Classification with Diffuse or Incomplete Information

AMAURY CABALLERO, KANG YEN, YECHANG FANG
Department of Electrical & Computer Engineering
Florida International University
10555 W. Flagler Street, Miami, Florida 33174
USA
caballer@fiu.edu, yenk@fiu.edu, yfang003@fiu.edu

Abstract. The problem of classification has been studied by many authors, and different methods have been developed. In this paper a combination of rough sets and fuzzy logic for classification is adopted. Rough set theory helps in minimizing the number of attributes that influence the selection. Using this technique, a group of rules can be extracted. When information is diffuse and the number of obtained values for each attribute is large, so is the number of rules. Even worst is hidden information in the data that makes the process complicated. Due to this fact, an interval of values is defined for each attribute, moving from the minimum to the maximum obtained values in the database. This is what is defined as interval-valued information systems. For discriminating between solutions that may give more than one possible object due to their similarity, a fuzzy logic discrimination is proposed, which is simple, and gives accuracy not less than other methods.

Key Words: Fuzzy Logic, Rough Sets, Classification, Databases.

1 Introduction

When analyzing an information system or a database, frequently we have found problems like attributes redundancy, missing or diffuse values, which are due in general to noise and missing partial data. Rough set theory is a very useful approach for minimizing the number of attributes necessary to represent the desired category structure by eliminating redundancy. The lack of data or complete knowledge of the system makes 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 obtained from different sensors, which are corrupted by noise and outliers. The present work is devoted to the analysis of these situations and the way they can be solved using rough and fuzzy sets. An example has been included to demonstrate the concepts of using rough and fuzzy sets, in classification applications.

2 Methodologies

2.1 Rough Sets

Rough Set Theory (RST), an extension of conventional set theory, supports approximations in decision making, and was introduced by Pawlak in 1982 [1]. This theory can be used as a tool to recover data dependencies and to reduce the number of attributes contained in a given data set by using the data alone, without additional information [2].

Pawlak has defined an information system by a pair $K = (U, A)$, where U is the universe of all objects and A is a finite set of attributes [3]. Any attribute $a \in A$ is a total function of a: $U \rightarrow V_a$, where V_a is the domain of a. Let $a(x)$ denote the value of attribute a for object x, where $[x]_A$ is the corresponding equivalent class containing object x. We say that (x, y) is A-indiscernible for the equivalence relation $I(A)$, where $U/I(A)$ denotes the partition determined by the relation $I(A)$, if (x, y) belongs to $I(A)$ [4]. Then two operations can be defined on any subset X of U. The *lower approximation* (positive region), is defined as follows:

$$\underline{A}(X) = \{x \in U : [x]_A \subseteq X\}$$

It is the union of all equivalent classes in $[x]_A$ which are subsets of X. The lower approximation is the complete set of objects that can be unequivocally classified as belonging to set X. This is called the positive region if A and X are equivalence relations over U:

$$POS_A(U) = \cup_{X \in U/A} A(X)$$

The *upper approximation* has the definition

$$\overline{A}(X) = \{x \in U : [x]_A \cap X \neq 0\}$$

It is the union of all equivalence classes in $[x]_A$ which are non-empty intersection with the target set. The upper approximation is the complete set of objects that are possible members of the target set X . The concept of *reduct* is important, which is the minimum number of attributes that can characterize the knowledge in the information system as a whole, or a subset of it. In any case, the *reducts* are not unique. The set of attributes which is common to all *reducts* is called *core*. Rough set attribute reduction provides a tool by which knowledge is extracted from a dataset, retaining the information content, without affecting the knowledge involved.

In the literature there are several methods for attribute reduction. R. Jensen [2] analyzes the problem of finding a *reduct* of an information system. It becomes clear that the perfect solution to locating such a subset is to generate all possible subsets and retrieve those with a maximum rough set dependency degree, but this solution is not practical for medium and large databases. A method for practical reducing the number of attempts is the *QUICK REDUCT* algorithm [2]. This algorithm starts off with an empty set and adds in turn, one at a time, those attributes that result in the greatest increase in the rough dependency metric until it produces its maximum possible value for the dataset. For very large datasets, one criterion for stopping the search could be to terminate the process when there is no further increase in the dependency. Another method uses the *discernibility matrix* approach. A discernibility matrix of a decision table $D = (U, C \bigcup \Omega)$ is a symmetric $|U| \times |U|$ matrix, where Ω is the set of decision features, C is the set of all conditional features, and the entries correspond to

$$f_D(a_1^*, ..., a_m^*) = \wedge\{\vee d_{ij}^* \mid d_{ij} \neq \varnothing\}$$
$$1 \leq j \leq i \leq |U|$$

Where

$$d_{ij}^* = \{a^* \mid a \in d_{ij}\}$$

Each d_{ij} contains those attributes that differ between objects i and j. The discernibility function f_D for minimizing the attributes a_j using the discernibility matrix is a Boolean function given for each term by

$$f_i(a_1,...,a_m) = \wedge\{\vee d_{ij}\}$$

Finding the set of all prime implicants of the discernibility function, all the minimal reducts of the system may be determined.

2.2. Fuzzy Sets

Fuzzy logic is a multi-valued Boolean logic that helps describing concepts that are commonly encountered in the real world, using linguistic variables. The range of possible values of a linguistic variable represents the universe of discourse of that variable.

One of the basic concepts in fuzzy logic is that of the membership functions. In general any function $A: X \to [0, 1]$ can be used as a membership function describing a fuzzy set. Differently from the Boolean logic, which only consider one of two possible states for a proposition, fuzzy logic states that it can be more states, which are defined by the membership functions.

When designing a fuzzy system, the expert encounters besides the selection of the input and output functions and their universe of discourse, the problem of optimizing the number and characteristics of the membership functions as well as the rules that control the process. Another issue is the way the classification is going to be made. Many simple decision processes are based on a single attribute, such as minimizing cost, maximizing profit, etc. An example of this can be found in [5]. Often, however, decisions must be made in an environment where more than one attribute affect the decision, and the relative value of each of these attributes can be different [6]. This is the case that often appears in classification problems.

When solving a multi-attribute classification problem, it is necessary to acquire information regarding the attributes belonging to the different classes (objects) and to rank or weight the relative importance of each of the attributes. The typical multi-attribute classification problem involves the selection of one object u_i from a group of objects, given a collection, or a set of attributes that are important to the classification.

Having the universe of m attributes $A = a_1, a_2, \ldots, a_m$ and a set of n objects $U = u_1, u_2, \ldots, u_n$, the degree of membership of alternative a in u_i, denoted by $\mu u_i(a)$, is the degree to which alternative a satisfies the criteria specified for this objective. All this leads to the concept of the *Compatibility Index* (*CI*), presented by Cox [7], and applied in several works [8] which is a method for calculating the similarity of a situation with previously imposed conditions using fuzzy sets. In this case, a weakness in the contribution of one attribute can be compensated by an increase in strength of the remaining attributes. This index can be calculated as the average of the degrees of membership (μ) for each object considering all the participating attributes. If all the attributes are given the same weight, the compatibility index can be calculated using the expression:

$$CI_i = CI_i = \sum_{k=1}^{m} \mu_k / m$$

Where $i = 1,\ldots, n$ is the object number and $k = 1,\ldots, m$ is the attribute number.

3. Classification of Diffuse or Incomplete Dataset

The theory of rough sets is used for the analysis of categorical data. The values that can be obtained for the different attributes characterizing the objects of interest may be very large, which can lead to the creation of too many rules.

As indicated by Y. Leung [9]: "These rules may be accurate with reference to the training data set; their generalization ability will most likely be rather low since perfect match of attribute values of the condition parts in real numbers is generally difficult if not impossible. To make the identified classification rules more comprising and practical, a preprocessing step which can transform the real numbered attribute values into a sufficiently small number of meaningful intervals is thus necessary." This method of approach is known as granular computing. Y.Y. Yao [10, 11] stated that when a problem involves incomplete, uncertain, or vague information, it may be difficult to differentiate distinct elements and one is forced to consider granules. This approach leads to the simplification of practical problems. R Jensen [2] presents a fuzzy-rough feature selection handling noisy data, to perform discretization before dealing with the data set. Several algorithms and examples are presented [2].

A generalization of rough set models based on fuzzy lower approximation is presented by Wang [4]. The concept of tolerant lower approximation is introduced for dealing with noisy data. When dealing with incomplete information, the typical approach is to use fuzzy logic, deriving inference rules from training examples using the available information and assuming that the characteristics of the missing part are similar to that of the previously obtained. Literature defines many methods to derive membership functions and fuzzy rules from training examples [12, 13]. Among them can be cited the Batch Least Square Algorithm, the Recursive Least Square Algorithm, the Gradient Method, the Learning from Example method, etc. E. A. Rady [14] introduced the modified similarity relation, for dealing with incomplete information systems, which is dependent on the number of missing values with respect to the number of the whole defined attributes for each object. Also neural networks and genetic algorithms have been used for classification with incomplete data. E. Granger, [15] present a work, analyzing the situations of limited number of training cases, missing components of the input patterns, missing class labels during training, and missing classes, using the fuzzy ARTMAP neural network.

4 Example for Diffuse Dataset

In this paper, the solution of the iris classification problem is used like a practical example. This problem has been solved by several authors. One example of solving classification problems using fuzzy sets and rough theory is presented by T. Ying-Chieh, [16]. He uses a rough classification approach and a minimization entropy algorithm for solving the iris classification problem. Another solution for solving the same problem is developed by Shyi-Ming Chen and Yao-De Fang [17]. These examples use Table 1, originally developed by R. Fisher [18]. The classes are defined as SL-sepal length, SW-sepal width, PL-petal length, and PW-petal width. These classes will be used in our example, where the mean and standard deviation, values have been calculated, as well as the minimum and maximum for each class have been included. The results are presented in Table 2.

Table 1. Iris Classification [15]

	Setosa				Versicolor				Virginica			
No.	SL	SW	PL	PW	SL	SW	PL	PW	SL	SW	PL	PW
1	5.1	3.5	1.4	0.2	7	3.2	4.7	1.4	6.3	3.3	6	2.5
2	4.9	3	1.4	0.2	6.4	3.2	4.5	1.5	5.8	2.7	5.1	1.9
3	4.7	3.2	1.3	0.2	6.9	3.1	4.9	1.5	7.1	3	5.9	2.1
4	4.6	3.1	1.5	0.2	5.5	2.3	4	1.3	6.3	2.9	5.6	1.8
5	5	3.6	1.4	0.2	6.5	2.8	4.6	1.5	6.5	3	5.8	2.2
6	5.4	3.9	1.7	0.4	5.7	2.8	4.5	1.3	7.6	3	6.6	2.1
7	4.6	3.4	1.4	0.3	6.3	3.3	4.7	1.6	4.9	2.5	4.5	1.7
8	5	3.4	1.5	0.2	4.9	2.4	3.3	1	7.3	2.9	6.3	1.8
9	4.4	2.9	1.4	0.2	6.6	2.9	4.6	1.3	6.7	2.5	5.8	1.8
10	4.9	3.1	1.5	0.1	5.2	2.7	3.9	1.4	7.2	3.6	6.1	2.5
11	5.4	3.7	1.5	0.2	5	2	3.5	1	6.5	3.2	5.1	2
12	4.8	3.4	1.6	0.2	5.9	3	4.2	1.5	6.4	2.7	5.3	1.9
13	4.8	3	1.4	0.1	6	2.2	4	1	6.8	3	5.5	2.1
14	4.3	3	1.1	0.1	6.1	2.9	4.7	1.4	5.7	2.5	5	2
15	5.8	4	1.2	0.2	5.6	2.9	3.6	1.3	5.8	2.8	5.1	2.4
16	5.7	4.4	1.5	0.4	6.7	3.1	4.4	1.4	6.4	3.2	5.3	2.3
17	5.4	3.9	1.3	0.4	5.6	3	4.5	1.5	6.5	3	5.5	1.8
18	5.1	3.5	1.4	0.3	5.8	2.7	4.1	1	7.7	3.8	6.7	2.2
19	5.7	3.8	1.7	0.3	6.2	2.2	4.5	1.5	7.7	2.6	6.9	2.3
20	5.1	3.8	1.5	0.3	5.6	2.5	3.9	1.1	6	2.2	5	1.5
21	5.4	3.4	1.7	0.2	5.9	3.2	4.8	1.8	6.9	3.2	5.7	2.3
22	5.1	3.7	1.5	0.4	6.1	2.8	4	1.3	5.6	2.8	4.9	2
23	4.6	3.6	1	0.2	6.3	2.5	4.9	1.5	7.7	2.8	6.7	2
24	5.1	3.3	1.7	0.5	6.1	2.8	4.7	1.2	6.3	2.7	4.9	1.8
25	4.8	3.4	1.9	0.2	6.4	2.9	4.3	1.3	6.7	3.3	5.7	2.1
26	5	3	1.6	0.2	6.6	3	4.4	1.4	7.2	3.2	6	1.8
27	5	3.4	1.6	0.4	6.8	2.8	4.8	1.4	6.2	2.8	4.8	1.8
28	5.2	3.5	1.5	0.2	6.7	3	5	1.7	6.1	3	4.9	1.8
29	5.2	3.4	1.4	0.2	6	2.9	4.5	1.5	6.4	2.8	5.6	2.1
30	4.7	3.2	1.6	0.2	5.7	2.6	3.5	1	7.2	3	5.8	1.6
31	4.8	3.1	1.6	0.2	5.5	2.4	3.8	1.1	7.4	2.8	6.1	1.9
32	5.4	3.4	1.5	0.4	5.5	2.4	3.7	1	7.9	3.8	6.4	2
33	5.2	4.1	1.5	0.1	5.8	2.7	3.9	1.2	6.4	2.8	5.6	2.2
34	5.5	4.2	1.4	0.2	6	2.7	5.1	1.6	6.3	2.8	5.1	1.5
35	4.9	3.1	1.5	0.2	5.4	3	4.5	1.5	6.1	2.6	5.6	1.4
36	5	3.2	1.2	0.2	6	3.4	4.5	1.6	7.7	3	6.1	2.3
37	5.5	3.5	1.3	0.2	6.7	3.1	4.7	1.5	6.3	3.4	5.6	2.4
38	4.9	3.6	1.4	0.1	6.3	2.3	4.4	1.3	6.4	3.1	5.5	1.8
39	4.4	3	1.3	0.2	5.6	3	4.1	1.3	6	3	4.8	1.8
40	5.1	3.4	1.5	0.2	5.5	2.5	4	1.3	6.9	3.1	5.4	2.1
41	5	3.5	1.3	0.3	5.5	2.6	4.4	1.2	6.7	3.1	5.6	2.4
42	4.5	2.3	1.3	0.3	6.1	3	4.6	1.4	6.9	3.1	5.1	2.3
43	4.4	3.2	1.3	0.2	5.8	2.6	4	1.2	5.8	2.7	5.1	1.9
44	5	3.5	1.6	0.6	5	2.3	3.3	1	6.8	3.2	5.9	2.3
45	5.1	3.8	1.9	0.4	5.6	2.7	4.2	1.3	6.7	3.3	5.7	2.5
46	4.8	3	1.4	0.3	5.7	3	4.2	1.2	6.7	3	5.2	2.3
47	5.1	3.8	1.6	0.2	5.7	2.9	4.2	1.3	6.3	2.5	5	1.9
48	4.6	3.2	1.4	0.2	6.2	2.9	4.3	1.3	6.5	3	5.2	2
49	5.3	3.7	1.5	0.2	5.1	2.5	3	1.1	6.2	3.4	5.4	2.3
50	5	3.3	1.4	0.2	5.7	2.8	4.1	1.3	5.9	3	5.1	1.8

Table 2. Attributes for Each Class

Attributes	Setosa				Versicolor				Virginica			
	x_{av}	σ	Min	Max	x_{av}	σ	Min	Max	x_{av}	σ	Min	Max
SL	5.0	0.35	4.3	5.8	5.94	0.52	4.9	7.0	6.59	0.64	4.9	7.9
SW	3.42	0.38	2.3	4.4	2.77	0.31	2.0	3.4	2.91	0.32	2.2	3.8
PL	1.45	0.11	1.0	1.9	4.26	0.47	3.0	5.1	5.55	0.55	4.5	6.9
PW	0.24	0.11	0.1	0.6	1.33	0.20	1.0	1.8	2.03	0.27	1.4	2.5

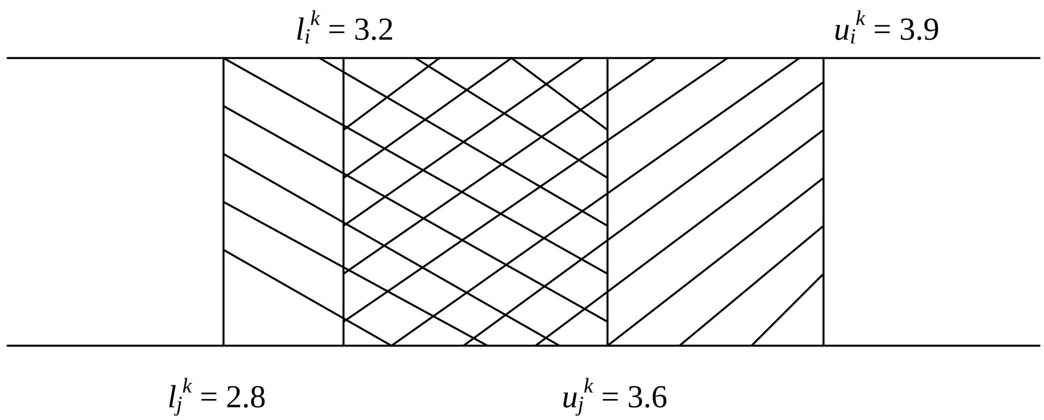

Fig.1. Region of Coincidence of Classes i and j for the Attribute k

The solution is divided in two parts. In the first part, the useful attributes for the classification is found based on the rough theory. For this part, the approach introduced by Leung Yee [9] is applied to the iris classification problem with small changes. In the second part, fuzzy logic is employed for discriminating between similar situations.
For clarifying the procedure, the following definitions are presented.

4.1 Misclassification Rates

Let α_{ij}^k denote the misclassification error between the classes i and j for attribute k. Probability that objects in class u_i are misclassified in class u_j according to attribute k is:

$$\alpha_{ij}^k = 0 \quad \textit{if } [l_i^k, u_i^k] \cap [l_j^k, u_j^k] = 0;$$

$$\alpha_{ij}^k = \min\{(u_i^k - l_j^k, u_j^k - l_i^k)/(u_i^k - l_i^k)\},1\}, \quad \textit{if } [l_i^k, u_i^k] \cap [l_j^k, u_j^k] \neq 0.$$

Where l_i^k and u_i^k are the minimum and maximum values for the object i and attribute k; l_j^k and u_j^k are the minimum and maximum values for the object j and attribute k.
For clarifying the concept, we present a numerical example on Fig.1. The region of intersection of the values of the two objects represents the zone, where the classification becomes problematic.
For the presented case, the misclassification error between the classes i and j for attribute k is given by

$\alpha_{ij}{}^{k} = min\{(3.9–2.8, 3.6 –3.2)/(3.9 – 3.2)\}$

$\alpha_{ij}{}^{k} = 0.57$

Note that in general

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

The maximum mutual classification error between classes u_i and u_j for attribute k is

$$\beta_{ij}{}^{k} = max \{ \alpha_{ij}{}^{k}, \alpha_{ji}{}^{k}\}$$

where $\beta_{ij}{}^{k} = \beta_{ji}{}^{k}$. The permissible misclassification rate between classes u_i and u_j in the system k is

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

Defining a parameter α as the specified admissible classification error, $0 < \alpha < 1$. 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 [6].

4.2 α-Tolerance Relation Matrix

For a given permissible misclassification rate $\alpha \in [0, 1]$ and an attribute subset $B \subseteq A$, a binary relation on U is defined by

$$R_B^{\alpha} = \{(u_i, u_j) \in U \times U \mid \beta_{ij}^{k} > \alpha, \forall a_k \in B\}$$

The errors that objects in class u_i being misclassified into class u_j in the system are defined as $\alpha_{ij} = min\ \{\alpha_{ij}{}^{k}: k \leq m\}$ and are given in Table 3.

Table 3.Error of Misclassification of Object u_i into u_j

α_{ij}	u_1	u_2	u_3
u_1	1	0	0
u_2	0	1	0.25
u_3	0	0.29	1

The maximal mutual classification error between classes is defined by

$$\beta_{ij}{}^{k} = max\ \{\alpha_{ij}{}^{k}, \alpha_{ji}{}^{k}\}$$

In the present example is given by:

$\beta_{12}{}^{1} = 0.6$ $\beta_{13}{}^{1} = 0.6$ $\beta_{23}{}^{1} = 1$
$\beta_{12}{}^{2} = 0.78$ $\beta_{13}{}^{2} = 0.94$ $\beta_{23}{}^{2} = 0.86$
$\beta_{12}{}^{3} = 0$ $\beta_{13}{}^{3} = 0$ $\beta_{23}{}^{3} = 0.29$
$\beta_{12}{}^{4} = 0$ $\beta_{13}{}^{4} = 0$ $\beta_{23}{}^{4} = 0.5$

Selecting α = 0.2, the permissible misclassification rate for this example is shown on Table 4.

Table 4. Permissible Misclassification Rate Between Classes u_i and u_j

β_{ij}	u_1	u_2	u_3
u_1	1	0	0
u_2	0	1	0.29
u_3	0	0.29	1

The matrix for the α-Tolerance relations, where all $\beta_{ij} > \alpha$ is represented by 1.

$$R_A{}^{0.2} = \begin{bmatrix} 100 \\ 011 \\ 011 \end{bmatrix}$$

From the matrix, it is clear that object u_1 - Setosa can be uniquely defined from the given attributes, but objects u_2 - Versicolor and u_3 - Virginica may not be separated. This situation can be expressed by

$$S_A{}^{0.2}(u_1) = \{u_1\}$$
$$S_A{}^{0.2}(u_2) = S_A{}^{0.2}(u_3) = \{u_2, u_3\},$$

where $S_A{}^{0.2}(u)$ denotes that these are the sets of objects which are possible indiscernible by A within u within the misclassification rate $\alpha = 0.2$. From the previous results, the 0.2-discernibility set is given on Table 5.

Table 5. Discernibility Set

	u_1	u_2	u_3
u_1			
u_2	a_3, a_4		
u_3	a_3, a_4		

The obtained functions are

$$f_1^{0.2} = a_3 \vee a_4$$
$$f_2^{0.2} = a_3 \vee a_4$$

Using rough sets, it has been demonstrated that the important attributes for the classification are a_3: PL-petal length, and a_4: PW-petal width.

From the previous results, the following rules can be extracted:

$R(u_1)$ – IF $a_3 \in [1, 1.9]$ or $a_4 \in [0.1, 0.6]$ THEN it is u_1 – Setosa.

$R(u_2)$ – IF $a_3 \in [3.0, 5.1]$ or $a_4 \in [1.0, 1.8]$ THEN it can be u_2 – Versicolor or u_3 – Virginica.

$R(u_3)$ – If $a_3 \in [4.5, 6.9]$ or $a_4 \in [1.4, 2.5]$ THEN it can be u_2 – Versicolor or u_3 – Virginica.

Rule 1 is clear for Setosa classification. Note that from R(u_2) and R(u_3), it is possible to develop other rules that substitute them:

$R(u_4)$ – IF $a_3 \in [3.0, 4.5]$ or $a_4 \in [1.0, 1.4]$ THEN it is u_2 – Versicolor.

$R(u_5)$ – If $a_3 \in [5.1, 6.9]$ or $a_4 \in [1.8, 2.5]$ THEN it is u_3 – Virginica.

$R(u_6)$ – If $a_3 \in [4.5, 5.1]$ or $a_4 \in [1.4, 1.8]$ THEN it can be u_2 – Versicolor or u_3 – Virginica.

In order to select between u_2 and u_3, into the coincident interval, one possibility is to use fuzzy logic. The authors propose the following procedure:

1) For each of the objects u_i in the fired rule, find the degree of membership with the imposed conditions, for each of the participating attributes a_k.
2) Find the compatibility index (*CI*) for each object.
3) Compare the different compatibility indexes and select the object with the greater one.

For the example, it was proposed for each interval and each type, a bell shape membership function with the maximum value coincident with the mean value for the interval, and defining the domain from the minimum to maximum values in the interval. The compatibility indexes are calculated using for this case, the different measurements obtained from Table 1. Rule $R(u_6)$ has been used taking into consideration only the petal length (a_3) and petal width (a_4), as stated by the rule. It has been tested the values for measurements 1, 5, and 10. This test is shown in Table 6. The values for versicolor were compared with the membership function "versicolor", and the same was done for virginica. This is shown in columns 2 and 3. Latter, the values for versicolor were compared with the membership function "virginica" and the values for virginica with the membership function "versicolor". This is shown in columns 4 and 5. As can be seen from Table 6, there is a big difference in the compatibility index between the proper and the wrong comparison. A test was made to all the objects in table 1. In this case, there are 50 virginica and 50 versicolor examples. The algorithm fails in one versicolor and one virginica. This gives an average classification rate of 98% for the analyzed table.

Table 6. Compatibility Indexes for Different Measurements from Table 1

No.	Comp. Index Vers./Vers.	Comp. Index Virg./Virg.	Comp. Index Vers./Virg.	Comp. Index Virg./Vers.
1	0.59	0.46	≤ 0.1	≤ 0.1
5	0.60	0.79	≤ 0.1	≤ 0.1
10	0.99	0.42	0.3	≤ 0.1

5 Conclusions

The combination of rough sets and fuzzy logic is a powerful tool for classifying objects into a database. The attribute minimization using rough set theory is extremely useful when dealing with large databases. If the number of possible values for the attributes is large, the selection of interval values is mandatory. Fuzzy logic can be used together with the rough theory for obtaining a unique response, in case where this is not possible using the rough theory alone. Several methods are presented for the classification with imprecise or missing information. It is not possible to affirm that one method is always better than other. In any case, there is always some uncertainty regarding the final result. This uncertainty is related to the accepted admissible classification error (α) in the case of imprecise information and to the characteristics and quantity of the missed information, in the case of incomplete information. The solution of the iris classification using a method that differs from those originally presented in the literature shows a different approach for obtaining equivalent results. The solution of the iris classification in the present work, results simpler than other methods due to the fact that simpler and/or fewer rules have been used. The usefulness of the method is more evident when dealing with databases containing a large number of objects and attributes, but the example presented in this paper serves for the purpose of giving an idea of the presented method.

References

[1] Z. Pawlak, Rough Set, *Int'l J. of Computer & Information Sciences,* No. 11, 1982, pp.341-356.

[2] R. Jensen, *Combining Rough and Fuzzy Sets for Feature Selection*, Ph.D. Dissertation, U. of Edinburgh, 2005,

[3] Z. Pawlak, *Rough Sets*: *Theoretical Aspects of Reasoning About Data*, Dordrecht: Kluwer, 1991.

[4] Feng-Hsu Wang, On Acquiring Classification Knowledge from Noisy Data based on Rough Set, *Expert Systems with Applications*, No.29, 2005, pp.49-64

[5] Kaleem F., et Al, *Intelligent Spatial Filters Based on Fuzzy Logic.* WSEAS Transactions on Circuits. Issue 1, Volume 2, January 2003. pp. 252 – 258.

[6] Caballero A**.**, et Al. *Model for Clustering Objects under Specified Conditions* . WSEAS Transactions on Systems and Control, Issue 1, Vol. 2, January 2007. pp 86 – 91.

[7] E.D. Cox, *Fuzzy Logic for Business and Industry*, Rockland: Charles River Media, 1995

[8] Caballero A., K. Yen. *Fuzzy Model for Risk Evaluation in Construction Projects.* WSEAS Transactions on Business and Economics, Issue 4, Vol. 1, October 2004. pp 275 – 280.

[9] 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.

[10] Y.Y. Yao, Information Granulation and Rough Set Approximation, *Int'l J. of Intelligent Systems*, Vol.16, No.1, 2001, pp.87-104.
[11] Y.Y. Yao, Information Granulation and Approximation in a Decision Theoretic Model of Rough Sets, *Rough-Neuro Computing with Words*, Edited by L. Polkowski et Al. Springuer, Berlin, 2003, pp.491-518.
[12] J.T. Ross, *Fuzzy Logic with Engineering Applications*, John Wiley and Sons, Inc., 2004.
[13] K. Pasino, S. Yurkovich, *Fuzzy Control*, Addison Wesley Longman, Menlo Park, California.
[14] E. Rady, et Al., A Modified Rough Set approach to Incomplete Information /systems, *J. of Applied Mathematics and Decision Sciences*, 2007.
[15] E. Granger, et Al., *Classification of Incomplete Data Using the Fuzzy ARTMAP Neural Network*, Technical Report. Boston U., Center for Adaptive Systems and Dept. of Adaptive Neural Systems, January 2000.
[16] T. Ying-Chieh, T. et Al., Entropy-Based Fuzzy Rough Classification Approach for Extracting Classification Rules, *Expert systems with Applications*, No. 31, 2006, pp.436-443.
[17] Shy-Ming Chen and Yao-De Fang, A New approach for Handling the Iris Data Classification Problem, *Int. Journal of Applied Science and Engineering*, No.3, 2005, pp. 37-49.
[18] R. Fisher, The Use of Multiple Measurement in Taxonomic Problem, *Ann. Eugenic*, No. 7, 1936, pp. 179-188.