JJMIE

Volume 8 Number 6, December 2014
ISSN 1995-6665
Pages 409 - 415

Jordan Journal of Mechanical and Industrial Engineering

Fuzzy Rules Extraction Based on Deterministic Data (Case Study: Bank's Customers Rating)

Ramin Sadeghian[* a], Behnam Gholamaliei [b], Leila Payandeh Peyman [c]

[a] *Corresponding author, Assistant Professor at Tehran Payam Noor University, Industrial Engineering Group, Tehran, Iran*

[b] *Master of Science student at Bu Ali Sina University, Engineering Department, Industrial Engineering Group, Hamedan, Iran*

[c] *Master of Science student at Bu Ali Sina University, Engineering Department, Industrial Engineering Group, Hamedan, Iran*

Received 20 April 2014 *Accepted 20 Nov 2014*

Abstract

Financial institutions and banks are the type of organizations whose customer rating is very valuable. In the present paper, two algorithms are proposed and compared that can extract some fuzzy rules from deterministic data. Each fuzzy rule may be used as a class of customers. In the first algorithm, a method is proposed based on both experiment and fuzzy theory. In the second algorithm a heuristic approach is proposed. Additionally, for more explanations a case study in a real bank is presented.

© 2014 Jordan Journal of Mechanical and Industrial Engineering. All rights reserved

Keywords: *Bank's Customers Rating, Fuzzy Rules Extraction, Deterministic Data Clustering, Linguistic Labels*.

1. Introduction

The knowledge acquisition from the existing data and information is essential for today's complicated world. The data mining and rules extraction from the raw data can be one method for knowledge acquisition [1]. In this paper, two algorithms are proposed to extract some fuzzy rules from the raw data. The extracted rules may be used in each field and situation. For instant, each rule may show a class of customers and consequently a service level of them.

The extracted rules can be applied for rating the customer of a specific bank. The customer rating can be very valuable for both customers and servicers. For example, top customers are considered as the customers of service level (1). These customers can receive the special services, while the service of other levels may be limited. However, the results of the present paper may also be used in any field, other than banking.

In many cases, banks use a set of raw and uncategorized information, which has been taken from the accounts of their customers. The information can be applied more effectively by processing.

By utilizing data mining methods, researchers and users will be able to extract some guidance from the given information. Of course, if there is a considerable ambiguity in the existing data, the deterministic data cannot be very helpful. Therefore, two proposed heuristic algorithms are able to extract fuzzy rules from the deterministic data. It is noteworthy that fuzzy rules and relations can describe uncertainty cases better than deterministic data [2-4].

Extracting the database of fuzzy rules from numerical data can be obtained by both tabling and clustering [5]. In the tabling approach, the input and output data are divided into several ranges, so a linguistic label is assigned to each range. Therefore, a table is shaped and the rules are determined [4]. Figure (1) shows a sample of tabulation of two-dimensional data.

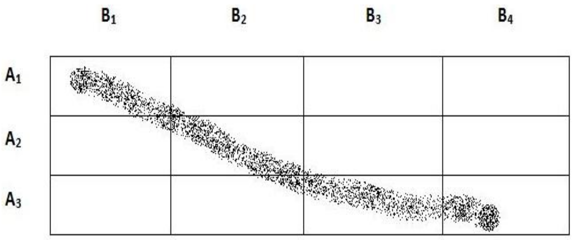

Figure 1. A sample of tabulation of two-dimensional data

The second approach utilizes clusters. The clustering approach places elements into a group based on some similarities and into different groups based on some dissimilarities [6].

In this approach, the data can be clustered by some specified clustering techniques, and each cluster may be considered as a rule. One of the advantages of this approach, of which the previous approach lacks, is that the fuzzy membership functions can be calculated easier [7-9]. Figure (2) shows a sample of two-dimensional deterministic data clustering.

[*] Corresponding author. e-mail: sadeghian@pnu.ac.ir.

 © 2014 Jordan Journal of Mechanical and Industrial Engineering. All rights reserved - Volume 8, Number 6 (ISSN 1995-6665)

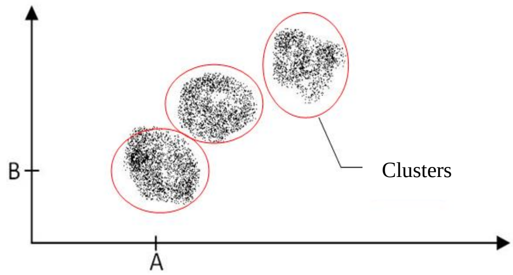

Figure 2. a sample of two-dimensional deterministic data clustering

Different approaches are used in describing the fuzzy sets such as "descriptive approach" and "approximate approach". The "descriptive approach" uses the terms as linguistic labels and its main objective is obtaining a qualitative form of the model [10, 11]. The "approximate approach" extracts fuzzy rules from data without any linguistic interpretation [12-14]. The fuzzy rules are the rules whose variables on left and right hand are fuzzy or linguistic.

Lan *et al.* provided a method to extract some nested fuzzy rules [15]. Ishibuchi *et al.* applied the genetic algorithm to select the best rule among many initial rules [16]. The genetic algorithm has also been applied in other studies [17, 18]. Some researchers used neural networks algorithm to extract rules [19, 20]. Novak *et al.* benefited from an algorithm, called GUHA, to find the appropriate linguistic labels [21]. The GUHA algorithm was among the first data mining methods and was presented in 1970's by Hajek [22].

Fuzzy rule extraction was presented in some applications. Kulluk *et al.* and Quteishat *et al.* presented some novel approaches for fuzzy rule extraction from neural networks [35, 36]. Wang *et al.* and Liu *et al.* proposed some methods for extracting fuzzy rules from fuzzy decision trees [37, 38]. Papageorgiou used fuzzy rule-extraction techniques in medical fields [39]. Also Chen *et al.* applied a rule extraction based approach in predicting derivative use for financial risk hedging by construction companies [40] and Chalaris introduced extraction of rules based on questionnaires [41].

2. Definitions

One of the assumptions of this paper is that each rule has multi-input sources (such as conditions) and only one output (as a consequence). The assumed rules are described as follows:

$$\begin{aligned} R_h : if\ \ x_1\ \ is\ \ A_h^1\ \ and\ \cdots\ and \\ x_p\ \ is\ \ A_h^p\ \ then\ \ y\ \ is\ \ B_h \end{aligned} \quad (1)$$

Where R_h $(1 \le h \le k)$ are the name of rules, x_j $(1 \le j \le p)$ are input variables, y is a output variable (consequence) and A_h^j, B_h are the linguistic labels of variables. Therefore the data can be displayed as following:

$$R_h \equiv (x_{1t}, x_{2t}, \ldots, x_{pt}, y_t)\ ;\ t = 1,2,\ldots,n \quad (2)$$

Note x_{1t}, x_{2t}, …, x_{pt}, y_t are the well-known values of variables.

The main idea of the present paper is that the placed similar data in a cluster may be described as a rule.

For this purpose, a suitable clustering algorithm (for example k-means) is used to cluster the existing data [1]. Next, the steps of the proposed algorithms (one of proposed algorithms) are done to find the extracted fuzzy rules. It is clear that the left hand side of the rules has n-1 variables and the right hand side of the rules has one variable only. So, according to Eq. (2), an n-dimensional space is obtained. Hence, when clustering is done, each cluster shows an n-dimensional sphere.

Since the obtained rules are formed from the set of linguistic labels, the rules can be defined as a Cartesian product similar to Eq. (3):

$$R: L_1 \times L_2 \times \ldots \times L_p \times T \quad (3)$$

Where Li is a set of labels used for the input variables and T is a set of labels used for the output variable.

Due to the nature of the data and clusters, the clustering techniques can be divided into four general categories:

- Deterministic data, deterministic clusters
- Deterministic data, fuzzy clusters
- Fuzzy data, deterministic clusters
- Fuzzy data, fuzzy clusters

In a clustering technique, if an element belongs to several clusters with different membership degrees, there is a fuzzy clustering. The fuzzy clustering methods can be performed by deterministic or fuzzy data. In this paper, two algorithms are presented to cluster the deterministic data as several fuzzy clusters.

3. Algorithms

3.1. Algorithm (I)

- Step 0: Begin
- Step 1: Determine the linguistic variables and labels for any data by experts - (in this case study by the bank's experts).
- Step 2: Use a suitable clustering technique to cluster the existing data.
- Step 3: Assign a fuzzy number to each member of clusters.
- Step 4: Compare the fuzzy numbers with their corresponding linguistic labels; and then select the label that is closest to the number.
- Step 5: According to the assigned labels in Step 4, extract and rewrite a fuzzy rule for each cluster.
- Step 6: End

The steps of the above algorithm are explained below.

Step 1: Define the linguistic variables

As mentioned in previous sections, fuzzy rules can be obtained using a set of linguistic labels. The definition of linguistic variables and labels may affect the efficiency of the algorithm. The arguments of linguistic variables are described by fuzzy numbers. Generally, a linguistic variable may be defined in four levels. In the first level, "variable name" (e.g., account balance), in the second level, "linguistic labels" (e.g., high, medium and low), in the third level, "membership function" of linguistic labels, and in the fourth level, "universal set."

Assume that the data are arranged using a table whose rows and columns represent the elements (e.g., bank customers) and the attributes of elements (e.g., customer account information), respectively. A linguistic variable is defined for each column. The first and second levels of linguistic variables may be defined by modeler, while their third and fourth levels should be defined by experts (e.g., expert in banking).

If you are not able to find an expert, you may use some another data or software similar to the report of Baturone et al. [5]. Table (1) shows a sample of the assignment of linguistic labels to fuzzy numbers.

Table 1. The assignment of linguistic labels into fuzzy number[23]

		Linguistic Variables	Fuzzy Numbers
	Sample1	Sample2	Sample3
Very Low			(0.0, 0.0, 0.1, 0.2)
Low		(0.0, 0.0, 0.2, 0.4)	(0.1, 0.25, 0.25, 0.4)
Medium	(0.4, 0.6, 0.6, 0.8)	(0.2, 0.5, 0.5, 0.8)	(0.3, 0. 5, 0. 5, 0.7)
High	(0.6, 0.8, 0.8, 1.0)	(0.6, 0.8, 1.0, 1.0)	(0.6, 0.75, 0.75, 0.9)
Very High			(0.8, 0.9, 1.0, 1.0)

Step 2: Data clustering

A cluster is a set of objects; its objects are similar to each other in some attributes, and are different from other objects in another attributes. For example, in the case of bank's customers, a cluster can include a set of customers whose account information have similarities, but are different from other customers. A good clustering technique may keep the mentioned feature. By utilizing this feature, the bank's managers may identify and forecast the behavior of customers to obtain better results [1]. A successful clustering is achieved through the execution of existing clustering algorithms. The use of different algorithms leads to different results, but there is not any approach for selecting the best algorithm [24]. In this paper, the K-means technique is used for clustering [1]. Despite that fact that the K-means technique was proposed for the first time over 50 years ago, it is still widely used for clustering. Simplicity, efficiency and empirical success are the main reasons for its popularity [25].

Definitely, any other clustering technique could also be used in this step. Clustering algorithms for high dimensional data were investigated in [26, 27]; time series data clustering were reviewed in [28]; the clustering problem in the data stream domain were studied in [29, 30]; and an overview of the approaches to clustering mixed data were given in [31].

Identifying the number of clusters in a data set (often labeled as *k*) is a fundamental issue in clustering analysis. To estimate the value of *k*, many studies have been presented and discussed in the literature [32].

Sun et al. [33] presented an algorithm based on the fuzzy k-means to determine the number of clusters automatically. It consists of a series of fuzzy k-means clustering procedures with the number of clusters varying from 2 to *k*. By investigating the clustering results for different values of *k*, the exact number of clusters is obtained in a given data set. The same approach for determining *k* is also used in the present paper.

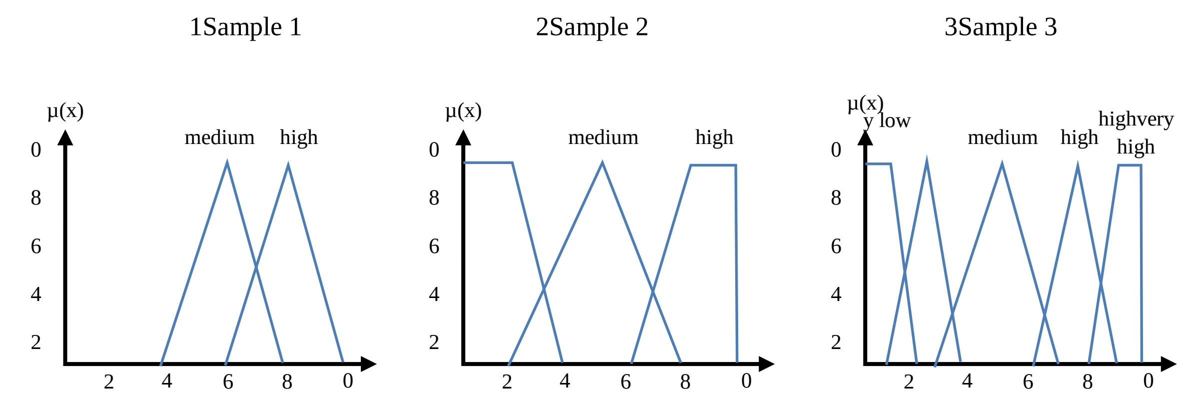

Figure 3. The assignment of linguistic labels into fuzzy numbers

Step 3: Determining the fuzzy numbers for the attributes of placed objects within clusters

Suppose the data is written in a table. The rows and columns of the table indicate the information of the bank's customers and the customer's accounts, respectively. Each attribute can be shown by a fuzzy number. In this paper, the triangular fuzzy numbers are used. In this step, for determining the fuzzy numbers the mean (μ) and standard deviation (σ) of data (data in each column) are applied. Hence, a triangular fuzzy number, such as (μ-σ, μ, $\mu + \sigma$), can be defined (see Figure (4)). Similarly, determine a triangular fuzzy number for each attribute (each column of table) and cluster.

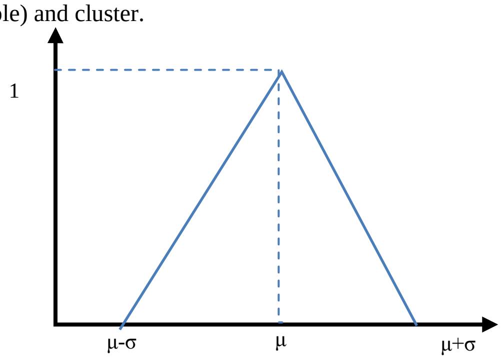

Figure 4. Determining triangular fuzzy number by μ, σ

Step 4: Select the appropriate label

In this step, you have to select the best label for any determined fuzzy number. A suitability criterion for a label means that its membership function is close to the desired fuzzy number. For measuring the closeness between two fuzzy numbers, there are several methods. In this paper two criteria as "Degree of Similarity" (*DOS*) and "Degree of Inclusion" (*DOI*) are applied [1]. The criteria can be defined as follows:

$$S(N,L) = \|N \cap L\| / \|N \cup L\| \quad (4)$$

$$I(N,L) = \|N \cap L\| / \|L\| \quad (5)$$

Where the symbol ||.|| shows the cardinality of a set, also N is an obtained fuzzy number in step (3) and L is the membership function of a linguistic label. $S(N,L)$ (degree of similarity between N and L) is used to determine the best L (as a linguistic label). L is selected if its *DOS* is the greatest. If the *DOS* of several linguistic labels are equal, *DOI* may be used to determine the best L. In this state, L is selected if its *DOI* is the greatest. Therefore, a linguistic label is assigned to each attribute.

Step 5: Rewriting the rules

In this step, the extracted rules can be rewritten like if–then rules. The number of rules may be equal to the number of clusters. For instance, a rule for cluster h can be written as follows:

$$R_h : if \;\; x_1 \;\; is \;\; A_h^1 \;\; and \;\; \cdots \;\; and \;\; x_p \;\; is \;\; A_h^p \;\; then \;\; y \;\; is \;\; B_h \quad (6)$$

Where x_1, x_2, ..., x_p, y are linguistic variables and Aih, Bh are the assigned labels in step (4). By rewriting the rules for all clusters, a rule base is obtained which can describe the behavior of the system in any conditions.

3.2. Algorithm (II)

- Step 0: Begin
- Step 1: Define a linguistic label and variable for each data by experts.
- Step 2: Cluster the deterministic data by an appropriate technique.
- Step 3: Calculate the "Degree of Dependency" to the cluster (*DoD*) for each attribute of objects within a cluster separately.
- Step 4: Find the equation of *DoD* changes in terms of "attribute values" for each attribute within a cluster separately.
- Step 5: Modify the obtained equations in step (4) according to the data concentrations.
- Step 6: Assign the suitable labels to the obtained equations in step (5).
- Step 7: Extract and rewrite the rules.
- Step 8: End.

Below, the above mentioned steps are explained.

Step 1: Defining linguistic variables

It is similar to step (1) in the first algorithm.

Step 2: Data clustering

It is similar to step (2) in the first algorithm.

Step 3: Calculating DoD

Each cluster includes a number of customers and their accounts have several attributes. A *DoD* can be calculated for each customer and attribute. If the number of customers is m and the number of attributes is n, the number of *DoDs* can be equal to $m \times n$. A *DoD* may be determined by equation (7):

$$DoD_{ijh} = 1 - \left| \frac{x_{ijh} - m_{jh}}{\max_i (x_{ijh} - m_{jh})} \right| \quad (7)$$

Where DoD_{ijh} is the degree of dependency to cluster h for each attribute j and customer account i. Also x_{ijh} is the value of attribute j and customer account i in cluster h and also *mjh* is the median of attribute j in cluster h. The symbol |.| shows an absolute function too.

Step 4: Finding the equation of *DoD* changes in terms of changes in the value of attributes

In discrete cases, the value of the attributes can be represented by fuzzy sets that include a number of ordered pairs, where the first element shows the value of an attribute and the second element shows *DoD*.

In the present paper, the values are continuous, so the horizontal axis shows the values of an attribute and the vertical axis shows *DoD*. A regression equation may be fitted for the existing data. This curve can be considered as a fuzzy number to assign the best linguistic label to the data (see Figure 5)

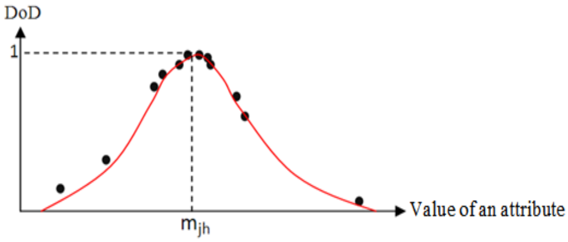

Figure 5. A sample of fitting a curve on data

Step 5: Improving the obtained equations according to the data concentrations

Generally, the values of attributes are concentrated in specified areas of domain (Figure 5). Also the existence of some useless data in first and end of domain can widen the domain and curve. The widened domain and curve can lead to some errors in the selection of appropriate linguistic labels and a reduction in the accuracy of the results. For this purpose, exclude a range of data from the beginning and end of domain. If the length of the old domain is d_1 and the length of the new domain is d_2, the multiplying proportion d_2/d_1 in the length of the graph horizontal axis may reduce the width of the chart and increase its focus on areas where the density of data is higher. This action will increase the accuracy of linguistic labels selection.

Step 6: Find the appropriate labels

Select a label that is closer to the obtained equation in step (5). The current step is similar to step (4) of the first algorithm, so you can implement it by equations (4) and (5).

Step 7: Rewriting the rules

This step is similar to step (5) of the first algorithm.

4. Case Study

The customer relationship management and customer requirements management in banks and financial institutions are very important. So rating customers based on the analyzed criteria can help organizations to present more favorable services. Data mining [14] and the presented algorithms in previous sections can be used as tools for evaluating, predicting customer behavior and customers rating in terms of some criteria. In this paper, it is shown how to implement the proposed algorithms. For this purpose, the information of 2500 accounts from a bank in Iran is studied [34]. The afore-mentioned information is clustered by the *k*-means method and is considered 5 clusters corresponding to 5 levels of customers rating.

In this case study, the following criteria of customers account information are considered:

1. **average:** The average of daily balance for each customer during one year.
2. **val-cheq:** The average amount of returned checks for each customer during one year.
3. **tran-bed:** Total debtor turnover of each customer during one year.
4. **tran-bes:** Total creditor turnover of each customer during one year.
5. **rem-bes:** The sum of the creditor balance of each customer during one year.
6. **remained:** The remaining creditor balance of each customer at the end of year.

The criterion "tran-bed" can be considered as a consequent of other criteria, so it is located in right hand side of the rules. The steps of the first algorithm can be performed as follows:

Step (1) Defining the linguistic variables (see Figures (6) to (11)).

Step (2) Clustering data by *k*-means method: since 5 clusters have less error, so 5 clusters are considered.

Steps (3-5) Selecting appropriate labels, their calculations, rules extraction and their rewriting.

Finally the five rules are obtained as follows:

- R1: If *(remained = VL)* and *(tran_bes = L)* and *(rem_bes = L)* and *(val_cheq = VS)* and *(average = VL)* $\Longrightarrow$ *(tran_bed = L)*
- R2: If *(remained = VL)* and *(tran_bes = S)* and *(rem_bes = L)* and *(val_cheq = VS)* and *(average = VL)* $\Longrightarrow$ *(tran_bed = S)*
- R3: If *(remained = VL)* and *(tran_bes = M)* and *(rem_bes = L)* and *(val_cheq = VS)* and *(average = VL)* $\Longrightarrow$ *(tran_bed = M)*
- R4: If *(remained = M)* and *(tran_bes = VS)* and *(rem_bes = S)* and *(val_cheq = S)* and *(average = M)* $\Longrightarrow$ *(tran_bed = VS)*
- R5: If *(remained = VL)* and *(tran_bes = M)* and *(rem_bes = L)* and *(val_cheq = VS)* and *(average = VL)* $\Longrightarrow$ *(tran_bed = M)*

Third and fifth rules are similar, so the fifth rule is removed, so only the first four rules can be considered.

Also by the second algorithm, five rules are obtained as following:

- R_1: If *(remained = VL)* and *(tran_bes = L)* and *(rem_bes = L)* and *(val_cheq = VS)* and *(average = VL)* $\Longrightarrow$ *(tran_bed = L)*
- R_2: If *(remained = L)* and *(tran_bes = S)* and *(rem_bes = L)* and *(val_cheq = VS)* and *(average = L)* $\Longrightarrow$ *(tran_bed = S)*
- R_3: If *(remained = VL)* and *(tran_bes = M)* and *(rem_bes = L)* and *(val_cheq = VS)* and *(average = L)* $\Longrightarrow$ *(tran_bed = M)*
- R_4: If *(remained = M)* and *(tran_bes = VS)* and *(rem_bes = S)* and *(val_cheq = VS)* and *(average = M)* $\Longrightarrow$ *(tran_bed = VS)*
- R_5: If *(remained = VL)* and *(tran_bes = M)* and *(rem_bes = L)* and *(val_cheq = VS)* and *(average = VL)* $\Longrightarrow$ *(tran_bed = L)*

Note that the extracted fuzzy rules from each of the two algorithms are different.

By extracting the fuzzy rules according to the needs of each bank, the customers can be rated into several levels. A bank can present services to its customers based on the determined levels. For example, the customers of the first cluster can be considered as high level customers (first level) because their account balance and turnover is large and their returned check rate is low; these customers can be useful for a bank. Therefore, a bank can prepare more services for this level of customers. The reward and punishment can also be applied similarly for other levels.

5. Conclusions

In this paper, two algorithms were presented to extract fuzzy rules from deterministic data. The proposed algorithms were performed in a case study on bank customers. One of the applications of fuzzy rules for banks may be rating their customers. Each customer in each level can receive some services depending on its level.

For investigating the validity of the algorithms, the information of 300 new customer accounts was extracted. The data were compared with the obtained rules from previous algorithms. The criteria values as input data and the value of "tran-bed" as output data were considered. The calculations show that the first algorithm 95.7% and the second algorithm 92.7% confirm the views of experts. Of course the confirmation of both algorithms is acceptable.

The advantage of the above methods over the other methods is that they can reduce the time and the required effort for extracting rules.

References

[1] Ghazanfari, M., Rezaei, M., *An introduction to fuzzy sets theory*, 1st ed. Tehran: Iran University Science and Technology; 2006.

[2] Araki, S., Nomura, H., Hayashi, I., and Wakami, N., "*A self-generating method of fuzzy inference rules*", Fuzzy Engineering toward Human Friendly Systems, (1991), 1047-1058.

[3] Ichihashi, H., "*Iterative fuzzy modeling and a hierarchical network*", IFSA, Bruselas, (1991), 49-51.

[4] Wang, L., and Mendel, J.M., "*Generating fuzzy rules by learning from exampl*es", IEEE Trans, Systems Man Cybernet, Vol. 22 (1992), 1414-1427.

[5] Baturone, I., Moreno-Velo, F.J., Gersnoviez, A., "*Identifying fuzzy systems from numerical data with Xfuzzy*", EUSFLAT – LFA, (2005).

[6] Walt Van, C.M., Barnard, E., "*Data characteristics that determine classifier performance*", In Proceedings of the Sixteenth Annual Symposium of the Pattern Recognition Association, South Africa, 2006, 160-165.

[7] Delgado, M., Gomez-Skarmeta, A.F., Martin, F., "*A methodology to model fuzzy systems using fuzzy clustering in a rapid-prototyping approach*", Fuzzy Sets and Systems, Vol. 97 (1998), 287-30.

[8] Setnes, M., "*Supervised Fuzzy Clustering for Rule Extraction*", IEEE TRANSACTIONS ON FUZZY SYSTEMS, Vol. 8 (2000) No. 4.

[9] Klawonn, F., Kruse, R., "*Constructing a Fuzzy Controller from Data*", Fuzzy Sets and Systems, Vol. 85 (1997), 177-193.

[10] Pedrycz, W., "*An identification algorithm in fuzzy relational systems*", Fuzzy Sets and Systems, Vol. 13 (1984), 153-167.

[11] Peng, X., and Wang, P., "*On generating linguistic rules for fuzzy models*", Uncertainty and Intelligent Systems, (1988), 185-192.

[12] Nie, J., and Linkens, D.A., "*Learning control using fuzzified self-organizing radial basis function network*", IEEE Trans, Fuzzy Systems, Vol. 1 (1993), 280-287.

[13] Nomura, H., Hayashi, I., and Wakami, N., "*A self-tuning method of fuzzy control by descent method*", IFSA, Bruxelles, (1991), 155-158.

[14] Simpson, P.K., and Jahns, G., "*Fuzzy Min-Max neural networks for function approximation*", Proc. IEEE lnternat. Conf. on Fuzzy Systems, (1993), 1967-1972.

[15] Abe, S., and Lan, M.S., "*A classifier using fuzzy rules extracted directly from numerical data*" Proc. 2nd IEEE Int'l Conf. on Fuzzy Systems (FUZZ-IEEE '93), San Francisco, CA, (1993), 1191-1198.

[16] Ishibuchi, H., Nakashima, T., Murata, T., "*Three-objective genetic-based machine learning for linguistic rule extraction*", Information Sciences, Vol. 136 (2001), 109-133.

[17] Cordon, O., del Jesus, M.J., Herrera, F., "*Genetic learning of fuzzy rule-based classification systems cooperating with fuzzy reasoning methods*", International Journal of Inteligent Systems, Vol. 13 (1998), 1025-1053.

[18] Ishibuchi, H., Nakashima, T., "*Improving the performance of fuzzy classifier systems for pattern classification problems with continuous attributes*", IEEE Transactions on Industrial Electronics, Vol. 46 (1999), 1057-1068.

[19] Abe, S., Lan, M.S., Thawonmas, R., "*Tuning of fuzzy classifier derived from data*" , International Journal of Approximate Reasoning, Vol. 14 (1996), 1-24.

[20] Mitra, S., De, R.K., Pal, S.K., "*Knowledge-based fuzzy MLP for classification and rule generation*", IEEE Transactions on Neural Networks, Vol. 8 (1997), 1338-1350.

[21] Novak, V., Perfilieva, I., Dvorak, A., Chen, G., Wei, Q., Yan, P., "*Mining pure linguistic associations from numerical data*" International Journal of Approximate Reasoning, Vol. 48 (2008), 4–22.

[22] Hajek, P., "*The question of a general concept of the guha method*", Kybernetika, Vol. 4 (1968), 505–515.

[23] Boroushaki,S., Malczewski,J., "*Using the fuzzy majority approach for GIS-based multicriteria group decision-making*" , Computers & Geosciences, Vol. 36 (2010), 302–312.

[24] Grekousis, G., Hatzichristos, T., "*Comparison of two fuzzy algorithms in geodemographic segmentation analysis: The Fuzzy C-Means and GustafsoneKessel methods*", Applied Geography, Vol. 34 (2012).

[25] Anil K. Jain, "*Data clustering: 50 years beyond K-means*", Pattern Recognition Letters, Vol. 31 (2010).

[26] Kriegel, H.P., Kroger, P., Zimek, A., "*Clustering high-dimensional data: a survey on subspace clustering, pattern-based clustering and correlation clustering*", ACM Transactions on Knowledge Discovery from Data, Vol. 3 (2009), No. 1.

[27] Bai, L., Liang , J.Y., Dang, C.Y., Cao, F.Y., "*A novel attribute weighting algorithm for clustering high-dimensional categorical data*", Pattern Recognition, Vol. 44 (2011) No. 12.

[28] Liao, T.W., "*Clustering of time series data survey*", Pattern Recognition, Vol. 38 (2005) No. 11.

[29] Aggarwal, C.C., Han, J.W., Wang , J.Y., Yu, P.S., "*A framework for clustering evolving data streams*", in: Proceedings of the 29th VLDB Conference, Berlin, Germany, 2003.

[30] Cao, F.Y., Liang, J.Y., Bai, L., Zhao, X.W., Dang, C.Y., "*A framework for clustering categorical time-evolving data*", IEEE Transactions on Fuzzy Systems, Vol. 18 (2010) No. 5.

[31] Hunt, L., Jorgensen, M., "*Clustering mixed data*", WIREs Data Mining and Knowledge Discovery, Vol. 1 (2011) No. 4.

[32] Mirkin, B., "*Choosing the number of clusters*", WIREs Data Mining and Knowledge Discovery, Vol. 1 (2011) No. 3.

[33] Sun, H.J., Wang, S.R., Jiang, Q.S., "*FCM-based model selection algorithms for determining the number of clusters*", Pattern Recognition, Vol. 37 (2004) No. 10.

[34] Tarokh, Mohammad J. and Sharifian, Kobra, "*Clustering the bank customers and its analysis by the customer value pyramid*", 6th IIEC conference, 2009.

[35] Sinem Kulluk, Lale Özbakır, Adil Baykasoğlu, "Fuzzy DIFACONN-miner: A novel approach for fuzzy rule extraction from neural networks", Expert Systems with Applications, Vol. 40 (2013) No. 3, 938-946.

[36] Anas Quteishat, Chee Peng Lim, "A modified fuzzy min–max neural network with rule extraction and its application to fault detection and classification", Applied Soft Computing, Vol. 8 (2008) No. 2, 985-995.

[37] Xianchang Wang, Xiaodong Liu, Witold Pedrycz, Lishi Zhang, "Fuzzy rule based decision trees", Pattern Recognition, Vol. 48 (2015) No. 1, 50-59.

[38] Xiaodong Liu, Xinghua Feng, Witold Pedrycz, "Extraction of fuzzy rules from fuzzy decision trees: An axiomatic fuzzy sets (AFS) approach", Data & Knowledge Engineering, Vol. 84 (2013), 1-25.

[39] Elpiniki I. Papageorgiou, "A new methodology for Decisions in Medical Informatics using fuzzy cognitive maps based on fuzzy rule-extraction techniques", Applied Soft Computing, Vol. 11 (2011) No. 1, 500-513.

[40] Jieh-Haur Chen, Li-Ren Yang, Mu-Chun Su, Jia-Zheng Lin, "A rule extraction based approach in predicting derivative use for financial risk hedging by construction companies", Expert Systems with Applications, Vol. 37 (2010) No. 9, 6510-6514.

[41] Manolis Chalaris, Ioannis Chalaris, Christos Skourlas, Anastasios Tsolakidis, "Extraction of Rules based on Students' Questionnaires", Procedia - Social and Behavioral Sciences, Vol. 73 (2013), 510-517.