The use of fuzzy logic in classification

Amaury Caballero^{1,*}

¹Department of Electrical & Computer Engineering
Florida International University. Miami, USA

Abstract. When determining the degree of coincidence of any multi-feature obtained information, received in the form of a fuzzy vector, to a pre-established known pattern, two general steps should be followed. The first step is to eliminate the features that have little or no effect to the final results and to maintain only those that will influence the pattern recognition. This step could be defined as the classification process and is imperative for the simplification of the problem. One example of classification that could considerably reduce system costs is when using sensors distributed along an industrial process to manage information at a central location. Several methods could be used for classification, such as statistical methods, rough sets, fuzzy logic or information theory. The second step is to find out the correlation between the received fuzzy vector and the vector defining the known pattern using the previously selected features. For this part, the use of fuzzy logic is extremely convenient. The present work analyzes some of the methods used for classification and pattern recognition based on concrete and practical examples

1. Introduction

When analyzing an information system or a database, frequently we face problems like attributes redundancy, missing or diffuse values, which are due in general to noise, and missing partial data. Several approaches for minimizing the number of attributes necessary to represent the desired category structure by eliminating redundancy have been developed. The lack of data or complete knowledge of the system makes, when developing a model, a practically impossible task using conventional means. This lack of data can be attributed to sensors failure, or simple due to incomplete system information. At last, diffuse values could be related to noise or imprecise measurements from sensors. In many applications, the information is collected from different sensors, which are corrupted by noise and outliers. Different methods have been presented [1, 2, 3, 4, 5, 6]. The present work is devoted to the use of fuzzy logic only as instrument for solving this problem.

For dealing with interval-valued information systems one frequently used procedure is the discretization. Yee Leung et al. [1] have presented a very useful method for obtaining rules, based on rough sets for discriminating the minimum number of attributes, and giving a first approach in the objects classification. Some aspects of the method are reproduced here.

2. Attribute reduction

Procedure:

1. Table preparation: From the original table, a new one presenting the minimum and maximum values for each parameter and for each object is generated.

2. Define the misclassification rates (α_{ij}^k). The region of coincidence or misclassification rate of the two attributes may vary from 0 when there is not coincidence at all, to 1 when the two attributes coincide completely.

In general, the probability that objects in class u_i are misclassified into class u_j according to attribute k can be represented by [1]:

$$\alpha_{ij}^k = 0 \text{ if } [l_i^k, u_i^k] \cap [l_j^k, u_j^k] \equiv \alpha; \quad (1)$$

$$\alpha_{ij}^k = \min\{(u_i^k - l_j^k, u_j^k - l_i^k)/(u_i^k - l_i^k), 1\}, \text{ if } [l_i^k, u_i^k] \cap [l_j^k, u_j^k] \neq \emptyset. \quad (2)$$

Where α is the permissible misclassification rate which gives the permitted error in the classification.

Note that in general

$$\alpha_{ij}^k \neq \alpha_{ji}^k.$$

3. Define α_{ij} as the error that class u_i being misclassified into class u_j in the system k .

$\alpha_{ij} = \min \{ \alpha_{ij}^k : k \leq m \}$, where m is the maximum number of attributes

4. Find out the maximum mutual classification error between classes (β_{ij}^k)

$$\beta_{ij}^k = \max \{ \alpha_{ij}^k, \alpha_{ji}^k \} \quad (3)$$

where $\beta_{ij}^k = \beta_{ji}^k$.

5. For each pair of classes, find out the permissible misclassification rate between classes u_i & u_j in the system k

$$\beta_{ij} = \min \beta_{ij}^k \text{ for } 1 \leq k \leq m$$

Let's define the permissible misclassification rate between classes as α . If $\beta_{ij} \leq \alpha$, there must exist an attribute A_k so that, by using A_k , the two classes U_i and U_j can be separated within the permissible misclassification rate α .

3. Attribute Reduction Using Information Theory

Several methods using information theory have been developed. The presented method here is developed in detail in [6]. The general procedure is as follows: It results logical to think that when the two classes coincide for some parameter k , the information obtained from this parameter for discriminating between classes i and j is 0, and that it increases as the coincidence diminishes. This leads to the representation of this information, from Shannon and Hartley definition, using a logarithmic scale. Here the logarithm is used to provide the additivity characteristic for independent uncertainty.

For expressing it with logarithms base 10, it is given as

$$I_{ij}^k = -(\log \beta_{ij}^k) \text{ [Hartley]}$$

Similarly, the minimum information required for the classification between two classes i and j for an attribute k and permissible misclassification rate α is given by

$$I_{\alpha}^k = L = -\log \alpha \text{ [Hartley].}$$

If $I_{ij}^k \geq I_{\alpha}^k$, the two classes can be separated using the attribute k . An algorithm has been created for solving this task.

4. Fuzzy logic classification

The major task of fuzzy-based pattern classification is the extraction of knowledge from numerical data to build a rule base, which will permit the classification of new data members. One way of calculating the similarity is given below [7, 8]:

Let $P^*(X)$ be a group of fuzzy sets with $A_i \neq 0$, and $A_i \neq X$. Defining two fuzzy sets from this family of sets, $A, B \in P^*(X)$, the expression

$$(A, B) = (A \bullet B) \wedge (\overline{A \oplus B})$$

describes the degree of similarity of the two sets A and B . When the approaching degree approaches unity, the two sets will have higher degree of similarity, and an approaching degree near zero implies the distinction between the two fuzzy sets.

Defining a new data sample B with m fuzzy attributes, the approaching degree concept can be applied to compare the new data pattern

$$B = \{B_1, B_2, \dots, B_m\},$$

with some known data pattern A_i . Each of the known patterns A_i is characterized by the same m attributes and given by

$$A_i = \{A_{i1}, A_{i2}, \dots, A_{im}\},$$

where $i = 1, 2, \dots, k$ describes k -patterns.

For each of the known k -patterns, the approaching degree expression is given by

$$(B, A_i) = \sum_{j=1}^m \omega_j (B_j, A_{ij}),$$

where ω_j is a normalizing weighting factor, taken unitary in this work.

Then sample B is closest to sample A_j if

$$(B, A_j) = \max_{1 \leq i \leq k} \{(B, A_i)\}$$

The collection of fuzzy sets

$$B = \{B_1, B_2, \dots, B_m\},$$

can be reduced to a collection of crisp singletons, $B = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ where each sample ($\mathbf{x}_i$) is a vector of features,

$$\mathbf{x}_i = \{x_{i1}, x_{i2}, x_{i3}, \dots, x_{im}\}$$

and the above mentioned weighted equation can be expressed as [12]

$$\mu_{A_i}(\mathbf{x}) = \sum_{j=1}^m \omega_j \cdot \mu_{A_{ij}}(\mathbf{x}_j)$$

If it is considered that the different attributes are not of equal importance, then relative weights must be calculated.

The method consists in creating a fuzzy model for each parameter based on the previous obtained information, and for each new received sample find out the degree of compatibility (IC), given by the previous equation, with each class. The maximum obtained from this expression is the class to which the received information belongs to. The previous concepts are applied in the following examples.

5. Abalone database

This database [9] has been created in order to predict the age of abalone from physical measurements. Here it is used only like one example, and the obtained conclusions not necessarily reflect all the possibilities that could be obtained in the classification.

The authors selected 150 entries from the database and divided them in three groups: Group I, from 6 to 10 years old, Group II, from 11 to 15 years old, and Group III, from 16 to 20 years old. The attributes are as follows: A- Length, B- Diameter, C- Height, D- Whole Weight, E- Shucked Weight, F- Viscera Weight, and G- Shell Weight. The elaborated information is presented in Table 1.

From Table 2, showing the maximum mutual classification error between classes in logarithmic form, it is clear that is not possible to discriminate between groups II and III, using the given attributes

$\sum_{k=1}^m I_{ij}^k < I_{\alpha}$) and the given error of misclassification

$\alpha = 0.2$. It is possible to discriminate between groups I and III, using attributes D or A. Finally, there is a possibility of discriminating between groups I and II, using fuzzy logic and practically all the attributes. But, in general as per the obtained results in the example, the conclusion could be that the selected attributes in this case are not useful for determining the abalone age, using intervals of 5 years. As per Table 2, the discrimination between all the selected classes can be done with an error of 0.9, which is in general not acceptable.

Table 1. Attributes for Each Class

Attribute	Group I			
	x_{av}	σ	Min	Max
A	0.46	0.09	0.27	0.59
B	0.37	0.06	0.2	0.47
C	0.12	0.02	0.07	0.18
D	0.54	0.24	0.12	1.06
E	0.22	0.37	0.05	0.49
F	0.12	0.06	0.03	0.27
G	0.17	0.07	0.07	0.34
Attribute	Group II			
	x_{av}	σ	Min	Max
A	0.57	0.06	0.38	0.72
B	0.46	0.05	0.31	1.06
C	0.16	0.03	0.10	0.24
D	1.04	0.37	0.28	2.55
E	0.41	0.15	0.11	1.07
F	0.23	0.08	0.07	0.54
G	0.32	0.12	0.10	0.76
Attribute	Group III			
	x_{av}	σ	Min	Max
A	0.62	0.05	0.54	0.74
B	0.50	0.04	0.42	0.58
C	0.19	0.02	0.16	0.24
D	1.43	0.41	0.94	2.50
E	0.53	0.25	0.35	0.93
F	0.29	0.08	0.17	0.49
G	0.47	0.15	0.27	0.78

Table 2. Mutual Classification Error between Classes U_i and U_j , expressed in Hartley for the Abalone

l_{ij}^k	U_{12}	U_{13}	U_{23}	Σ
A	0.18	λ	0.05	0.23
B	0.23	0.51	0	0.74
C	0.14	0.6	0	0.74
D	0.08	λ	0	0.08
E	0.07	0.06	0	0.13
F	0.07	0.07	0	0.14
G	0.05	0.58	0.04	0.68
$\sum_{k=1}^m l_{ij}^k$	0.82		0.09	

Figure 1 a) and b) show the membership functions for groups I and II for the attributes A, B, C, and D. Parameter D was presented only for making obvious that it can't be used for discrimination. Three values from the group I have been taken aleatory from the abalone table [9] like examples, and calculated with MATLAB the their compatibility index (CI) when using the membership functions of group I and group II. From Table 3 is clear that the CI between the selected examples from group I, when calculated with the group I G1-G1) is much bigger the CI when calculated with group I (G1-G2)I.

From Table 3 was clear that for comparing group 1 and Group II the quantity of information for discriminating is very small, so the analysis has been focused on parameters A, B, and C

Table 3. Compatibility Index

Example	G1-G1	G1-G2
1	0.34	0.054
2	0.98	0.22
3	0.98	0.362

It is important to say that here is not made a deep analysis of the abalone classification. The intention has been to present the possibilities of the method and the database is used only like one example.

6. Vertical handoff target selection in a heterogeneous wireless network

The information for this example has been taken basically from [10], Variations are introduced in the way of solving the problem. Wireless and mobile networking is becoming an increasingly important and

popular way to provide global information access to users on the move.

The handoff process has two major stages: handoff initiation, and handoff execution[2]. In the handoff initiation phase, a decision is made regarding the selection of the new Base Station (BS), or Access Point (AP), to which the Mobile Station (MS) will be transferred. In the execution phase, new radio links are formed between the BS/AP and MS, and resources are allocated [10]

There exist several methods for solving this problem [10]. Here is presented a simple method for the selection of the best network based in the RSS, the velocity, and the cost.

The following table 4 shows the input parameters for the selection. There are four (target networks) alternatives A1, A2, A3, and A4 from which it is necessary to select an optimum target network for the user

Table 4. Input Parameters

Network	A1	A2	A3	A4
RSS (dbm)	-87	-93	-83	-98
Velocity (Mbps)	90	100	82	50
Cost	52	42	38	30

Three Decision Makers (DM) with the different voting power are used. The DM1 have 41% of the voting power ($\zeta_1 = 0.41$) and DM2 have 34% of the voting power ($\zeta_2 = 0.34$) and the DM3 have 25% of the voting power ($\zeta_3 = 0.25$) respectively.

The desired condition for each decision maker is shown in Table 5, where the selected fuzzy variables are: very low (VL), low(L), medium (M), high(H), and very high (VH).

The membership functions are shown on Figure 2 a), and b). As there is not criterion for the membership functions selection, all of them have been taken similarly, only taking into consideration the different units. Table 6 shows the Compatibility Index calculated for each network.

From the table can be seen that the best option as considering the given conditions is the network A3.

Table 5. Desired condition for each attribute for each decision maker

Attributes	Decision Makers		
	DM1	DM2	DM3
RSS (C1)	H	M	M
Velocity (C2)	VH	H	H
Cost(C3)	M	L	VL

Table 6. Compatibility Indexes

DM	A1	A2	A3	A4
DM1	0.102	0	0.123	0
DM2	0	0.027	0	0.122
DM3	0.075	0	0.092	0
Σ	0.177	0.027	0.215	0.122

7. Conclusions

Several methods are developed for the classification with imprecise or missing information. The lack of data can be attributed to sensors failure, or simple due to incomplete system information. At last, diffuse values could be related to noise or imprecise measurements from sensors. It is not possible to affirm that one method is always better than other. In any case, the selection using fuzzy logic is simply, and extremely useful when dealing with large databases. The rough sets and information theory permit minimize the number of attributes. Both methods are based in the definition of the misclassification rates (α_{ij}^k). For any two classes, if this index is bigger than the defined permissible misclassification rate between classes α , the discrimination between these two classes is not possible. Fuzzy logic permits to obtain results from models created using MATLAB or other specialized software.

The first used example is the abalone database, in the analysis it is taken the dimensions only. In general as per the obtained results in the example, the conclusion could be that the selected attributes in this case are not always useful for determining the abalone age, using intervals of 5 years. Using fuzzy logic, it has been made the differentiation between groups I and II, but was not possible between II and III.

The network selected in the second example, is A3. The fundamental object of this example was to

show the simplicity of the method. For obtaining trusted results, it is necessary to make a complete analysis of the used membership functions, which can be done having a group of experts for solving this task.

References

1. Y. Leung, et Al., "A Rough Set Approach for the Discovery of Classification Rules in Interval-valued Information Systems", Int'l J. of Approximate Reasoning, No. 47, 2007, pp. 233-246.
2. Hem Jyotsana Parashar, et Al., "An Efficient Classification Approach for Data Mining" International J. of Machine Learning and Computing, Vol. 2, No. 4, August 2012.
3. S. M. Gorade et Al., "A Study of Some Data Mining Classification Techniques" International Research J. of Engineering and Technology (IRJET), Vol. 4 April 2017.
4. A. Caballero et Al. "Proposed method and Program for Classification of Information Systems", Proceedings from: 11th WSEAS International Conference on Circuits, Systems, Electronics, Control & Signal Processing (CSECS'12), Montreux, Switzerland, December 29-31, 2012
5. K. Yen, et Al., "Practical solution for the Classification in interval-Valued. Information Systems". WSEAS Transactions on Systems and Control, Issue 9, Vol. 5, September 2010. pp 735 – 744.
6. A. Caballero et Al., "Method for Optimizing the Number and Precision of Interval-Valued Parameters in a Multi-Object System", Proceedings from: "15th International Conference on Telecommunications and Informatics (TELE-INFO'16)", Ischia Italy, June 17-19, 2016.
7. J.T. Ross, *Fuzzy Logic with Engineering Applications*. John Wiley and Sons, Inc., 2004.
8. E.D. Cox, "Fuzzy Logic for Business and Industry" Rockland: Charles River Media Publisher, 1995.
9. Asuncion, A. & Newman, D.J. (2007). UCI Machine Learning Repository <http://www.ics.uci.edu/~mllearn/MLRepository.html>. Irvine, CA: University of California, School of Information and Computer Science.
10. M. Ramalingam, "Vertical Handoff Target Selection in a Heterogeneous Wireless Network Using Fuzzy Electre" Thesis for the Degree of Master of Science in Electrical Engineering, Florida International University, 2015.

a)

b)

Fig. 1 Membership functions for Abalone Groups I a) and II b)

a)

b)

Fig. 2

- a) Membership functions for RSS (dbm)
- b) Velocity. Membership Functions, (Mbps): 60, 65, 70, 75, 80,85, 90, 95, 100
Cost Membership Functions, (\$): 20, 25; 30, 35, 40, 45, 50, 55, 60