International Journal of Mechanical, Industrial and Aerospace Sciences
ISSN: 2517-9950
Vol:5, No:6, 2011

A Rough-set Based Approach to Design an Expert System for Personnel Selection

Ehsan Akhlaghi

***Abstract*—**Effective employee selection is a critical component of a successful organization. Many important criteria for personnel selection such as decision-making ability, adaptability, ambition, and self-organization are naturally vague and imprecise to evaluate. The rough sets theory (RST) as a new mathematical approach to vagueness and uncertainty is a very well suited tool to deal with qualitative data and various decision problems. This paper provides conceptual, descriptive, and simulation results, concentrating chiefly on human resources and personnel selection factors. The current research derives certain decision rules which are able to facilitate personnel selection and identifies several significant features based on an empirical study conducted in an IT company in Iran.

***Keywords*—**Decision Making, Expert System, Personnel Selection, Rough Set Theory

I. Introduction

PERSONNEL selection is the process of collecting and evaluating information about individuals and choosing those who match the qualifications needed to perform a predefined job in the best way. This process plays a determining role in human resource management and is crucial to the success of an organization.References [1] and [2] reviewed the personnel selection studies and found that the several main factors including change in organizations, change in work, change in personnel, change in the society, change of laws, and change in marketing have influenced personnel selection.In literature, there are a number of studies which use heuristic methods for employee selection. A fuzzy MCDM framework based on the concepts of ideal and anti-ideal solutions for the selection of the most appropriate candidate is presented in [3]. Also, a fuzzy number ranking method by metric distance for personnel selection problem was proposed in [4] and a personnel selection system based on fuzzy AHP was developed in [5]. In addition, researchers used fuzzy technique for order preference by similarity (TOPSIS) based on the veto threshold for ranking job applicants [6].Recently, owing to the advancements in information technology, researchers have developed decision support systems and expert systems to improve the outcomes of human resource management [7]. A model to design an expert system for effective selection and appointment of the job applicants developed in [8].Although the applications of expert systems or decision support systems on personnel selection and recruitment are increasing (e.g., [9] and [10]), little research has employed rough-sets based approaches for personnel selection as the present study does. In [11] authors presented an empirical study using RST to analyze human resource data for personal selection and human capital enhancement. Yet, a number of researches used rough-sets to deal with different kinds of decision-making problems. A rough set based approach to distributor selection in supply chain management was developed in [12] and several certain decision rules for distributor selection based on a study conducted in China were extracted. In addition, a novel extension to Fuzzy QFD (Quality Function Development) methodology by combining fuzzy arithmetic operations with the concepts of rough number and rough boundary interval which are derived from rough set theory was presented in [13].This paper, proposes a rough set based methodology which effectively inducts recruitment rules. Moreover, the weight of each input attribute is incorporated in the proposed approach so as to enhance quality of the derived rules.The rest of the paper is organized as follows: A brief overview on rough set theory and rough set-based rule identification will be presented in the next section. Section 3 describes a case study conducted in an IT company. Section 4 shows the results and proves their validity. Section 5 consists of conclusion and further research directions.

Ehsan Akhlaghi is with the Industrial Engineering Faculty, Islamic Azad University, Tehran South Branch, Tehran, Iran (phone: +98-912-3062328; fax: +98-21-44436566; e-mail: st_e_akhlaghi@azad.ac.ir).

II. Rough sets

A. Definitions

The rough set theory, firstly introduced by Pawlak [14], has proved to be a powerful tool for uncertainty and has been applied to data reduction, rule extraction, data mining and granularity computation [15].

In rough set theory, data is represented as data or decision tables. A data table can be expressed by a 4-tuple information system $S=(U,Q,V,f)$ where $U=\{x_1,x_2,...,x_n\}$ is a finite set of objects (universe), $A=\{a_1,a_2,...,a_m\}$ is a finite set of attributes (criteria). Moreover, V is the set of all attribute values such that $V=\bigcup_{a\in A}V_a$ and V_a is the domain of the attribute (criterion) a, and $f:U\times A\to V$ is an information function such that $f(x_i,a)\in V_a$ for each $a\in A$ and $x_i\in U$. The set A is usually divided into set C of condition attributes and set D of decision attributes $A=C\cup D$.

In RST, the objects in universe U can be described by various attributes in A. When two different objects are described by the same attributes, then these two objects are classified as one kind in the information system *S*, thus we call their relationship is indiscernibility relation. In mathematical word, an indiscernibility relation *IND(B)* generated by attribute subset $B \subset A$ on U, is defined as follows:

$$IND(B) = \{(x_i, x_j) \in U \times U \mid f(x_i, a) = f(x_j, a), \forall a \in B\} \quad (1)$$

As an example, Table I shows a decision table with six objects that are characterized with three condition attributes (A1, A2, A3) and one decision attribute (B).

Reduct is a fundamental construct in RST which is defined as a minimal sufficient subset of attributes ($\mathrm{Re}\,duct \subseteq A$) that produces the same categorization of objects as the collection of all attributes.For generating reducts, the algorithm shown in Fig. 1 can be used, according to [15].An instance of generated reducts (nine reducts) for the objects in Table I is shown in Table II.

TABLE II
REDUCTS GENERATED FOR OBJECTS IN TABLE I

Reduct No.	Object No.	A1	A2	A3	B
1	1	x	0	x	1
2	2	1	x	x	2
3	2	x	2	x	2
4	3	x	0	x	1
5	4	x	1	x	0
6	4	x	x	2	0
7	5	0	3	x	1
8	5	0	x	1	1
9	6	2	3	x	2

[a]Sign “x” implies that the attribute is not considered to determine the output of an object.

TABLE III
SELECTED ATTRIBUTES FOR PERSONNEL SELECTION

Sign	Attribute	Weight
A1	Technical Skill	0.9
A2	Creativity	0.88
A3	Self-Organization	0.7
A4	Reasoning Ability	0.73
A5	Motivation	0.8
A6	Team working	1.0

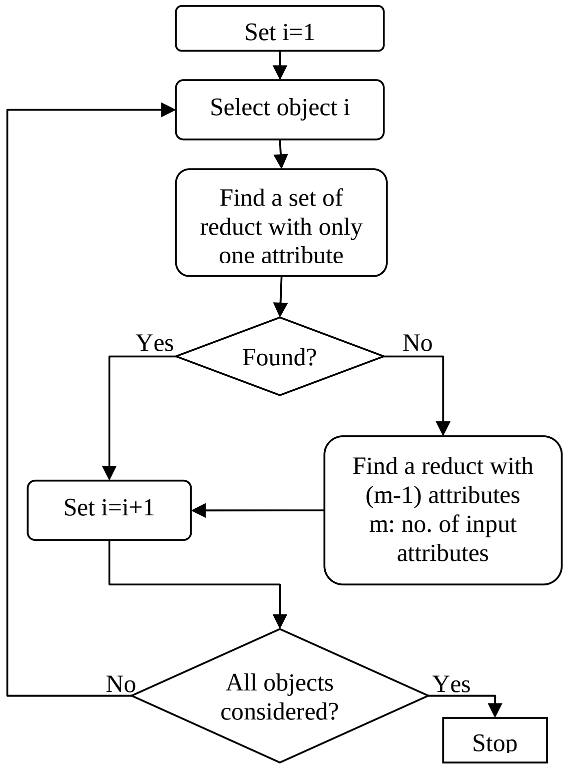

Fig. 1 Reduct generation flow chart

B. Rule Identification

In order to produce decision rules, a conceptual framework which has the following steps: problem definition, data preparation, data partition, reduct generation, and rule validation has been introduced in [12]. In the first two steps, the problem should be identified clearly, preprocessing tasks should be done, and the data set should be randomly divided into two parts; training data set and testing data set. The training data set is used to build the model and the testing data set is used to detect over fitting of the modeling tools. According to a study [17], the partitioning can be performed using bootstrapping method and the ratios for training and testing data would be 0.632 and 0.368, respectively.

In the third step (reduct generation), the algorithm previously mentioned in Fig. 1 can be used. Using the algorithm all the existing reducts can be generated, but in the real circumstances we know that all the attributes of an object have not the same importance. Thus we choose weight for each attribute. Afterwards, we can select the reducts with the greater weights of attributes. This may reduce the final number of reducts (more important ones) to produce rules. For example, if the weight of attribute A3 in Table II is very low, we can ignore the reducts using A3 from the rule generation process (reducts number 6 and 8). In this case, an instance rule with object 6 would be:

International Journal of Mechanical, Industrial and Aerospace Sciences
ISSN: 2517-9950
Vol:5, No:6, 2011

IF (condition-attribute A1 is 2) AND
(condition-attribute A2 is 3)
THEN decision-attribute B is 2

In the last step the generated rules should be validated by comparing any decision rule with the objects in the testing data. By dividing the total matched objects by the total incorrectly matched objects, the accuracy of each rule will be determined. Comparing the calculated accuracy of a rule with a predefined threshold, we decide to confirm or reject the rule. The rejected rules should be removed.

III. Case Study

In order to recruit IT professionals for an engineering company in Iran, ten attributes were collected by senior manager of the company from the literature and consulting with experienced professionals. Then, these attributes were sent to ten experts from several successful IT companies. The experts selected six main attributes. Later, these six attributes were sent to the experts again for weighting. The average weight from all ten experts considered as the final weight of each attribute. Table III depicts the extracted attributes and corresponding weightings.

Afterwards, a team of company managers were asked for providing the score of six attributes for their personnel. By scoring ninety two staff, according to aforementioned partitioning method (%63.2 training data and %36.8 testing data), 58 objects were selected for training randomly (Fig. 2), and the rest (34 objects) for testing (Fig. 3).

Each attribute has been classified into three levels: "low" represented by "0", "middle" represented by "1", and "high" presented by "2". Also, the output attribute, which is a measure of personnel performance, was classified into "weak" (0), "normal" (1), and "good" (2). Table IV shows several candidate rules.

IV. Results

Using the RSES software (Warsaw University Rough Set Exploration System, Logic Group, Inst. Mathematics) version 2.2, (Fig. 4) data analysis was performed and the derived candidate rules were examined by testing data set. The RSES generates candidate decision rules and shows the number of matched objects for each candidate rule. The number of candidate rules in this study was 1271 rules (Fig. 5). The decision rules supported with more matched objects would be considered as the process results. The results are listed as 4 rules depicted in Table V.

Table: Personnel_Part_0.632

58 / 7	A1	A2	A3	A4	A5	A6	O
O:1	0	1	1	1	2	0	0
O:2	2	1	1	0	0	0	0
O:3	1	2	0	1	2	2	1
O:4	2	1	2	1	2	2	2
O:5	2	0	0	1	1	0	0
O:6	2	0	1	0	1	0	0
O:7	1	0	1	2	0	0	0
O:8	0	1	2	1	1	0	1
O:9	1	0	0	2	2	1	1
O:10	1	1	2	0	0	2	1

Fig. 2 Training data

Table: Personnel_Part_0.368

34 / 7	A1	A2	A3	A4	A5	A6	O
O:1	1	2	1	2	2	1	1
O:2	1	1	1	0	1	1	1
O:3	2	2	2	1	0	0	1
O:4	2	2	1	2	0	2	2
O:5	2	1	1	1	2	2	2
O:6	1	2	2	2	2	1	2
O:7	2	2	2	0	0	0	1
O:8	2	2	1	2	2	2	2
O:9	1	1	1	0	1	2	1
O:10	1	1	1	0	1	1	1

Fig. 3 Testing data

TABLE IV
SEVERAL CANDIDATE RULES FOR PERSONNEL SELECTION

Reduct No.	A1	A2	A3	A4	A5	A6	B
1	0	x	x	x	x	x	0
2	1	x	1	1	2	x	1
3	2	x	1	x	1	2	1
4	x	0	0	0	x	x	0
5	2	0	x	x	x	2	1
6	2	2	x	x	x	2	2
Weight	0.9	0.88	0.7	0.73	0.8	1.0	

TABLE V
FINAL EXTRACTED RULES AND TESTING RESULTS

Rule	Matched	Total	Accuracy
IF "Technical Skill" is "medium" AND "Team working" is "high" THEN the performance is "good"	47	49	95.9%
IF "Technical Skill" is "medium" AND "Creativity" is "high" AND "Motivation" is "medium" THEN the performance is "normal"	29	33	87.8%
IF "Motivation" is "high" AND "Reasoning Ability" is "high" THEN the performance is "normal"	12	17	70.5%
IF "Creativity" is "low" AND "Team working" is "middle" AND "Self-Organization" is "low" THEN the performance is "weak"	21	26	80.7%

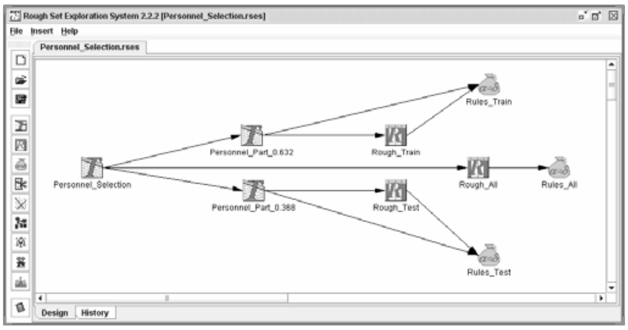

Fig. 4 RSES software for data analysis

Rule set: Rules_All

(1-1271)	Match	Decision rules
8	11	(A1=1)&(A3=2)=>(O={1[4],2[6],0[1]})
9	4	(A1=0)&(A3=0)=>(O={0[4]})
10	1	(A1=0)&(A2=1)&(A4=1)&(A5=2)=>(O={0[1]})
11	7	(A1=1)&(A2=2)&(A4=2)&(A5=2)=>(O={1[3],2[4]})
12	1	(A1=2)&(A2=1)&(A4=0)&(A5=0)=>(O={0[1]})
13	1	(A1=1)&(A2=2)&(A4=1)&(A5=2)=>(O={1[1]})
14	4	(A1=2)&(A2=1)&(A4=1)&(A5=2)=>(O={2[4]})

Fig. 5 Several candidate decision rules generated by RSES

V. Conclusion

Personnel selection, as a crucial part of human resources management, is analyzed in this paper using rough set theory. The method generates several useful rules for incorporates to select job applicants effectively. Since the results have been checked for accuracy, they can be used in IT companies' recruitment procedure as it is used in an IT company in Iran.

Further research should be focusing on personnel selection in other fields and for other positions in a company. Also, it seems to be useful to weight experts' opinions in order to benefit more experienced experts' ideas. Furthermore, the number of experts and objects (employees) may be increased, by working on a bigger company or a group of companies as the case study. This study was conducted in Iran and may be influenced by Iranian management culture.

References

[1] W. C. Borman, M. A. Hanson, J. W. Hedge, "Personnel selection," Annual Review of Psychology, pp. 299–337, 1997.

[2] I. T. Robertson, M. Smith, "Personnel selection," Journal of Occupational and Organizational Psychology", pp.441–472, 2001.

[3] E. E. Karsak, "Personnel selection using a fuzzy MCDM approach based on ideal and anti-ideal solutions," Lecture Notes in Economics and Mathematical Systems, pp. 393–402, 2001.

[4] L. S. Chen, C. H. Cheng, "Selecting IS personnel use fuzzy GDSS based on metric distance method," European Journal of Operational Research, pp. 803–820, 2005.

[5] Z. Güngör, G. Serhadlioglu, S. E. Kesen, "A fuzzy AHP approach to personnel selection problem," Applied Soft Computing, vol 9, pp. 641–646, 2009.

[6] A. Kelemenis, D. Askounis, "A new TOPSIS-based multi-criteria approach to personnel selection," Expert Systems with Applications, vol 37, pp. 4999-5008, 2010.

[7] C. F. Chien, L. F. Chen, "Data mining to improve personnel selection and enhance human capital: A case study in high-technology industry," Expert Systems with Applications, vol 34(1), pp. 280–290, 2008.

[8] M. S. Mehrabad, M. F. Brojeny, "The development of an expert system for effective selection and appointment of the jobs applicants in human resource management," Computers & Industrial Engineering, vol 53, pp. 306–312, 2007.

[9] R. S. Hooper, T. P. Galvin, R. A. Kilmer, J. Liebowitz, "Use of an expert system in a personnel selection process," Expert Systems with Applications, vol. 14(4), pp. 425–432, 1998.

[10] M. Nussbaum et al, "A decision support system for conflict diagnosis in personnel selection," Information and Management, vol. 36, pp. 55-62, Jan. 1999.

[11] C. F. Chien, L. F. Chen, "Using rough set theory to recruit and retain high potential talents for semiconductor manufacturing," IEEE Transactions on Semiconductor Manufacturing, vol 20, pp. 528-541, 2007.

[12] Z. Zou, T. L. Tseng, H. Sohn, G. Song, R. Gutierrez, "A rough set based approach to distributor selection in supply chain management," Expert Systems with Applications, vol. 38, pp. 106-115, 2011.

[13] L.Y. Zhai, L. P. Khoo, Z. W. Zhong, "Towards a QFD-based expert system: A novel extension to fuzzy QFD methodology using rough set theory," Expert Systems with Applications, vol 37, pp. 8888-8896, 2010.

[14] Z. Pawlak, "Rough sets," International Journal of Computer and Information Sciences, vol 11, pp. 341–356, 1982.

[15] C. C. Yeh, D. J. Chi, M. F. Hsu, "A hybrid approach of DEA, rough set and support vector machines for business failure prediction," Expert Systems with Applications, doi:10.1016/j.eswa.2009.06.088, 2009.

[16] Z. Pawlak, "Rough sets", Dordrecht, The Netherlands: Kluwer Academic Publishers, 1991.

[17] A. Kusiak, "Rough set theory: A data mining tool for semiconductor manufacturing," IEEE Transactions on Electronics Packaging Manufacturing, vol 24, pp. 44–50, 2001.