THE INTERNATIONAL JOURNAL OF SCIENCE & TECHNOLEDGE

Attribute Selection in Product Mix: A Rough Set Approach

P. Seethalakshmi
Department of Basic Sciences, Kongu Polytechnic College, Perundurai, Erode, Tamil Nadu, India
S. Vengataasalam
Department of mathematics, Kongu Engineering College, Perundurai, Erode, Tamil Nadu, India

Abstract:
Using an illustrative case study on the Product mix, this paper propose the advantages of the rough set approach over conventional techniques for the extraction of decision rules with the help of confidence and strength of an association. As an important concept of rough set theory (RST), an attribute reduct is a subset of attributes that are jointly sufficient and individually necessary for preserving a particular property of the given information table. Our approach focus on the selection of the best attributes in order to generate the effective reduct set and formulating the core of the attribute set.

Key words: *Membership, Rough set, Reduct and core, Decision rules*

1. Introduction

Best product mix is considered essential to successful production based business and to good management. Due to the present competitive market and rapid variation of customers need, companies need to identify the best product mix. A huge set of data has no practical relevance unless it can be mined to provide useful information pertaining to the interests of the organization. The model developed in this paper includes some important issues related to select the best product mix that results in the maximization of number of products produced and profit; and the minimization of production time, subject to material availability, machine production time and yield rates. These multi objectives are simultaneously optimized in the proposed model.

Over the past ten years, RST has indeed become a topic of great interest to researchers and has been applied to many domains. In the theory of rough set, it is a kind of relatively brief and efficient algorithm to start with discernibility matrix for calculation of core and reduction(Rong-Rong Chen, Yen-I. Chiang, P. Pete Chong, Yung-Hsiu Lin, Her-Kun Chang, 2011). The so called discernibility matrix is a property set with elements in universe arranged in rows and columns and the elements in matrix distinguishing the row or column elements belonging to different categories. The element is a core when only one property is needed to distinguish the elements belonging to different categories. Adding properties in core is reduction. In fact, reduction is a property set including at least one property of each non empty element of discernibility matrix.

Given a dataset with discretized attribute values, it is possible to find a subset (termed as reduct) of the original attributes using RST that are the most informative; all other attributes can be removed from the dataset with very little information loss.

The main advantage of the rough set approach to product mix is that the decision rule performance model which is very convenient for decision support, because it generates simpler rules and removes the relevant attributes which reflects the quality of an attribute selection. First, a sample of solutions is taken for consideration. Next, from this data of information we need to express it as a function which maps each real number to a membership degree using fuzzy set. Then, by applying rough set approach, a decision table can be reduced by removing redundant attributes without any information loss. The concepts of reduct and core set are then used for rule discovery from the database. The fundamental concepts of the rough set approach are briefly explained in the following section.

2. Materials and Methodology

The rough set approach has several advantages over the conventional methods (Shen & Loh, 2004).This theory deals with representation, learning and generalization of uncertain knowledge (Huanglin, 1996; Jianguo, 2002; Pawlak, 1991) and its distillate is to reduce knowledge and produce concise decision rules without any prior information beyond the data set to be dealt with. However, the conventional techniques cannot reduce the data dimensions efficiently and the persistent redundant attributes would affect the rule discovery process, leading to highly degraded rules (Zhong, Dong, & Ohsugu, 2001). How to develop measures to automatically extract and evaluate interesting, relevant, and novel rules becomes an urgent and practical topic in this area. The rough set theory developed by (Pawlak, 2007) is based on the assumption that with every object of the universe there is associated a certain amount of information (data, knowledge), expressed by means of some attributes used for object description. All these factors have opened up the scope for some of the newer techniques which have been developed in recent years (Beynon, Curry, & Morgan, 2001).

3. Fuzzy Set Approach for Classification

Fuzzy Logic was initiated in 1965 by L.A. Zadeh. It is basically a multivalued logic that allows intermediate values to be defined between conventional evaluations like true/false, yes/no, high/low, etc. The fuzzy system is constructed based on the fuzzy set theory, fuzzy if-then rules and fuzzy reasoning. The main objective of this theory is to develop a methodology for the formulation and solution of problems that are too complex or ill-defined to be suitable for analysis by conventional Boolean techniques. A fuzzy set can be defined as a set of ordered pairs A = {x, $\mu_A(x)$/ $x \in U$}. The function $\mu_A(x)$ is called the membership function for A, mapping each element of the universe U to a membership degree in the range [0, 1]. An element $x \in U$ is said to be in a fuzzy set if and only if $\mu_A(x) > 0$ and to be a full member if and only if $\mu_A(x) = 1$(H. J. Zimmermann,1991). Membership functions can either be chosen by the user arbitrarily, based on the user's experience or they can be designed by using optimization procedures (Jang,J.S.R.,1992; Horikowa, S., T. Furahashi and Y. Uchikawa,1992).

The triangular membership function is defined as

$$\mu_A(X) = \text{triangular}(x; a, b, c) = \boxed{\begin{cases} 0, x \le a \\ \dfrac{x-a}{b-a}, a \le x \le b \\ \dfrac{c-x}{c-b}, b \le x \le c \\ 0, c \le x \end{cases}}$$

4. Rough Set Approach

Rough set theory has been used successfully as a selection tool to discover data dependencies and reduce the number of attributes contained in a data set by purely structural methods. Rough sets remove superfluous information by examining attribute dependencies. It deals with inconsistencies, uncertainty and incompleteness by imposing an upper and lower approximation to set membership. Given a data set with discretized attribute values, it is possible to find a subset (termed as reduct)of the original attributes using rough sets that are the most informative ; all other attributes can be removed from the data set with minimal information loss(R.Jensen, Q. Shen, 2005).

The first step in the rough set-based approach to the problem is defining a decision table that contains the whole knowledge about the universe of discourse (U). Columns of the decision table are defined by conditional (C) and decision (D) attributes while rows (X) specify values of these attributes (A=CUD) for each object of the Universe, which allow to partition U into equivalence classes ($[x]_A$) based on the notion of indiscernibility relation(Z. Pawlak, 1996). The indiscernibility relation and resulting from its equivalence classes enable to describe sets of objects by their lower $\underline{A}X$ and upper approximations $\bar{A}X$. In the lower approximation the objects of the universe are included for which the entire equivalence class is also included in the considered set, while the upper approximation is constructed with these objects for which at least one element of the equivalence class is included in the set. Set difference between the upper and lower approximation being empty indicates that the set is crisp, otherwise it is said to be rough.

Suppose we are given two finite and non empty sets U and A where U is called the universe and A, a set of attributes. With attribute $a \in A$, we associate a set V_a (value set) called the domain of a.

Any subset B of A determines a binary relation IND (B) on U which will be called an indiscernibility relation,

$$IND(B)=\{(x,y) \in U / \forall\ a \in B, a(x)=a(y)\}, \qquad (1)$$

where IND(B) is an equivalence relation and is called B-indiscernibility relation.The indiscernibility relation will be used now to define basic concept of Rough set theory.

Let us consider $B \subseteq A$ and $X \subseteq U$. We can approximate X by using only the information contained in B by constructing lower approximation (2) and upper approximation (3) of x in the following way:

$$\underline{B}(x) = \{ x \in U: B(x) \subseteq x \} \qquad (2) \text{ and}$$

$$\overline{B}(x) = \{ x \in U: B(x) \cap X \neq \phi \} \qquad (3)$$

Equivalence classes contained within X belongs to the lower approximation whereas equivalence classes within X and along its border form the upper approximation. Let P and Q be sets of attributes including equivalence relation over U, then the positive region is defined as

$$POS_p(Q) = \bigcup_{x \in U/Q} \underline{P}X \qquad (4)$$

where $POS_P(Q)$ compromises all objects of U that can be classified to classes U/Q using the information contained within attributes P.

4.1. Data Analysis and Attribute Selection

An example of product mix problem (Roman Slowinski,2008)is given to extract decision rule to select efficient solution of product mix considered as relatively good. The data set includes 13 solutions as shown in Table I.

			Produced Quantity				
Solution	Profit	Total Time	X_A	X_B	X_C	Sales	Product Mix
S_1	165	120	0	0	10	250	Poor
S_2	172.6923	130	0.7692	0	10	265.3846	Poor
S_3	180.3846	140	1.538	0	10	280.7692	Good
S_4	141.125	140	3	3	4.916667	272.9196	Good
S_5	148.375	150	5	2	4.75	378.75	Good
S_6	139.125	150	5	3	3.58333	279.583	Poor
S_7	188.0769	150	2.3076	0	10	296.153	Poor
S_8	139	150	6	0	6	270	Poor
S_9	140.5	150	6	2	3.666	271.6667	Good
S_{10}	109.25	200	6	2	7.8333	375.833	Poor
S_{11}	189.375	200	5	5	5.416	385.4167	Poor
S_{12}	127.375	130	2.3	3	4.083	252.08	Poor
S_{13}	113.625	120	3	3	3.25	231.25	Poor

Table 1: Data set of attributes

From Table 1, we consider six condition attributes: Profit, Total time, Quantity Produced – X_A, X_B and X_C and Sales. And a decision attribute denoted as Product mix represents whether the quantity produced is good or not.

Initially, in order to represent a continuous fuzzy set, we need to express it as a function which maps each real number to a membership degree. A very common parametric function is the triangular membership function. Membership functions are usually predefined by experienced experts. They can be derived through automatic adjustments(T.P.Hong, C.H.Chen, Y.L.Wu, Y.C.Lee, 2004). Each attribute have three fuzzy regions (Low, Medium and High) described as follows:-

Profit: Low (0,110,140) Medium (125,155,185) High (170,200,230)
Total time: Low (0,120,140) Medium (130,150,170) High (160,180,200)
(X_A): Low (0, 0.75, 2.25) Medium (1.5, 3.0, 4.5) High (3.75, 5.25, 6.75)
(X_B): Low (0, 1, 2) Medium (1.5, 2.5, 3.5) High (3, 4, 5)
(X_C): Low (0, 3, 5) Medium (4, 6, 8) High (7, 9, 11)
Sales: Low (0,230,270) M edium (250,290,330) High (310,350,390)

Thus three Fuzzy membership values are produced for each solution according to the predefined membership functions. The fuzzified result is shown in Table 2.

Solution	Profit	Total Time	Produced quantity			Sales	Product Mix
			X_A	X_B	X_C		
S_1	M	L	L	L	H	M	Poor
S_2	H	M	L	L	H	L	Poor
S_3	M	L	M	L	H	M	Good
S_4	L	L	M	H	L	L	Good
S_5	M	M	H	L	L	H	Good
S_6	L	M	H	H	M	L	Poor
S_7	M	M	L	L	H	M	Poor
S_8	L	M	H	L	M	L	Poor

S_9	L	M	H	L	M	L	Good
S_{10}	L	H	H	L	M	H	Poor
S_{11}	M	H	H	H	L	H	Poor
S_{12}	M	M	L	H	M	M	Poor
S_{13}	L	L	M	H	L	L	Poor

Table 2: Data set in fuzzy form

4.2. Calculation of Upper and Lower Approximations

The decision attribute (Product Mix) have two values, Good and Poor. Each value may be classified into its partition. For example solution 3, 4, 5 and 9 belong to partition X_G and solution 1, 2,6,7,8,10,11,12 and 13 belong to partition X_P. X_G={3,4,5,9}, X_P={1,2,6,7,8,10,11,12,13}.

The accuracy for each value class of the decisional variable is calculated by dividing the lower to the upper approximation of each class. In our decision Set Solution, S4 and S13 have all the same condition attributes values but with different decision (S4 has Good but in S13 there is Poor) so, this condition set will be regarded as uncertain and it will be ignored from the calculation of the lower approximation; but we will include it in the calculation of upper approximation. Thus: lower approximation = {S3, S5, S9} (Good), upper approximation = {S3, S4, S5, S9, S13} (Good). The boundary region = {S4, S13} (this can be classified either as Good or not-Good (Poor)). Using the same to concept of poor, we find the objects to have the lower approximation = {S1,S2,S6,S7,S8,S10,S11,S12}, the upper approximation { S1,S2,S4,S6,S7,S8,S10,S11,S12,S13}, and the boundary region we get is = { S4, S13}. The quality of lower approximation is (3 + 8)/(4 + 9), or 0.84615; the accuracy for Good is 0.6 (3/5); the accuracy for Poor is 0.80 (8/10); and the accuracy of the whole classification is (3 + 8)/(5 + 10), or 0.73.

	Good	**Poor**
Number of records	4	9
Number of lower approximation	3	8
Number of upper approximation	5	10
Accuracy	0.6	0.8

Table 3: The accuracy of classification using all condition attributes

Same or indiscernible objects may be represented many times and some of the attributes may be superfluous (redundant). That is, their removal cannot affect the classification.

Solution	**Profit**	**Total Time**	**Produced quantity**			**Sales**	**Product Mix**
			X_A	X_B	X_C		
S_1	M	L	L	L	H	M	Poor
S_2	H	M	L	L	H	L	Poor
S_3	M	L	M	L	H	M	Good
S_5	M	M	H	L	L	H	Good
S_6	L	M	H	H	M	L	Poor
S_7	M	M	L	L	H	M	Poor
S_8	L	M	H	L	M	L	Poor
S_9	L	M	H	L	M	L	Good
S_{10}	L	H	H	L	M	H	Poor

S_{11}	M	H	H	H	L	H	Poor
S_{12}	M	M	L	H	M	M	Poor

Table 4: Data set in consistency form

4.3. Reduction of Attributes

Now we will generate the rules based on reduct and core of Table 2. Reduct is the reduced set of relation that conserves the same inductive classification of relation. The set A of attributes is the reduct (or covering) of another set B of attributes if A is minimal and the indiscernibility relations, defined by A and B are same.

CoreB = ∩ ReductB, where RED(B) is the set off all reducts of B, the core is the intersection of all reducts and will include in every reduct. Therefore, the core is an important subset of attributes. Reduct of table4 are {Profit, Total time, produced quantity XA, XB, XC, Sales} and core of the table2 is attribute Sales. We cannot eliminate attribute sales because this is the most important attribute of the Table4. 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→ D is the ratio of number of example that contain x U D to the number of example that contain x.

For Table 4 we can calculate the strength of all the attributes as follows:

- (Profit =M) → (D =Good) strength of this particular rule comes out to be 66%.
- (Profit =L) → (D =Good) strength of this particular rule comes out to be 33%.
- (Profit =L) → (D =Poor) strength of this particular rule comes out to be 38%.
- (Profit =M) → (D = Poor) strength of this particular rule comes out to be 50%.
- (Profit =H) → (D = Poor) strength of this particular rule comes out to be 13%.
- (Total time =L) → (D =Good) strength of this particular rule comes out to be 33%.
- (Total time =M) → (D =Good) strength of this particular rule comes out to be 66%.
- (Total time =L) → (D =Poor) strength of this particular rule comes out to be 13%.
- (Total time =M) → (D = Poor) strength of this particular rule comes out to be 63%.
- (Total time =H) → (D = Poor) strength of this particular rule comes out to be 25%.
- Similarly we can find the strength of rules for the other attributes.

From these calculations we can easily find that attribute y is indispensible among other attributes because the strength of rules for attribute y is maximum. The reduct of the set {Profit, Total time, produced quantity XA, XB, XC, Sales} is {Profit, Total time, produced quantity XB, Sales}. Table4 can be reduced to Table 5 as follows.

Solution	**Profit**	**Total Time**	**Produced quantity**	**Sales**	**Product Mix**
			$\mathbf{X_A}$		
S_1	M	L	L	M	Poor
S_2	H	M	L	L	Poor
S_3	M	L	M	M	Good
S_5	M	M	H	H	Good
S_6	L	M	H	L	Poor
S_7	M	M	L	M	Poor
S_8	L	M	H	L	Poor
S_9	L	M	H	L	Good
S_{10}	L	H	H	H	Poor
S_{11}	M	H	H	H	Poor
S_{12}	M	M	L	M	Poor

Table 5: Data after the reduction of attributes

Reduce Table 5 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.

Solution	Profit	Total Time	Produced X_A	Sales	Product Mix
S_1	M	L	L	M	Poor
S_2	H	M	L	L	Poor
S_3	M	L	M	M	Good
S_5	M	M	H	H	Good
S_7	M	M	L	M	Poor
S_{10}	L	H	H	H	Poor
S_{11}	M	H	H	H	Poor

Table 6: Ultimate version of given data

It is clear from the table that we get the same result after the reduction also and the data has got reduced and the following table is the ultimate version of the Table 1.Then build the discerning matrix. Discern = $(dis_{ij})_{7x6}$ where $dis_{ij} = \{r/r \in C, r(Equiv_i) \neq r(Equiv_j)$. Reduct í of an equivalence class should be able to distinguish Equiv í from all other equivalence classes. Reduct í should be the joint of the entries in the íth row of the discerning matrix. Finally, the decision table can be built to extract the rules.

Class	Solution	Profit	Total Time	ProducedquantityX_A	Sales	Product Mix
Equiv 1	S_1	-	L	-	-	Poor
Equiv 2	S_2	H	-	-	L	Poor
Equiv 3	S_3	-	-	M	-	Good
Equiv 4	S_5	-	M	-	-	Good
Equiv 5	S_7	-	M	L	M	Poor
Equiv 6	S_{10}	L	-	-	-	Poor
Equiv 7	S_{11}	M	H	-	-	Poor

Table 7: Decision table for Rule extraction

4.4. Extraction of rules

Here, the attributes needed for the classification has been reduced. Thus superfluous data is removed from our table and we can extract decision rules in IF-THEN form. Here the condition attribute value (Profit, total time, produced quantity X_A and sales) is used as the rule antecedent and class label attribute (Product mix) as the rule consequent. Hence, we can extract the following decision rules:

- If (Total time, low) then (Product mix , Poor)
- If (Profit,High) and (Sales,low) then (Product mix , Poor)
- If (Produced quantity X_A, medium)then (Product mix , Good)
- If (Total time, medium)then (Product mix , Good)
- If (Total time&Sales, medium) and (Produced quantity X_A ,low)then (Product mix , Poor)
- If (Profit,low) then (Product mix, Poor)
- If (Profit, medium) and (Total time, high) then (Product mix, Poor).

5. Conclusion

In this paper we have introduced a new approach for the selection of important attributes on the basis of strength of an association. At the same time, under the framework of this space, the consistent attribute set is established, from which we can obtain the approach to attribute reductions in information systems. The object of this work is to extract decision rules for modification of attributes, which minimizes or removes disadvantages.

As a direction for future research, attempts may be made towards testing this method using some large databases and broadening the comparative studies to include comparisons with other feature selection and dimensionality reduction techniques.

6. References

1. Rong-Rong Chen, Yen-I. Chiang, P. Pete Chong, Yung-Hsiu Lin, Her-Kun Chang(2011).Rough set analysis on call center metrics, Applied Soft Computing 11, pp.3804–3811.
2. Shen, L., & Loh, H. T. (2004). ‘Applying rough sets to market timing decisions’. Decision Support Systems, 37(4), 583e597.
3. Huanglin, Z., (1996). Rough Sets Theory and its Applications. Chongqing University Press, Cbongqing.

4. Jianguo, T., (2002). Reliability analysis to deal with imperfect information by rough set theory. Control Dec, 17(2): 255-256.
5. Pawlak, Z., (1991). Rough Set-Theoretical Aspects of Reasoning about Data. Kluwer Academic Publishers, Dordrecht.
6. Zhong, N., Dong, J., & Ohsugu, S. (2001).'Using rough sets with heuristics for feature selection'. Journal of Intelligent Information Systems, 16(3), 199e214.
7. Pawlak, Z., & Skowron, A. (2007). 'Rudiments of rough sets'. Information Sciences, 177(1), 3e27.
8. Beynon, M., Curry, B., & Morgan, P. (2001). 'Knowledge discovery in marketing: An approach through rough set theory'. European Journal of Marketing, 35(7/8), 915e937.
9. Zadeh, L.A. (1965): Fuzzy sets. Inf. Control 8, 338–353.
10. H. J. Zimmermann, (1991).Fuzzy Set Theory and Its Applications, Kluwer Academic Publishers,.
11. Jang, J.S. R., (1992)Self-Learning Fuzzy Controllers Based on Temporal Back-Propagation, IEEE Trans. On Neural Network, Vol. 3, September, pp. 714723.
12. Horikowa, S., T. Furahashi and Y. Uchikawa,(1992).On Fuzzy Modeling Using Fuzzy Neural Networks with Back-Propagation Algorithm, IEEE Trans. on Neural Networks, Vol.3, Sept.,pp.801-806.
13. R.Jensen, Q. Shen, (2005). "Fuzzy-Rough Data Reduction With Ant Colony Optimization", Fuzzy Sets and Systems, vol. 149, no.1, pp. 5-20,.
14. Z. Pawlak, Rough Set Rudiments,(1996). Institute of Computer Science Report,Warsaw University of Technology, Poland, pp. 1–47.
15. Roman Slowinski(2008): Rough Set Approach to Multiple Criteria Decision Aiding , Poznan University of Technology, Poland ,Chania, October 1st,.148-161.
16. T.P. Hong, C.H. Chen, Y.L. Wu, Y.C. Lee, (20040.Using Divide-and- Conquer GA Strategy in Fuzzy Data Mining, The Ninth IEEE Symposium on Computers and Communications