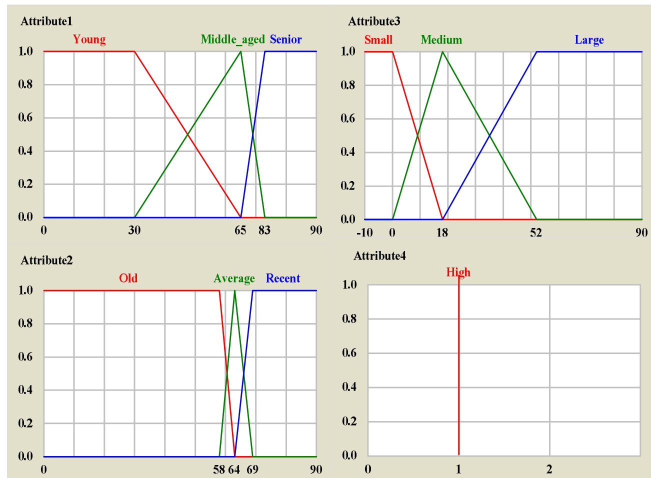

Fig. 2. Linguistic terms of four attributes for case 1.

After performing the fuzzy ART2, the data can be grouped into 3 clusters, in which clusters 1 to 3 contain 131, 115 and 59 patterns, respectively. We use DHP and interesting measures to find the association rules for each cluster. Because of the differences among the volumes of patterns in clusters, minimum support (min_sup) and minimum confidence (min_conf) can be set to different values.

So, we pro rata set minimum support and confidence in accordance with percentage of the number of each cluster. Fig. 3 illustrates two cases of the data ratio in the clusters. For example, cluster 1 in case 1 contains 43% of the 305 patterns, so we set both minimum support and confidence to 57%.

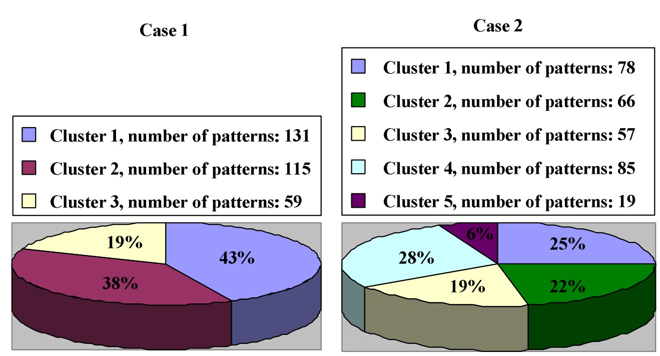

Fig. 3. Cluster results for case 1 and case 2.

The result of this case is compared with the case that only uses DHP to mine all 305 patterns without using Fuzzy ART2. Table II compares the results. Note that, min_sup and min_conf for mining all 305 patterns will be set lower than that of each cluster because each cluster after clustering has the similar properties.

From Table II, we can see that the results of association rules of combining 3 clusters (after mining separately clusters) and the results after mining from all patterns (without fuzzy ATR2) are the same but the latter needs more time to process.

The purpose of the other experiment is to find out the effect of the number of membership functions on the association rules. The number of membership functions for the patient age attribute and the number of positive auxiliary nodes attribute will be respectively changed to 5 membership functions as shown in Fig. 4 (case 2, the vigilance value is set to 3.4) and the vigilance parameter is set to the same as the case 1.

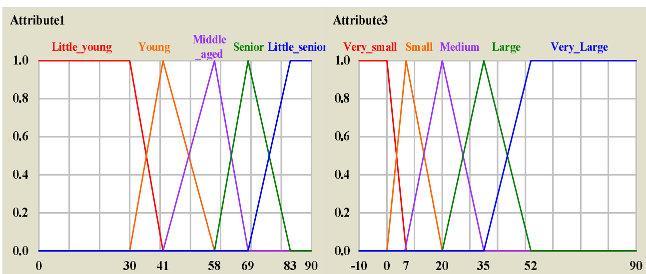

Fig. 4. Linguistic terms of attribute 1 and attribute 3 for case 2.

The results are presented in Table III and Table IV. In this case, the dataset is clustered into 5 clusters (larger than that of case 1) with the numbers of patterns respectively are 78, 66, 57, 85 and 19. And the same as the case 1, the association rules are similar with or without using fuzzy ART2.

Two cases take full advantage of fuzzy ART2 in finding association rules. It separates the dataset into some smaller groups to reduce the computational cost.

V. Conclusions

This study proposes a novel approach for finding association rules from medical data. The combination of fuzzy model and ART2 neural network is developed to cluster the fuzzified dataset into several groups with similar properties. The groups are then particularly exploited for finding the association rules. The approach shows that the more computationally efficiency can be obtained by reducing processing time. An application considered is to group the patients who had undergone surgery for breast cancer into groups that have the similar properties to use group association rules. The experimental results show that the discovered association rules are exactly the same with the conventional approach. The merit of this research is that the computational time is significantly reduced and the approach can be applied to a variety of domains.

References

[1] M. H. Margahny and A. A. Mitwaly, "Fast algorithm for mining association rules," in *Proc. of ICGST Int. Conf. on A*rtificial *Intelligent and Machine Learning*, Cairo, Egypt, Dec. 2005.

[2] A. Savasere, E. Omieccinski and S.B. Navathe, "An efficient algorithm for mining association rules in large databases," in *Proc. 21th Int. Conf. on Very Large Data Bases*, Zurich, Switzerland, pp.432-444, Sep. 1995.

[3] R. Li and Z. Wang, "Mining classification rules using rough sets and neural networks," *European Journal of Operational Research*, vol. 157, no. 2, pp.439-448, Sep. 2004.

[4] A.J.C. Sharkey, G.O. Chandorth and N.E. Sharkey, "A multi-net system for the fault diagnosis of a diesel engine," *Journal of Neural Computing and Application*, vol. 9, no. 2, pp.152-160, 2000.

[5] D. S. S. Lee, B. J. Lithgow and R. E. Morrison, "New fault diagnosis of circuit breakers," *IEEE Trans. on Power Delivery*, vol. 18, no. 2, pp. 454-459, Apr. 2003.

[6] J.M. Benitez, J.L. Castro and I. Requena, "Are artificial neural networks black boxes," *IEEE Trans. on Neural Networks*, vol. 8, no. 5, pp.1156-1164, Sep. 1997.

[7] E. Kolman and M. Margaliot, "Are artificial neural networks white boxes," *IEEE Trans. on Neural Networks*, vol. 16, no. 4, pp.844-852, July 2005.

[8] Y.-P. Huang, T.-W. Chang, L.-J. Kao and F.E. Sandnes, "Using fuzzy SOM strategy for satellite image retrieval and information mining," *Journal of Systemics, Cybernetics and Informatics,* vol. 6, no. 1, pp.1-6, Nov. 2008.

[9] S.A. Mingoti and J.O. Lima, "Comparing SOM neural network with fuzzy c-means, k-means and traditional hierarchical clustering algorithms," *European Journal of Operational Research*, vol. 16, no. 3, pp.1742-1759, Nov. 2006.

[10] R.J. Kuo and Y.T. Su, "Integration of ART2 neural network and fuzzy sets theory for market segmentation," *Int. Journal of Operations Research*, vol. 1, no. 1, pp.67-68, 2004.

[11] B.S. Ahn, S.S. Cho and C.Y. Kim, "The integrated methodology of rough set theory and artificial neural network for business failure prediction," *Expert Systems with Application*, pp. 65–74, 2000.

[12] G. Tayfur and V.P. Singh, "ANN and fuzzy logic models for simulating event-based rainfall-runoff," *Journal of Hydraulic Engineering*, vol. 132, no. 12, pp.1321-1330, Dec. 2006.

[13] A. Majumdar, P.K. Majumdar and B. Sarkar, "Application of an adaptive neuro-fuzzy system for the prediction of cotton yarn strength from HVI fibre properties," *Journal of the Textile Institute*, vol. 96, no.1, pp.55-60, Jan. 2005.

[14] T.-P. Hong, C.-H. Chen, Y.-C. Lee and Y.-L. Wu, "Genetic-fuzzy data mining with divide-and-conquer strategy," *Journal of Evolutional Computation*, vol. 12, no.2, pp.252-265, April 2008.

[15] R.J. Schalkoff, *Artificial Neural Networks*, McGraw-Hill, New York, USA, 1997.

[16] J. Gao, S. Li and F. Qian, "A method of improvement and optimization on association rules apriori algorithm," in *Proc. 6th World Congress on Intelligent Control and Automation*, Dalian, China, vol. 2, pp.5901-5905, June 2006.

[17] L. Ji, B. Zhang and J. Li, "A new improvement on apriori algorithm," in *Proc. of Int. Conf. on Computational Intelligence and Security*, vol. 1, pp.840-844, Nov. 2006.

[18] Y.-P. Huang, L.-J. Kao and F.E. Sandnes, "Efficient mining of salinity and temperature association rules from ARGO data," *Expert Systems with Applications*, vol. 35, no. 1-2, pp.59-68, Aug. 2008.

[19] L.V.S. Lakshmanan, C.K.-S. Leung and R.T. Ng, "The segment support map: scalable mining of frequent itemsets," in *Proc. of the ACM SIGKDD Explorations Newsletter*, vol. 2 no. 2, pp.21-27, Dec. 2000.

[20] J.S. Park, M.-S. Chen and P.S. Yu, "Using a hash-based method with transaction trimming for mining association rules," *IEEE Trans. on Knowledge and Data Engineering*, vol. 9, no. 5, pp.813-825, Oct. 1997.

[21] J. Mata, J.L. Alvarez and J.C. Riquelme, "Evolutionary computing and optimization: An evolutionary algorithm to discover numeric association rules," in *Proc. of the 2002 ACM symposium on Applied computing*, Madrid, Spain, pp.590-594, Mar. 2002.

[22] R. Natarajan and B. Shekar, "A relatedness-based data-driven approach to determination of interestingness of association rules," in *Proc. of the 20th Annual ACM Symposium on Applied Computing*, Santa Fe, New Mexico, USA, pp.551-552, Mar. 2005.

[23] http://archive.ics.uci.edu/ml/

TABLE I

ABBREVIATIONS OF THE PATIENT PARAMETERS

Attribute	No.	Parameter	Linguistic Term
Age of patient	1	Ly	Little_Young
	2	Yo	Young
	3	Ma	Middle_Aged
	4	Se	Senior
	5	Ls	Little_Senior
Patient's year of operation	1	Ol	Old
	2	Av	Average
	3	Re	Recent
Number of positive auxiliary nodes detected	1	Vs	Very_Small
	2	Sm	Small
	3	Me	Medium
	4	La	Large
	5	Vl	Very_Large
Survival status	1	Hi	High

TABLE II

THE ASSOCIATION RULES RESULTS FOR CASE 1.

Cluster 1	Cluster 2	Cluster 3	Combining Cluster 1, 2, 3	All Dataset
Sup. (57%) – Conf. (57%)	**Sup. (62%) – Conf. (62%)**	**Sup. (81%) – Conf. (81%)**		**Sup. (70%) – Conf. (70%)**
Hi <- Av (97.7, 97.7)	Av <- Hi (82.6, 96.8)	Ma <- Av (94.9, 98.2)	Hi <- Av (97.7, 97.7)	Av <- Hi (73.5, 96.9)
Av <- Hi (97.7, 97.7)	Hi <- Av (95.7, 83.6)	Av <- Ma (98.3, 94.8)	Av <- Hi (97.7, 97.7)	Hi <- Av (96.4, 73.9)
Ma <- Av (97.7, 100.0)	Me <- Hi (82.6, 100.0)	Me <- Av (94.9, 100.0)	Ma <- Av (97.7, 100.0)	Ma <- Hi (73.5, 100.0)
Av <- Ma (100.0, 97.7)	Hi <- Me (100.0, 82.6)	Av <- Me (100.0, 94.9)	Av <- Ma (100.0, 97.7)	Hi <- Ma (99.3, 74.0)
Me <- Av (97.7, 100.0)	Ma <- Hi (82.6, 100.0)	Me <- Ma (98.3, 100.0)	Me <- Av (97.7, 100.0)	Me <- Hi (73.5, 100.0)
Av <- Me (100.0, 97.7)	Hi <- Ma (100.0, 82.6)	Ma <- Me (100.0, 98.3)	Av <- Me (100.0, 97.7)	Hi <- Me (100.0, 73.5)
Ma <- Hi (97.7, 100.0)	Me <- Av (95.7, 100.0)	Me <- Av Ma (93.2, 100.0)	Ma <- Hi (97.7, 100.0)	Ma <- Av (96.4, 99.3)
Hi <- Ma (100.0, 97.7)	Av <- Me (100.0, 95.7)	Ma <- Av Me (94.9, 98.2)	Hi <- Ma (100.0, 97.7)	Av <- Ma (99.3, 96.4)
Me <- Hi (97.7, 100.0)	Ma <- Av (95.7, 100.0)	Av <- Ma Me (98.3, 94.8)	Me <- Hi (97.7, 100.0)	Me <- Av (96.4, 100.0)
Hi <- Me (100.0, 97.7)	Av <- Ma (100.0, 95.7)	Ma <- Av (94.9, 98.2)	Hi <- Me (100.0, 97.7)	Av <- Me (100.0, 96.4)
Me <- Ma (100.0, 100.0)	Ma <- Me (100.0, 100.0)	Av <- Ma (98.3, 94.8)	Me <- Ma (100.0, 100.0)	Me <- Ma (99.3, 100.0)
Ma <- Me (100.0, 100.0)	Me <- Ma (100.0, 100.0)	Me <- Av (94.9, 100.0)	Ma <- Me (100.0, 100.0)	Ma <- Me (100.0, 99.3)
Ma <- Av Hi (95.4, 100.0)	Me <- Hi Av (80.0, 100.0)	Av <- Me (100.0, 94.9)	Ma <- Av Hi (95.4, 100.0)	Ma <- Hi Av (71.2, 100.0)
Hi <- Av Ma (97.7, 97.7)	Av <- Hi Me (82.6, 96.8)	Me <- Ma (98.3, 100.0)	Hi <- Av Ma (97.7, 97.7)	Av <- Hi Ma (73.5, 96.9)
Av <- Hi Ma (97.7, 97.7)	Hi <- Av Me (95.7, 83.6)	Ma <- Me (100.0, 98.3)	Av <- Hi Ma (97.7, 97.7)	Hi <- Av Ma (95.8, 74.4)
Me <- Av Hi (95.4, 100.0)	Ma <- Hi Av (80.0, 100.0)	Me <- Av Ma (93.2, 100.0)	Me <- Av Hi (95.4, 100.0)	Me <- Hi Av (71.2, 100.0)
Hi <- Av Me (97.7, 97.7)	Av <- Hi Ma (82.6, 96.8)		Hi <- Av Me (97.7, 97.7)	Av <- Hi Me (73.5, 96.9)
Av <- Hi Me (97.7, 97.7)	Hi <- Av Ma (95.7, 83.6)		Av <- Hi Me (97.7, 97.7)	Hi <- Av Me (96.4, 73.9)
Me <- Av Ma (97.7, 100.0)	Ma <- Hi Me (82.6, 100.0)		Me <- Av Ma (97.7, 100.0)	Me <- Hi Ma (73.5, 100.0)
Ma <- Av Me (97.7, 100.0)	Me <- Hi Ma (82.6, 100.0)		Ma <- Av Me (97.7, 100.0)	Ma <- Hi Me (73.5, 100.0)
Av <- Ma Me (100.0, 97.7)	Hi <- Me Ma (100.0, 82.6)		Av <- Ma Me (100.0, 97.7)	Hi <- Ma Me (99.3, 74.0)
Me <- Hi Ma (97.7, 100.0)	Ma <- Av Me (95.7, 100.0)		Me <- Hi Ma (97.7, 100.0)	Me <- Av Ma (95.8, 100.0)
Ma <- Hi Me (97.7, 100.0)	Me <- Av Ma (95.7, 100.0)		Ma <- Hi Me (97.7, 100.0)	Ma <- Av Me (96.4, 99.3)
Hi <- Ma Me (100.0, 97.7)	Av <- Me Ma (100.0, 95.7)		Hi <- Ma Me (100.0, 97.7)	Av <- Ma Me (99.3, 96.4)

TABLE III
THE ASSOCIATION RULES RESULTS FOR CASE 2.

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Sup. (74%) – Conf. (74%)	**Sup. (79%) – Conf. (79%)**	**Sup. (81%) – Conf. (81%)**	**Sup. (72%) – Conf. (72%)**	**Sup. (94%) – Conf. (94%)**
Hi <- Ma (78.2, 95.1)	Vl <- Av (93.9, 95.2)	Vl <- Av (94.7, 94.4)	Vl <- Hi (92.9, 97.5)	Av <- Yo (100.0, 100.0)
Ma <- Hi (94.9, 78.4)	Av <- Vl (95.5, 93.7)	Av <- Vl (94.7, 94.4)	Hi <- Vl (96.5, 93.9)	Yo <- Av (100.0, 100.0)
Av <- Ma (78.2, 98.4)	Ma <- Av (93.9, 95.2)	Yo <- Av (94.7, 100.0)	Av <- Hi (92.9, 97.5)	Vl <- Yo (100.0, 100.0)
Ma <- Av (97.4, 78.9)	Av <- Ma (95.5, 93.7)	Av <- Yo (100.0, 94.7)	Hi <- Av (97.6, 92.8)	Yo <- Vl (100.0, 100.0)
Vl <- Ma (78.2, 98.4)	Ma <- Vl (95.5, 95.2)	Yo <- Vl (94.7, 100.0)	Yo <- Hi (92.9, 100.0)	Vl <- Av (100.0, 100.0)
Ma <- Vl (98.7, 77.9)	Vl <- Ma (95.5, 95.2)	Vl <- Yo (100.0, 94.7)	Hi <- Yo (100.0, 92.9)	Av <- Vl (100.0, 100.0)
Av <- Hi (94.9, 97.3)	Ma <- Av Vl (89.4, 94.9)	Yo <- Av Vl (89.5, 100.0)	Av <- Vl (96.5, 97.6)	Vl <- Yo Av (100.0, 100.0)
Hi <- Av (97.4, 94.7)	Vl <- Av Ma (89.4, 94.9)	Vl <- Av Yo (94.7, 94.4)	Vl <- Av (97.6, 96.4)	Av <- Yo Vl (100.0, 100.0)
Vl <- Hi (94.9, 98.6)	Av <- Vl Ma (90.9, 93.3)	Av <- Vl Yo (94.7, 94.4)	Yo <- Vl (96.5, 100.0)	Yo <- Av Vl (100.0, 100.0)
Hi <- Vl (98.7, 94.8)			Vl <- Yo (100.0, 96.5)	
Vl <- Av (97.4, 98.7)			Yo <- Av (97.6, 100.0)	
Av <- Vl (98.7, 97.4)			Av <- Yo (100.0, 97.6)	
Av <- Ma Hi (74.4, 98.3)			Av <- Hi Vl (90.6, 97.4)	
Hi <- Ma Av (76.9, 95.0)			Vl <- Hi Av (90.6, 97.4)	
Ma <- Hi Av (92.3, 79.2)			Hi <- Vl Av (94.1, 93.8)	
Vl <- Ma Hi (74.4, 98.3)			Yo <- Hi Vl (90.6, 100.0)	
Hi <- Ma Vl (76.9, 95.0)			Vl <- Hi Yo (92.9, 97.5)	
Ma <- Hi Vl (93.6, 78.1)			Hi <- Vl Yo (96.5, 93.9)	
Vl <- Ma Av (76.9, 98.3)			Yo <- Hi Av (90.6, 100.0)	
Av <- Ma Vl (76.9, 98.3)			Av <- Hi Yo (92.9, 97.5)	
Ma <- Av Vl (96.2, 78.7)			Hi <- Av Yo (97.6, 92.8)	
Vl <- Hi Av (92.3, 98.6)			Yo <- Vl Av (94.1, 100.0)	
Av <- Hi Vl (93.6, 97.3)			Av <- Vl Yo (96.5, 97.6)	
Hi <- Av Vl (96.2, 94.7)			Vl <- Av Yo (97.6, 96.4)	

TABLE IV
THE ASSOCIATION RULES FROM COMBINING 5 CLUSTERS FOR CASE 2.

Combining 5 clusters		All Datasets	
		Sup. (30%) – Conf. (30%)	
Hi <- Ma (78.2, 95.1)	Ma <- Hi Av (92.3, 79.2)	Hi <- Ma (40.5, 75.0)	Ma <- Hi Av (71.2, 41.7)
Ma <- Hi (94.9, 78.4)	Vl <- Ma Hi (74.4, 98.3)	Ma <- Hi (73.5, 41.3)	Vl <- Ma Hi (30.4, 98.9)
Av <- Ma (78.2, 98.4)	Hi <- Ma Vl (76.9, 95.0)	Av <- Ma (40.5, 96.0)	Hi <- Ma Vl (39.2, 76.7)
Ma <- Av (97.4, 78.9)	Ma <- Hi Vl (93.6, 78.1)	Ma <- Av (96.4, 40.3)	Ma <- Hi Vl (72.2, 41.6)
Vl <- Ma (78.2, 98.4)	Vl <- Ma Av (76.9, 98.3)	Vl <- Ma (40.5, 96.8)	Vl <- Ma Av (38.9, 96.6)
Ma <- Vl (98.7, 77.9)	Av <- Ma Vl (76.9, 98.3)	Ma <- Vl (96.7, 40.5)	Av <- Ma Vl (39.2, 95.8)
Av <- Hi (94.9, 97.3)	Ma <- Av Vl (96.2, 78.7)	Hi <- Yo (59.5, 72.5)	Ma <- Av Vl (93.1, 40.4)
Hi <- Av (97.4, 94.7)	Vl <- Hi Av (92.3, 98.6)	Yo <- Hi (73.5, 58.7)	Av <- Yo Hi (43.1, 96.2)
Vl <- Hi (94.9, 98.6)	Av <- Hi Vl (93.6, 97.3)	Av <- Yo (59.5, 96.7)	Hi <- Yo Av (57.5, 72.2)
Hi <- Vl (98.7, 94.8)	Hi <- Av Vl (96.2, 94.7)	Yo <- Av (96.4, 59.7)	Yo <- Hi Av (71.2, 58.3)
Vl <- Av (97.4, 98.7)	Yo <- Av Vl (89.5, 100.0)	Vl <- Yo (59.5, 96.7)	Vl <- Yo Hi (43.1, 97.7)
Av <- Vl (98.7, 97.4)	Vl <- Av Yo (94.7, 94.4)	Yo <- Vl (96.7, 59.5)	Hi <- Yo Vl (57.5, 73.3)
Yo <- Av (94.7, 100.0)	Av <- Vl Yo (94.7, 94.4)	Av <- Hi (73.5, 96.9)	Yo <- Hi Vl (72.2, 58.4)
Av <- Yo (100.0, 94.7)	Yo <- Hi Vl (90.6, 100.0)	Hi <- Av (96.4, 73.9)	Vl <- Yo Av (57.5, 96.6)
Yo <- Vl (94.7, 100.0)	Vl <- Hi Yo (92.9, 97.5)	Vl <- Hi (73.5, 98.2)	Av <- Yo Vl (57.5, 96.6)
Vl <- Yo (100.0, 94.7)	Hi <- Vl Yo (96.5, 93.9)	Hi <- Vl (96.7, 74.7)	Yo <- Av Vl (93.1, 59.6)
Yo <- Hi (92.9, 100.0)	Yo <- Hi Av (90.6, 100.0)	Vl <- Av (96.4, 96.6)	Vl <- Hi Av (71.2, 98.2)
Hi <- Yo (100.0, 92.9)	Av <- Hi Yo (92.9, 97.5)	Av <- Vl (96.7, 96.3)	Av <- Hi Vl (72.2, 96.8)
Av <- Ma Hi (74.4, 98.3)	Hi <- Av Yo (97.6, 92.8)	Av <- Ma Hi (30.4, 97.8)	Hi <- Av Vl (93.1, 75.1)
Hi <- Ma Av (76.9, 95.0)		Hi <- Ma Av (38.9, 76.5)	