Decision Rules making based on Rough Set Approach

Mr. Antoine Nzaramba, Prof. Wang Jia Yang, Mr. Gilbert Kipkirui Langat

ABSTRACT-Rough set theory is a field which in many years been a focus of study, it is known for many uses including generation of rules which will help in final decision.

In our approach we have come up with an elimination based technique which avoids the redundancy and reflects the global perspective of the whole set. By using the technique of Reduct and core we eliminate non necessary attributes and stay with the most important attributes which will help us for decision making.

Key terms: Rough set, Reduct, Core, Rule generation, Decision making, information system, Knowledge Discovery.

1. INTRODUCTION

Often, information on the surrounding world is imprecise, incomplete or uncertain. Still our way of thinking and concluding depends on information at our disposal. This means that to draw conclusions, we should be able to process uncertain and / or incomplete information.

Tools, which turned out to be particularly adequate for the analysis of various types of data, especially, when dealing with inexact, uncertain or vague knowledge, are the fuzzy set and the rough set theories. Rough sets and fuzzy sets capture two distinct aspects of imperfection in knowledge: indiscernibility and vagueness. The fuzzy set theory, introduced by Zadeh in 1965 [1], has already demonstrated its usefulness in chemistry and in other disciplines [2-9]. The rough set theory, introduced by Pawlak in 1985 [14, 15], although popular in many other disciplines [10], is nearly unknown in chemistry [11, 12].

To deal with the uncertainty problems, the concept of fuzzy set was introduced. Fuzzy set is defined by the membership function which can attain values from the closed interval [1], allowing partial membership of the elements in the set.

In the rough set theory, membership is not the primary concept. Rough sets represent a different mathematical approach to vagueness and uncertainty. Definition of a set in the rough set theory is related to our information knowledge and perception about elements of the universe.

The rough set methodology is based on the premise that lowering the degree of precision in the data makes the data pattern more visible [13], whereas the central premise of the rough set philosophy is that the knowledge consists in the ability of classification. In other words, the rough set approach can be considered as a formal framework for discovering facts from imperfect data. The results of the rough set approach are presented in the form of classification or decision rules derived from a set of examples.

The aim of this paper is to presents a new approach for determining the most important attribute on the basis of strength of an association; introduce the basic concepts of the Rough Set Theory and also to show some of its possible applications.

2. ROUGH SET THEORY

2.1. Basic concepts of the rough sets theory

2.1.1. Information system

- *Mr. Antoine Nzaramba: currently pursuing masters degree program in Computer Science in Central South University, China, PH-(+86)18569416743. E-mail: napoumen@yahoo.fr*
- *Prof. Wang Jia Yang: my supervisor, teacher in Central South University, China*
- *Mr. Gilbert Kipkirui Langat: currently pursuing PhD program in Computer Science in Central South University, China*

Information System is a group of components that interact to produce information. In our case of rough sets theory, Information System is equal to (U, A) where A is a non-empty, finite set of attributes and U is a non-empty, finite set of objects (the universe).

Suppose, we are given an information system

$$S = (U, A), X \subseteq U \text{ and } P \subseteq A$$

where U and A, are finite, nonempty sets and called as the universe, and the set of attributes, respectively. Set A will contain two disjoint sets of attributes, called condition and decision attributes and the system is donated by

$S = (U, C, D)$

Where C is called condition attribute and D is called decision attribute. With every attribute a ϵ A we associate a set Va, of its values, called the domain of a.

2.1.2. Indiscernibility relation

Indiscernibility is a central concept of Rough Set Theory which relates to the Process of grouping objects based on having the same series of attributes values. In other words the data are supposed to be similar with respect to this relation.

2.1.3. Lower and upper approximations

The rough sets approach to data analysis hinges on two basic concepts, namely the *lower* [$\underline{\mathbf{P(X)}}$] and the *Upper approximations* [**P(X)**] of a set.

Lower Approximation refers to the elements that doubtlessly belong to the set and Upper Approximation refers to the elements that possibly belong to the set

$$\underline{P(X)} = \cup x \in U \{P(X) : P(X) \in X\}$$

And

$$P(X) = \cup x \in U \{P(x): P(x) \cap X \neq \varphi\}$$

2.1.4. Boundary region

The boundary region is given by the set difference **P(X)-** $\underline{\mathbf{P(X)}}$ consists of those objects that can neither be ruled in nor ruled out as members of the target set X.

If the boundary region is empty i.e **P(X) =** $\underline{\mathbf{P(X)}}$ then the set is crisp or definable otherwise the set is rough or undefinable. Rough set theory can determine whether there is any redundant information in the data and if it is there, then can we find essential data required for our applications.

There are four types of undefinable sets in U:

1. If $P(X) \neq \phi$ and $P(X) \neq U$, X is called roughly definable in U;
2. If $P(X) \neq \phi$ and $P(X) = U$, X is called externally undefinable in U;
3. If $P(X) = \phi$ and $P(X) \neq U$, X is called internally undefinable in U;
4. If $P(X) = \phi$ and $P(X) = U$, X is called totally undefinable in U,

where ϕ denotes an empty set.

2.1.5. Accuracy of approximation

The accuracy of the approximation to the set X from the elementary subsets is measured as the ratio of the lower and the upper approximation size. The ratio is equal to 1, if no boundary region exists, which indicates a perfect classification. In this case, deterministic rules for the data classification can be generated. Thus, a set X with accuracy equal to 1 is crisp otherwise X is rough.

Accuracy of approximation= card(P(X)) / card(P(X))

With **card**: cardinality.

2.1.6. Core and reduct of attributes

Reduct and core are the two most important concept of rough set theory. Reduct is a reduced subset of original set which retains the accuracy of original set. Reduct is often used in the attribute selection process to reduce unnecessary attributes towards decision making application.

In a decision table we may find multiple reduct and some rule would appear more frequently in some reduct than others. There are so many methods of finding reduct of a decision table. The reducts can be obtained by using the reduct generation algorithms. Using the discernibility matrix, the reduct of a decision table can be found [1]. The core can be found as the set of all singleton entries in the discernibility matrix. The reduct is the minimal element in the discernibility matrix, which intersects all the element of the discernibility matrix.

3. REDUCTION OF DATA AND RULES FINDING

Rough sets have many applications in the field of Knowledge Discovery in Databases (KDD), such as feature selection, data reduction, discretization, etc. When a dataset contains irrelevant (dispensable) features the same may be eliminated and thereby reducing the dimension of the problem. Thereafter, Rough sets can be used to find subsets of relevant (indispensable) features.

The volume of data is increasing day by day. In many real applications, it is very difficult to find which attributes are important for a particular task and which attributes are not so important. Hence identifying the relevant features is important for the reduction of the volume of data. The aim of data reduction is to find a minimal subset of relevant attributes that have all the essential information of the data set, thus the minimal subset of the attributes can be used instead of the entire attributes set for rule discovery.

3.1 Decision Table

Rough set theory can be considered as an extension of classical set theory. The basic concept of the RST is the notion of approximation space, which is with every object of universe we associate some information i.e. Data and Knowledge. Every example of the Rough set is organized in the form of information table, whose columns are labeled as condition and decision attributes and rows of the table contain the example. Entries in the table represent the attribute values. So a knowledge representation system containing the set of attributes A now called condition attributes and the set of decision attributes D is called a decision table.

Table 1 is a decision table whose decision attribute is LOOK and condition attributes are {HAT, SHIRT, TROUSER, SHOES}.

Table 1: Decision Table

	HAT	SHIRT	TROUSER	SHOES	LOOK
1	Black	Red	Black	Sneakers	Smart
2	Black	Red	Bleu	Oxford	Smart
3	Red	Red	Black	Oxford	Smart
4	Red	White	Black	Sneakers	Good
5	Black	White	Black	Monk	Good
6	Red	White	Red	Sneakers	Acceptable
7	Bleu	White	Red	Sneakers	Acceptable
8	Bleu	Grey	Red	Sneakers	Acceptable
9	Bleu	Grey	Bleu	Monk	Good
10	Bleu	Grey	Bleu	Monk	Smart

From Table 1 it is easy to see that for example 9 and 10 all the values of the condition attributes are same except for the values of decision attributes. We can say that Table1 is inconsistent because example 9 and 10 are conflicting (or are inconsistent) for both examples the value of all condition attribute is the same, yet the decision value is different.

3.2 Calculation of Lower and Upper Approximations

Rough set theory offers a tool to deal with inconsistencies. For each concept X the greatest definable set contained in X and the least definable set containing X are computed. The former set is called a lower approximation of X the latter is called an upper approximation of X. In the case of Table 1, the elementary sets are {1}, {2}, {3}, {4}, {5}, {6}, {7}, {8}, {9, 10}

Now, let us consider the concept for the table1. We can define decision attributes and elementary set associated with the decision as subset of the set of all examples with

the same value of decision. Such subset are called concept. There are three concepts in Table1.

A1 = {1, 2, 3, 10} for decision **Smart**

A2 = {4, 5, 9} for decision **Good**

A3 = {6, 7, 8} for decision **Acceptable**

We can easily find lower and upper approximation of these three concepts.

$\underline{P(X)}$ = (1, 2, 3, 4, 5, 6, 7, 8) (Lower approximation)

P(X) = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10) (Upper approximation)

The Boundary region = P(X) - $\underline{P(X)}$ = {9, 10}

From the table 1, we can eliminate the conflicting examples which are 9 and 10.

Then we will get the table 2

Table 2: Decision Table after eliminating conflict examples

	HAT	SHIRT	TROUSER	SHOES	LOOK
1	Black	Red	Black	Sneakers	Smart
2	Black	Red	Bleu	Oxford	Smart
3	Red	Red	Black	Oxford	Smart
4	Red	White	Black	Sneakers	Good
5	Black	White	Black	Monk	Good
6	Red	White	Red	Sneakers	Acceptable
7	Bleu	White	Red	Sneakers	Acceptable
8	Bleu	Grey	Red	Sneakers	Acceptable

3.3 Rule Generation

Based on reduct and core of the table 2 we are going to generate the rules. Reduct is the reduced set of relation that conserves the same inductive classification of Relation. The set P of attributes is the reduct of another set Q of attributes if P is minimal and the indiscernibility relations, defined by P and Q are same.

Core = ∩ reduct

Reduct of table2 are { HAT, SHOES, TROUSER}, { HAT, SHIRT, TROUSER }, { SHIRT, SHOES, TROUSER } and core of the table2 is attribute TROUSER. We cannot eliminate attribute TROUSER because this is the most important attribute of the Table2. By using the confidence or strength (α) we will find another indispensible attribute of the table. The confidence or strength for an association rule x→ LOOK is the ratio of number of example that contain x U LOOK to the number of example that contain x.

We can calculate the strength of attribute (α) HAT, SHIRT and SHOES as follows:

α = example that contain X U Look / number of example that contain X

We can find the strength of rules for attribute HAT

(HAT =Black) and (Look=Smart), α = 66 %

(HAT = Red) and (Look=Smart), α = 33 %

(HAT = Red) and (Look= Good), α = 33 %

(HAT =Black) and (Look= Good), α = 33 %

(HAT = Red) and (Look= Acceptable), α = 33 %

(HAT =Bleu) and (Look= Acceptable), α = 100 %

We can find the strength of rules for attribute SHIRT

(SHIRT = Red) and (Look=Smart), α = 100 %

(SHIRT = White) and (Look= Good), α = 50 %

(SHIRT = White) and (Look= Acceptable), α = 50 %

(SHIRT = Grey) and (Look= Acceptable), α = 100 %

We can find the strength of rules for attribute SHOES

(SHOES = Sneakers) and (Look=Smart), α = 20 %

(SHOES = Oxford) and (Look=Smart), α = 100 %

(SHOES = Sneakers) and (Look= Good), α = 20 %

(SHOES = Monk) and (Look= Good), α = 60 %

(SHOES = Sneakers) and (Look= Acceptable), α = 100 %

After the calculation of α we can easily find that attribute SHIRT is indispensible among other attributes because the strength of rules for attribute SHIRT is maximum. The reduct of the set {HAT, SHIRT, SHOES, TROUSER} is {SHIRT, TROUSER}.

Table2 can be reduced to Table 3 as follows:

Table 3: Decision Table after calculating the strength of all attributes

	SHIRT	TROUSER	LOOK
1	Red	Black	Smart
2	Red	Bleu	Smart
3	Red	Black	Smart
4	White	Black	Good
5	White	Black	Good
6	White	Red	Acceptable
7	White	Red	Acceptable
8	Grey	Red	Acceptable

We can reduce Table3 by eliminating the same values of decision and condition attributes i.e we can merge different rows that has the same values for condition and decision attributes. This method is called Row Reduction

Table 4 will be:

Table 4: Decision Table after eliminating same values of decisions and conditions attributes

	SHIRT	TROUSER	LOOK
1	Red	Black	Smart
2	Red	Bleu	Smart
3	White	Black	Good
4	White	Red	Acceptable
5	Grey	Red	Acceptable

Find out the core of each example:

We will find the core of the Table 4 in such manner that the table will remain consistent.

If we eliminate TROUSER = Black there are two decision values Smart and Good. It means that based on attribute TROUSER we cannot make a unique decision, thus the value of SHIRT cannot be eliminated.

Similarly if we eliminate SHIRT = White there are two decision values Good and Acceptable It means that based on attribute SHIRT we cannot make a unique decision, thus the value of TROUSER cannot be eliminated.

Now the table 5 will be:

Table 5: Decision Table

	SHIRT	TROUSER	LOOK
1	Red	*	Smart
2	Red	*	Smart
3	White	Black	Good
4	*	Red	Acceptable
5	*	Red	Acceptable

Table 5 shows the core of each example. We can further reduced Table 5 by merging duplicate rows. Now we again eliminate the identical rows.

Finally the table 6 will be:

Table 6: Final table

	SHIRT	TROUSER	LOOK
1	Red	*	Smart
2	White	Black	Good
3	*	Red	Acceptable

Now, no further reduction is possible. Table 6 gives us the decision rules. Followings are the decisions rules based on reduct and core:

1. If SHIRT is Red then LOOK is Smart
2. If SHIRT is White and TROUSER is Black then LOOK is Good
3. If TROUSER is Red then LOOK is Acceptable

4. CONCLUSION

This paper presents a new approach for determining the most important attribute on the basis of strength of an

association. It is one of the most promising and new analytical approach of the Rough set theory that can be used for framing new decision rules. The application of this approach may be used extensively in the fields of knowledge discovery, data mining or any other field concerning attribute reduction and feature selection. As a direction for future research attempts may be made towards testing this method using some large databases and comparing this method with some others existing methods.

5. REFERENCES

[1] L.A. Zadeh, Fuzzy sets, Inf. Control 8 1965. 338–353.

[2] J.R. Mansfield, M.G. Sowa, G.B. Scarth, R.L. Somorjai, H.H. Mantsch, Analysis of spectroscopic imaging data by fuzzy C-means clustering, Anal. Chem. 69 1997. 3370–3374.

[3] O.N. Jensen, P. Mortensen, O. Vorm, M. Mann, Automation of matrix-assisted laser desorptionrionization mass spec-trometry using fuzzy logic feedback control, Anal. Chem. 69 1997. 1706–1714.

[4] M. Otto, Fuzzy theory. A promising tool for computerized chemistry, Anal. Chim. Acta 235 1990. 169–175.

[5] J.M. Mendel, Fuzzy logic systems for engineering: a tutorial, Proc. IEEE 3 3. 1995. 345–377.

[6] G.N. Chen, Assessment of environmental water with fuzzy clustering, analysis of fuzzy recognition, Anal. Chim. Acta 271 1992. 115.

[7] G.J. Postma, F.M. Hack, P.A.A. Janssen, L.M.C. Buydens, G. Kateman, A data based approach on analytical methods ap-plying fuzzy logic in the search strategy and flow charts for the representation of the retrieved analytical procedures, Chemometr. Intell. Lab. Syst. 25 1994. 285–295.

[8] M. Otto, H. Bandemer, A fuzzy method for component iden-tification and mixture evaluation in the ultraviolet spectral range, Anal. Chim. Acta 191 1986. 193–204.

[9] Y. Hu, J. Smeyers-Verbeke, D.L. Massart, An algorithm for fuzzy linear calibration, Chemometr. Intell. Lab. Syst. 8 1990. 143–155.

[10] W.P. Ziarko Ed. ., Rough Sets, Fuzzy Sets and Knowledge Discovery, Springer, New York, 1994.

[11] T.W. Collette, A.J. Szladow, Use of rough sets and spectral data for bulding predictive models of reaction rate constants, Appl. Spectrosc. 48 1994. 1379–1386.

[12] B. Walczak, D.L. Massart, Multiple outliers revisited, Chemometr. Intell. Lab. Syst. 41 1998. 1–15.

[13] R. Slowinski Ed. ., Intelligent Decision Support. Handbook of Applications and Advances of the Rough Sets Theory, Kluwer Academic Publishers, Dordrecht, 1992.

[14] Pal S.K., Skowron (Eds.), A. 1999. Rough Fuzzy Hybridization: A new trend in decision making. Springer-Verlag, Berlin.

[15] Pawlak, Z. 1982. Rough sets, International Journal of Computer and Information Sciences 11: 341–356.

[16] Pawlak, Z. 1991. Rough Sets: Theoretical Aspects of Reasoning about Data, System Theory, Knowledge Engineering and Problem Solving, vol. 9, Kluwer Academic Publishers, Dordrecht, The Netherlands.