Original Article

A New Method of Rule Extraction from Quantitative Data Using the Theory of Rough Sets and Fuzzy Logic

Narges Khazravi *, Seyyed Majid Alavi

Department of Mathematics, Islamic Azad University, Hamedan Branch, Iran.

Abstract

The extraction of rules from unclassified quantitative data is one of the works done in the fuzzy-rough logic. When the rough and fuzzy algorithms are blended, the quantitative data will be shown as fuzzy variables. Then, the rules are extracted from the approximation from below and above. The investigation of the membership functions of each feature in this process is necessary and time-consuming. Monotonic membership function will be introduced in this paper. The investigation of this monotonic function is substituted for the investigation of the membership functions of the information system features. The proposed method is a simple and efficient method.

Keywords: rough set, fuzzy set, approximation from below, approximation from above

INTRODUCTION

The theory of rough sets was presented by Pawlak (1982); since then, this theory has expanded. It is complicated for human to interpret raw information. Hence, raw information needs to be analyzed so that it can be understandable for human. Here, the fuzzy logic is used to change the quantitative data into fuzzy variables; in the final step, rules are extracted from data [1-3]. This paper is organized such that the concepts of the rough set are summarized in section 2, section 3 investigates the fuzzy-rough set, rule extraction from quantitative data is elaborated in section 4, the proposed method is explained in section 5, section 6 compares the proposed method and the previous method, and conclusion is mentioned in section 7.

Theory of rough sets

This theory is a powerful tool for arguing the ambiguities and a process for rule extraction without losing the major data. Theory of rough has been taken into consideration due to the following reasons:

1. Analysis of the hidden information from data
2. No need for additional data information

Consider the information system of I=(U,S). U is a nonempty finite set of samples and S is a nonempty finite set of features such that $a = U \to V_a$ and for each $a \in S$, V_a is a set of values that the attribute a can take; an equivalence relation is considered for each $B \subseteq S$:

$$IND(B) = \{(x, y) \in U^2 | \forall a \in B, a(x) = a(y)\}$$

If $(x \& y) \in IND(B)$, x and y can be recognized by the features of the set B. The equivalence class of B is shown with $[x]_B$. Consider $X \subseteq U$. X can be estimated by using the available information in the set of P features; it is obtained by making approximation from below and above for set X.

$$\bar{B}X = \{x | [x]_B \cap X \neq \emptyset\}$$

$$\underline{B}X = \{x | [x]_B \subset X\}$$

The binary relation of $< \overline{BX}$ and $\underline{BX} >$ is known as the rough set. The decision system of $(\boldsymbol{U}$ and $\boldsymbol{S} \cup d)$ is a special kind of information system which is used for classification or prediction. d is the attribute of the decision and $[x]_{\boldsymbol{d}}$ are decision classes.

Address for correspondence: Mrs Narges Khazravi, Department of Mathematics, Islamic Azad University, Hamadan Branch, Iran.
E-mail: mkhazravi@ymail.com

How to cite this article: Khazravi, N., Alavi, S. M. A New Method of Rule Extraction from Quantitative Data Using the Theory of Rough Sets and Fuzzy Logic. Arch Pharma Pract 2020;11(S1):23-7.

The fuzzy-rough sets

In practice, there are some databases in which the values of some features are accurate and some of them are real. The theory of rough faces a problem in these cases. This theory cannot specify whether the values of two attributes are similar to each other or not. Discretization of database is one of the ways to face this problem. This strategy changes the values of the features to accurate values; however, a part of the information is lost. The fuzzy-rough theory is the basic way to solve this problem. In fact, the combination of the fuzzy and rough set mostly focuses on the fuzzification of approximation from above and below. For this purpose, the two following rules are used:

1. Set X is generalized to the fuzzy set in U and this is made possible that objects can belong to the set with different membership degrees.
2. Objects are placed at equivalence classes based on their similarities and are allowed to belong to more than one class. In fact, an equivalent approximation is made between objects and fuzzy relation R in model X which relate the similarity degree to each pair of objects.

Rule extraction from quantitative data by the theory of rough sets and fuzzy logic:

Rules are extracted from a set of quantitative data that consist of n objects and each object is of m features. The set of data is first placed at equivalence classes based on the decision feature. It is complicated for human to interpret this raw information. Hence, data are changed into linguistic variables in the fuzzy theory and for each feature in the fuzzy set, an appropriate membership function that indicates the overlapping degree of set X and the intended equivalence class is considered.

$$\mu_X^{B_k(x)} = \frac{card(B_k(x) \cap X)}{card B_k(x)}$$

In the next step, the features' degree of dependence (independence) on the intended feature is obtained as follows.

$$K = \gamma(B \text{ and } D) = \Sigma_{x\epsilon U/D} \frac{|\underline{B}(X)|}{|U|}$$

$$K = \frac{|POS_B(D)|}{|U|}$$

$$POS_B(D) = \cup_{x\epsilon U/D}\, \underline{B}(X)$$

If K=1, then, the feature with the defined relation is dependent on the intended feature; if K<1, the feature with the defined relation dependent on the intended feature with K degree. The approximation from above and below is obtained for each feature in relation to the decision feature. It is calculated as follows $B_k \subseteq B$

$$\underline{B}_k(\mathrm{X}) = \left\{\left(B_k(x) \text{ and } \mu_{\mathrm{X}}^{B_k(x)}\right) |x\epsilon U \text{ and } C(B_K(x) \text{ and } \mathrm{X}) \leq \beta\right\}$$

$$1 \leq k \leq |B(x)|$$

$$\bar{B}_k(X) = \left\{\left(B_k(x) \text{ and } \mu_X^{B_k(x)}\right) |x\epsilon U \text{ and } C(B_K(x) \text{and} X) \leq 1-\beta\right\}$$

$$1 \leq k \leq |B(x)|$$

The degree of accuracy of the decision feature in each cluster is calculated from the approximation from above and below as mentioned below and shown in the following table:

If $S = (U \text{ and } A)$ $B \subseteq A$ and $X \subseteq U$, then, the trust degree is calculated from the following formula:

$$\alpha_B(X) = \frac{|\underline{B}(X)|}{|\overline{B}(X)|}$$

If $\alpha_B(X) = 1$, then, set X is accurate and if $\alpha_B(X) < 1$, set X is a rough set.

At last, rules are extracted based on what mentioned above.

The proposed method

The proposed method is based on the monotonic membership function. This membership function includes all membership functions defined for each feature. The monotonic membership function is a piecewise-defined function. In fact, instead of working with several membership functions [2-4], only one monotonic fuzzy function is used that covers the membership functions. It is simpler to work with this monotonic fuzzy function.

The algorithm of the proposed method:

1. The input of the fuzzy membership function algorithm and data of the table
2. Converting the membership functions of the linguistic variables to a fuzzy monotonic membership function for each feature
3. Substitution of the table

Comparison of the proposed method and the previous method [2]

This comparison is shown by an example.

Consider the table of quantitative data:

Table 1. Set of quantitative data

Object	Systolic pressure(sp)	Diastolic pressure(Dp)	
$obj^{(1)}$	122	80	**N**
$obj^{(2)}$	155	90	**H**
$obj^{(3)}$	130	92	**N**

$obj^{(4)}$	87	68	**L**
$obj^{(5)}$	165	93	**H**
$obj^{(6)}$	150	100	**H**
$obj^{(7)}$	95	75	**L**

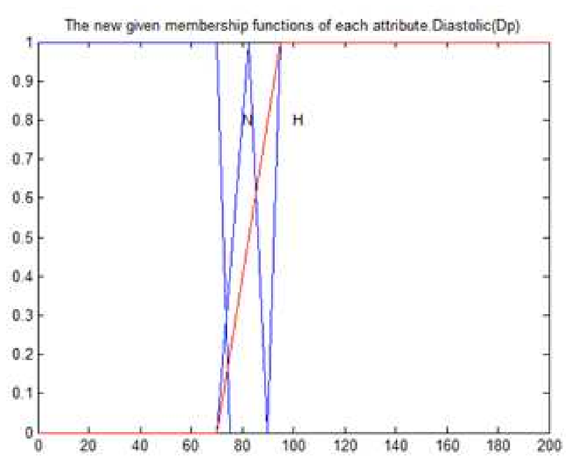
Quantitative data are converted to fuzzy variables and the membership functions are specified for fuzzy variables. In the next step, the equivalence classes are determined for decision feature and the set of features. Then, the degree of dependence of each feature is obtained in relation to the decision feature and the approximation from above and below is calculated [1-5].

In the proposed method, the membership functions of SP and DP are monotonic piecewise-defined functions and defined as follows:

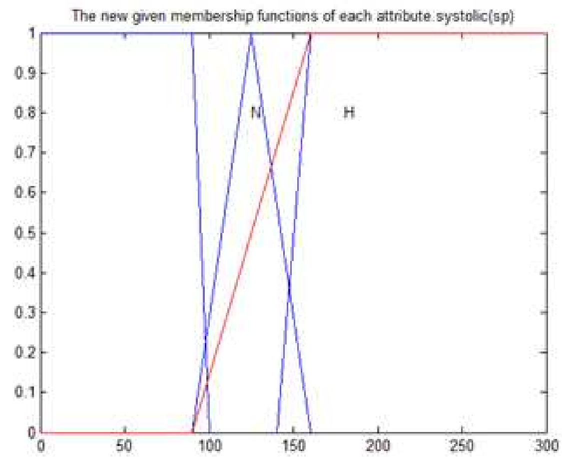
$$f_{sp}(x) = \begin{cases} 0 & 90 \leq x \\ Y = \dfrac{1}{70}x - \dfrac{9}{7} & 90 < x < 160 \\ 1 & x \geq 160 \end{cases}$$

$$f_{DP}(x) = \begin{cases} 0 & x \leq 70 \\ Y = \dfrac{1}{25}x - \dfrac{14}{5} & 70 < x < 95 \\ 1 & x \geq 95 \end{cases}$$

It is shown in the two following figures that the monotonic membership function has substituted for fuzzy membership functions.



The figure of the membership functions of the feature (DP)



The figure of the membership functions of the feature (SP)

Table 2. Data set of the monotonic function

Object	Systolic pressure(sp)	Diastolic pressure(Dp)	BlOOd pressure(Bp)
$obj^{(1)}$	0.4	0.4	**N**
$obj^{(2)}$	0.9	0.8	**H**
$obj^{(3)}$	0.5	0.8	**N**
$obj^{(4)}$	0	0	**L**
$obj^{(5)}$	1	0.9	**H**
$obj^{(6)}$	0.8	1	**H**
$obj^{(7)}$	0.07	0.2	**L**

Data clustering has been used in the proposed method and MATLAB software has been used for showing random data.

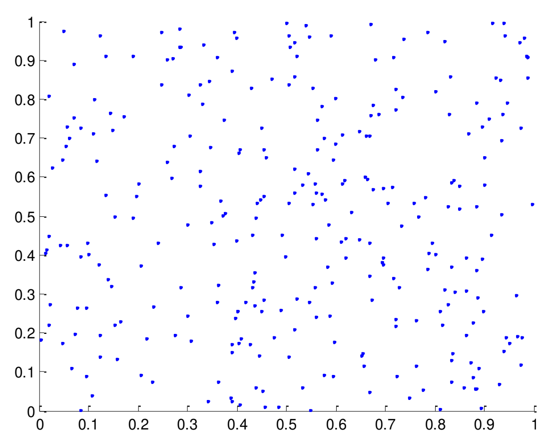


Figure of data clustering

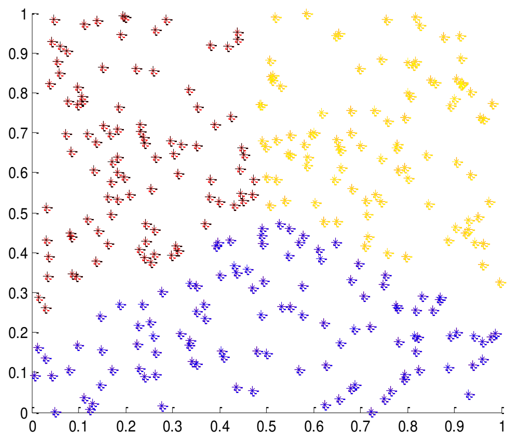


Figure of data distribution

The cluster centers are shown in the following table and MATLAB software has been used for showing random data.

Table 3. Cluster centers

	C	
Cluster 1	0.8	0.8
Cluster 2	0.1	0.03
Cluster 3	0.4	0.4

Each of the data in Table 2 has been placed in one of the clusters as shown in the following table.

Table 4. Belonging of data to clusters

IDX	Systolic pressure(sp)	Diastolic pressure(Dp)	BlOOd pressure(Bp)
3	0.4	0.4	**N**
1	0.9	0.8	**H**
1	0.5	0.8	**N**
2	0	0	**L**
1	1	0.9	**H**
1	0.8	1	**H**
2	0.07	0.2	**L**

The clusters of each feature are specified based on the following relations:

$$U/\{SP\} = \{N_{SP}\text{and}Z_{SP}\}$$

$$U/\{DP\} = \{N_{DP} \text{ and } Z_{DP}\}$$

$$U/\{SP \text{ and } DP\} = \otimes \left\{U \setminus IND_{(\{d\})}; d \in \{SP \text{ and } DP\}\right\}$$

$$U/\{SP \text{ and } DP\} = \{N_{SP} \cap N_{DP} \; and \; N_{SP} \cap Z_{DP} \; and \; Z_{SP} \cap N_{DP} \; and \; Z_{SP} \cap Z_{DP}\}$$

In this example, the decision feature has been divided into three equivalence classes as follows:

$$U/\{BP\} = \{(\{obj^{(1)} \; \& \; obj^{(3)}\} \; \& \; N) \text{ and } (\{obj^{(2)} \& obj^{(5)} \& \; obj^{(6)}\} \& \; H)\}$$

$$\text{and } (\{obj^{(4)} \& \; obj^{(7)}\} \& L)\}$$

According to the decision feature, the degree of accuracy of the obtained clusters and the features' dependence on the intended feature are obtained.

In clustering based on U / P, if $P = \{DP\}$ or $P = \{SP\}$ or $P = \{SP \text{ and } DP\}$ and the clustering of the decision feature is considered as $Q = \{BP\}$, the table of degree of accuracy will be obtained as follows:

$$POS_P(Q) = \cup \, \underline{P}_{x \in U/Q}(X) = \underline{P}(X_1) \cup \underline{P}(X_2) \cup \cdots \cup \underline{P}(X_n)$$

Table 5. Degree of accuracy

	SP	DP	SP & DP
Cluster 1	1	0	1
Cluster 2	1	1.3	1
Cluster 3	1	0	1

Table for the degree of dependence of the features on the decision feature is as follows:

Table 6. Degree of dependence

Dependence	BP
SP	1
DP	1.3
SP & DP	1

The following rules are extracted from the above calculations:

1. The size of BP can be exactly estimated according to the size of SP.
2. The size of BP cannot be estimated according to the size of DP.
3. Regarding SP and DP

CONCLUSION:

In this study, the quantitative data were converted to fuzzy variables using fuzzy logic and rough theory. Then, the rules were extracted from the approximation from below and above, the degree of dependence of a monotonic function which was substituted for membership functions, and clustering that was done by MATLAB software. This method is very simpler and faster than investigating all membership functions.

References

1. Chang Y, Liu H. Semi-supervised classification algorithm based on the KNN. In2011 IEEE 3rd International Conference on Communication Software and Networks 2011 May 27 (pp. 9-12). IEEE.
2. Hong TP, Wang TT, Wang SL. Mining fuzzy β-certain and β-possible rules from quantitative data based on the variable precision rough-set model. Expert Systems with Applications. 2007 Jan 1;32(1):223-32.
3. Mac Parthaláin N, Jensen R. Fuzzy-rough set based semi-supervised learning. In2011 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2011) 2011 Jun 27 (pp. 2465-2472). IEEE.
4. Z.Pawlak Rough sets. International journal of computer & information sciences. 1982 Oct 1;11(5):341-56.
5. Z.Pawlak, Rough Sets Rough Sets -Theoretical Aspect of Reasoning about Data Theoretical Aspect of Reasoning about Data, Kluwer Academic Pubilishers, 1991.