Rough Sets Based Association Rules Application for Knowledge-Based System Design

Shu-Hsien Liao[1] and Yin-Ju Chen[2]

[1] Department of Management Sciences and Decision Making, Tamkang University, No. 151, Yingjuan Road, Danshuei Jen, 251 Taipei, Taiwan, ROC
[2] Department of Management Sciences, Tamkang University, No. 151, Yingjuan Road, Danshuei Jen, 251 Taipei, Taiwan, ROC
michael@mail.tku.edu.tw, s5515124@ms18.hinet.net

Abstract. The Internet has emerged as the primary database, and technological platform for electronic business (EB), including the emergence of online retail concerns. Knowledge collection, verification, distribution, storage, and re-use are all essential elements in retail. They are required for decision-making or problem solving by expert consultants, as well as for the accumulation of customers and market knowledge for use by managers in their attempts to increase sales. Previous data mining algorithms usually assumed that input data was precise and clean, this assumes would be eliminated if the best rule for each particular situation. The Algorithm we used in this study however, proved to function even when the input data was vague and unclean. We provided an assessment model of brand trust as an example, to show that the algorithm was able to provide decision makers additional reliable information, in the hope of building a rough set theoretical model and base of resources that would better suit user demand.

Keywords: Machine Learning, Knowledge Representation, Knowledge-Based Systems, Rough sets, Association rules.

1 Introduction

In recent years, there has been a growing use of data mining and machine learning outside the computer science community. Such methods are being integrated into decision support systems for fields such as finance, marketing, insurance, and medicine [8]. Machine learning methods are well known for knowledge discovery. They can help to elicit knowledge (explicit and tacit) [9, 10] from data and generalize that knowledge to new, previously unseen cases [11]. Data with missing values may also be helpful if we need to guess at the correct classification of a condition using an ‘‘incomplete’’ table, for example by looking for rows that almost cover a condition [6]. One of the main tools of data mining is rule induction from raw data represented by a database. Real-life data are frequently imperfect: erroneous, incomplete, uncertain and vague [12]. Of the various data mining algorithms, this analysis uses the rough set algorithm due to its ability to deal with incomplete, imprecise or inconsistent information, which is typical in credit assessment analyses [4, 5]. Rough set theory is different from the fuzzy theory

J.-S. Pan, S.-M. Chen, and N.T. Nguyen (Eds.): ICCCI 2010, Part II, LNAI 6422, pp. 501–510, 2010.

and neural network. In addition, the majority of scholars mentioned that using association rules that need to face uncertainty or information inaccurate information. In the research, we further through the use of rough set theory to improve the thorny issue of the encounter for using association rules.

We can be to judge the decision of consumers relying on rules of thumb, when consumer choice factors taken into account are simple. But when the variety of choices as well as the growing number of factors to consider, how a simple analysis and consumer rule of thumb helping to determine the shopping behavior of consumers has become an important issue. At this time we may need more rigorous approach to help us to determine future consumer decision-making, and to find a complex combination of factors, and these effects of factors are tangible or intangible. The remainder of this paper is organized as follows. Section 2 reviews relevant literature correlate with the research and problem statement. Section 3 Mathematical models for new algorithm. Section 4 is presented an illustrative example. Closing remarks and future work are presented in Sect. 5.

2 Literature Review and Problem Statement

Machine learning can extract desired knowledge from existing training examples and ease the development bottleneck in building expert systems [9]. Knowledge-Based Systems are interactive computer programs that mimic and automate the decision-making and reasoning processes of human experts. Brand familiarity reflects the “share of mind” of a given consumer attained to the particular brand and the extent of a consumer's direct and indirect experience with a brand argue that brand familiarity is determined by strength of associations that the brand name evokes in consumer memory, and in this way it captures the consumer's brand attitude schemata [2]. The high levels of prior experience with a brand lead to the retention of stronger advertisement- brand links, making the attributes of previously familiar brands easier to recall [3]. In view of this, we hope to proceed to that the brand or product from the consumer's subjective or objective point of view preferences, according to the above-mentioned scholars, through establishing the brand image of the trust evaluation model by the ratio scale algorithms, combined with rough set theory, then find the influence degree of the consumer preferences variables between each other for the marketing decision-makers used.

3 Incorporation of Rough Set for Classification Processing

The knowledge in this Knowledge-Based System (KBS) consists of descriptions of domain classes and class hierarchies. We propose a classification of processed to describe these hierarchies, which could provide decision makers with addition information.

Definition 1: The questionnaire is $X = \{x_1, x_2, \cdots x_n\}$, where $x_i \in X$. The questionnaire answer item is $A_{ij} = \{(a_{11}, a_{12} \cdots a_{1j}), (a_{21,} a_{22} \cdots a_{2j}), \cdots (a_{i1}, a_{i2} \cdots a_{ij})\}$. The decision attribute is $D = \{d_1, d_2, \cdots, d_m\}$.

Example 1: Consumers of beer "brand recall" who sort and respondents related to what answer information are shown in Table 1 and 2.

Table 1. The original questionnaire answer

No.	Questionnaire X			Decision-making D
	Item1 x_1	Item2 x_2	Item3 x_3	Item4
	Gender A_1	Age A_2	Income A_3	Beer brand recall
1	Male a_{11}	30 a_{21}	35,000 a_{31}	As shown in Table 2.
2	Male a_{11}	40 a_{22}	60,000 a_{32}	As shown in Table 2.
3	Male a_{11}	45 a_{23}	80,000 a_{34}	As shown in Table 2.
4	Female a_{12}	30 a_{21}	35,000 a_{31}	As shown in Table 2.
5	Male a_{11}	40 a_{22}	70,000 a_{33}	As shown in Table 2.

Table 2. Beer brand recall ranking table

No.	First	Second	Third	Fourth
	Fifth	Sixth	Seventh	Eighth
1	Heineken	Miller	Taiwan light beer	Taiwan beer
	Taiwan draft beer	Tsingtao	Kirin	Budweiser
2	Taiwan beer	Heineken	Taiwan light beer	Kirin
	Taiwan draft beer	Tsingtao	Miller	Budweiser
3	Taiwan beer	Miller	Taiwan light beer	Heineken
	Taiwan draft beer	Tsingtao	Kirin	Budweiser
4	Heineken	Miller	Taiwan beer	Tsingtao
	Taiwan draft beer	Taiwan light beer	Budweiser	Kirin
5	Taiwan beer	Miller	Heineken	Tsingtao
	Taiwan draft beer	Taiwan light beer	Budweiser	Kirin

Definition 2: All the respondents, whose answers were the same, are denoted as $R_k^l = \bigcap_{i=1}^{n} A_{ij}(x_i)$, where $k = 1 \cdots n$ is rule number and $l = 1 \cdots c$ is the number of accumulated customers.

Example 2: According to Table 1, $R_1^1 = \{a_{11}, a_{21}, a_{31}\}$ indicated that male gender, age 30, income of 35,000.

Definition 3: We further discussion between the items and the items are there hidden relationships exist, that is, $R_k^l = \bigcap_{i=1}^{n} A_{ij}(x_i)$ cross-comparison of the relationship between the ratio of x_i ◦ $X_{A_{ij}} = \dfrac{\sum (A_{ij} - A_{ij})}{\sqrt{\sum (A_{ij} - A_{ij})^2}} = \dfrac{X_{A_i} - X_{A_j}}{\sqrt{(X_{A_i} - X_{A_j})'(X_{A_i} - X_{A_j})}}$,

$-1 < X_{A_{ij}} < 1$, if $X_{A_{ij}} < 0$ indicated a negative correlation between the two attributes; $X_{A_{ij}} > 0$ indicated a positive correlation between the two attributes; $X_{A_{ij}} = 0$ that is no related between the two attributes. If they are in categories (nominal) variables and numerical variables are mixed, $X_{A_{ij}} = \frac{\sum\left(\overline{A_{ij}} - \overline{A_{ij}}\right)}{\sqrt{\sum\left(\overline{A_{ij}} - \overline{A_{ij}}\right)^2}}\sqrt{P_{ij}} = \frac{\overline{X_{A_i}} - \overline{X_{A_j}}}{\sqrt{\left(\overline{X_{A_i}} - \overline{X_{A_j}}\right)'\left(\overline{X_{A_i}} - X_{A_j}\right)}}\sqrt{P_i P_j}$, where $\overline{A_{ij}}$ is denoted as mean of each category and P_{ij} is denoted as percentage of each category.

Example 3: Taking the ratio of the relationship between gender, age, and income as an example, as shown in Table 1, we find $X_{A_{12}} = 0$ indicates that there was no relation between sex and age; $X_{A_{23}} > 0$ indicates a positive correlation between age and income. An increase in age accompanied an increase in income; with lower age, income was reduced. It also indicated that there was no relation between gender and income. The ratios of the relationship between gender, age and income are shown in Table 3, where A_{ij}^{+} denotes a positive correlation between *i* and *j*. A_{ij}^{-} denotes a negative correlation between *i* and *j*. Δ_{ij} denotes no different between the two categories (nominal).

Table 3. The ratio of the relationship between gender, age and income

Y_R	Gender and age A_1A_2	Age and income A_2A_3	Income and gender A_3A_1
y_{12}	$\left(a_{11}^{+},+\right)$	$(+,+)$	$\left(+,a_{11}^{+}\right)$
y_{23}	$\left(a_{11}^{+},+\right)$	$(+,+)$	$\left(+,a_{11}^{+}\right)$
y_{34}	$(\Delta_{1112},-)$	$(-,-)$	$(-,\Delta_{1112})$
y_{45}	$(\Delta_{1112},+)$	$(+,+)$	$(+,\Delta_{1112})$
y_{51}	$\left(a_{11}^{-},-\right)$	$(-,-)$	$\left(-,a_{11}^{+}\right)$
$X_{A_{ij}}$	$X_{A_{12}} = 0$	$X_{A_{23}} > 0$	$X_{A_{31}} = 0$

Definition 4: We wanted to discover the hidden relationships between items, and whether they indicated another relationship with the set of ordinal attributes (D_c^{order}). Where $D_c^{order} = \{(d_{11}, d_{12}, \cdots d_{1j}), (d_{21}, d_{22}, \cdots d_{2j}), \cdots (d_{i1}, d_{i2}, \cdots d_{ij})\}$ is a set of ordinal attributes that could be divided into several subsets.

Example 4: The correlation between the recall of Beer brand sequence with age and income, as shown in Table 4.

Table 4. The relationship between beer brand recall sequence with age and income

Y_R	$X_{A_{23}}$	Decision-making attributes set(D)Beer brand recall sequence
y_{12}	$(+,+)$	$D_1^{order} = \{d_2^1, d_4^2, d_3^3, d_1^4, d5_5^5, d_6^6, d_7^7, d_8^8\}$
y_{23}	$(+,+)$	$D_2^{order} = \{d_1^1, d_2^2, d_3^3, d_7^4, d_5^5, d_6^6, d_4^7, d_8^8\}$
y_{34}	$(-,-)$	$D_3^{order} = \{d_1^1, d_4^2, d_3^3, d_2^4, d_5^5, d_6^6, d_7^7, d_8^8\}$
y_{45}	$(+,+)$	$D_4^{order} = \{d_2^1, d_4^2, d_1^3, d_6^4, d_5^5, d_3^6, d_8^7, d_7^8\}$
y_{51}	$(-,-)$	$D_5^{order} = \{d_1^1, d_4^2, d_2^3, d_6^4, d_5^5, d_3^6, d_8^7, d_7^8\}$

Definition 5: We want to discover the hidden relationship between the decision-making attribute (Beer brand recall sequence) and the ratio of the definition3 found $X_{A_{23}} > 0$ (Age and income have a positive ratio of inter-relationships). The hidden relationship is denoted as $\delta_D^A = \wedge X_{A_{ij}} \underset{R_k^l}{\wedge} f\left(y_{ij}^D, y_{ij}^A\right)$ representing all of the hidden relationships, where $f\left(y_{ij}^D, y_{ij}^A\right)$ is defined as:

$$f\left(y_{ij}^D\right) = \left\{ \begin{array}{l} y_{D_{ij}}^+ = \left\{ y_{ij}^A \in A_{ij}, y_{ij}^D \in D_{ij} : \dfrac{D_c^{order}}{D_c^{order}} > 0 \forall y_{ij}^D \right\} \\ y_{D_{ij}}^- = \left\{ y_{ij}^A \in A_{ij}, y_{ij}^D \in D_{ij} : \dfrac{D_c^{order}}{D_c^{order}} < 0 \forall y_{ij}^D \right\} \\ y_{D_{ij}}^0 = \left\{ y_{ij}^A \in A_{ij}, y_{ij}^D \in D_{ij} : \dfrac{D_c^{order}}{D_c^{order}} = 1 \forall y_{ij}^D \right\} \end{array} \right\}$$

$$f\left(y_{ij}^A\right) = \left\{ \begin{array}{l} y_{A_{ij}}^+ = \left\{ y_{ij}^A \in A_{ij}, y_{ij}^D \in D_{ij} : \dfrac{y_{ij}^A}{y_{ij}^D} > 0 \right\} \\ y_{A_{ij}}^- = \left\{ y_{ij}^A \in A_{ij}, y_{ij}^D \in D_{ij} : \dfrac{y_{ij}^A}{y_{ij}^D} < 0 \right\} \\ y_{A_{ij}}^0 = \left\{ y_{ij}^A \in A_{ij}, y_{ij}^D \in D_{ij} : \dfrac{y_{ij}^A}{y_{ij}^D} = 0 \right\} \\ y_{A_{ij}}^{D_i^{order}} = \left\{ y_{ij}^A \in A_{ij}, y_{ij}^D \in D_{ij} : \dfrac{y_{ij}^A}{y_{ij}^D} = 1 \right\} \end{array} \right\}$$

$[Y_R]_{cord}$ indicates a set defining the same hidden relationship between the decision-making attributes and the ratio found in definition3.

Example 5: We take the Customer No. 1 and No. 2 as an example of the relationship between beer brand recall sequence of decision-making attribute of Taiwan Beer d_1, as shown in table 5, is denoted as $y^+_{d_1}$ and according to definition3, the relationship between the ratio of x_i is shown $y^+_{A_{23}}$, then is mean that decision attribute and items with the same orientation relationship ($\delta^{A_{23}}_{d_1} = +$). We can be found from the above, as the "age and income" increases, the rank of Taiwan Beer, Tsingtao, and Budweiser are also rise, while with the "age and income" increases, the rank of Heineken and Kirin are also decreases, then $[Y_R]_{cord} = \{(d_1, d_6, d_8) \cap (d_2, d_7)\}$. In addition, Taiwan draft beer is always ranked fifth in the table.

Table 5. The relationship between beer brand recall

y^D_{ij} \ y^A_{ij}	$X_{A_{23}} > 0$					$f(y^A_{ij})$	δ
	y_{12} $(+,+)$	y_{23} $(+,+)$	y_{34} $(-,-)$	y_{45} $(+,+)$	y_{51} $(-,-)$		
Taiwan beer	$y^+_{d_1}$	$y^0_{d_1}$	$y^-_{d_1}$	$y^+_{d_1}$	$y^+_{d_1}$	same $y^+_{A_{23}}$	$\delta^{A_{23}}_{d_1} = +$
Heineken	$y^-_{d_2}$	$y^-_{d_2}$	$y^+_{d_2}$	$y^-_{d_2}$	$y^+_{d_2}$	opposite $y^-_{A_{23}}$	$\delta^{A_{23}}_{d_2} = -$
Miller	$y^+_{d_4}$	$y^-_{d_4}$	$y^+_{d_4}$	$y^+_{d_4}$	$y^-_{d_4}$	inconsistent $y^0_{A_{23}}$	$\delta^{A_{23}}_{d_4} = 0$
Taiwan draft beer	$y^0_{d_5}$	$y^0_{d_5}$	$y^0_{d_5}$	$y^0_{d_5}$	$y^0_{d_5}$	ranked fifth $y^{d^5_5}_{A_{23}}$	$\delta^{A_{23}}_{d_5} = 5$

Definition 6: We use the rough set theory concept of approximation in the algorithm. According to Definition3, the lower estimate, denoted as $Lower_{Y_R}$, is defined as the union of all these elementary sets, which contained in R^l_k. More formally:

$$Lower_{Y_R} = \{R^l_k \in A_{ij}(x_i) \,|\, [Y_R]_{core} \subset R^l_k\}$$

The upper estimate, denoted as $Upper_{Y_R}$, is the union of these elementary sets, which have a non-empty intersection with R^l_k.

$$Upper_{Y_R} = \{R^l_k \in A_{ij}(x_i) \,|\, [Y_R]_{core} \cap R^l_k \neq 0\}$$

The difference: $Boundary_{Y_R} = Upper_{Y_R} - Lower_{Y_R}$ is called a boundary of R^l_k.

Example 6: According to Definition3, we know

$$Lower_{y_{di}} = \{ X_{A_2}, X_{A_3} \} = \{age, income\} \text{ and}$$

$$Upper_{y_{di}} = \{(X_{A_1}, X_{A_2}), (X_{A_1}, X_{A_3}), \cdots (X_{a_{12}}, X_{a_{35}})\}$$
$$= \{(sex, age), (sex, income) \cdots (sex - female, income - 80{,}000)\}$$

Definition 7: Let us define an assessment model to establish brand trust E_i, as follows: $E_c = \alpha \times \beta(a_{ij} / x_i / d_{ij})$, where α refers to the weight of assessment model, using the upper and lower bounds established by Definition6, while the attribute is included in $Lower_{Y_R}$, depending on attribute ranking computing with double and while the attribute is included in $Upper_{Y_R}$, depending on attribute ranking. And $\beta = \lfloor (n+1) - D^{order} \rfloor$ is mean the rankings of the scores, where n is denoted the number of ordinal Decision-making attributes set and D^{order} is after that sort of ranking.

Example 7: $E_1^{x_2 x_3} = 2 \times 8 \times \text{Heineken}(d_2) = 16$, is meaning that the assessment model to establish brand trust of Customer No. 1 in the condition in considering the age and income is 16. $E_3^{x_2 x_3} = 2 \times 5 \times \text{Heineken}(d_2) = 10$, is mean that assessment model to establish brand trust of Customer No. 3 in the condition in considering the age and income is 10.

Definition 8: Finally, $\frac{\sum E_c}{c}$ is the total trust value found in Definition7 that can provide marketing decision-maker as the basis for brand preference.

Example 8: We can calculate the total trust value of the Heineken brand for all customers $\frac{\sum E_c}{c} = 13.5$, and the higher the total value of the trust, the higher the degree of trust. The degree of brand trust with the brand recall ranking of decision-making attribute of Taiwan Beer, as Table 6.

Table 6. The brand trust level

Customer number / D_{ij}	$E_c^{x_2 x_3}$					Total trust value
	1	2	3	4	5	
Taiwan beer	10	16	16	12	16	14
Heineken	16	14	10	16	12	13.6
Taiwan light beer	12	12	12	6	6	9.6
Miller	14	4	14	14	14	12

4 Example Application

4.1 Rough Set Implementation

In the research, one thousand questionnaires were distributed; 900 were returned, of which 115 were disqualified as incomplete or invalid. This left 785 valid questionnaires, yielding a valid completion rate of 78.5%. First, we tried to determine the potential relationships through the regression, as shown in Table 7.

Table 7. Potential relationships through the regression (Gender, Age, and Income)

	Linear combination
1	Age＊0.531+Income＊(-0.847)

Second, we used the concept of redaction of rough set theory to find the core value. The redaction set is shown in Table 8. Pos. Reg. indicates the positive region of the reduction table and SC indicates the reduction of the stability coefficient.

Table 8. Redact set (Gender, Age, Income and Brand recall ranking)

Set	Pos. Reg.	SC	Redacts
1	0.997	1	{ Income, Beer brand recall ranking2, Beer brand recall ranking3, Beer brand recall ranking4, Beer brand recall ranking5, Beer brand recall ranking6, Beer brand recall ranking7}
2	0.997	1	{ Age , Beer brand recall ranking1, Beer brand recall ranking2, Beer brand recall ranking3, Beer brand recall ranking4, Beer brand recall ranking5, Beer brand recall ranking6, Beer brand recall ranking7 }

Table 7 shows a negative correlation between age and income. According to Table 10, we find that{ Income, Beer brand recall ranking2, Beer brand recall ranking3, Beer brand recall ranking4, Beer brand recall ranking5, Beer brand recall ranking6 and Beer brand recall ranking7} are core attributes. The rule set is shown in Table 9.

Table 9. Rule set (Age, Income and Brand recall ranking)

	Match	Decision Rules
1	12	{(Income= Below NT$5,000)&(Beer brand recall ranking 2= Taiwan draft beer)&(Beer brand recall ranking 4= Budweiser) =>(Beer brand recall ranking 7= Heineken)}
2	11	{(Income= Below NT$5,000)&(Beer brand recall ranking 3= Taiwan light beer)&(Beer brand recall ranking 4= Budweiser)=>(Beer brand recall ranking 7= Heineken) }

Third, we tried to determine the potential relationship through regression, as shown in Table 10. This indicates a positive correlation between income and Heineken, and a negative correlation between income and Miller. A positive relationship existed between age and Taiwan_Beer, as well as between age and Taiwan_Light_Beer .

Table 10. Potential relationships through regression (Age, Income and Brand recall ranking)

	Linear combination
1	Income ＊0.346+Heineken＊0.938
2	Income＊0.049+Miller＊(-0.665)+Tsingtao＊0.744

According to the potential relationships found in Table 7 and Table 10, we established the degree of Brand Trust, as shown in Table 11.

Table 11. Total Brand trust

Brand	Brand trust	Brand	Brand trust
Taiwan beer d_1	12.77	Taiwan draft beer d_5	8.63
Heineken d_2	12.09	Tsingtao d_6	8.21
Taiwan light beer d_3	9.52	Kirin d_7	8.71
Miller d_4	6.51	Budweiser d_8	5.55

4.2 Association Rule Generation

Finally, we took the general attributes including Channels, Consumer Behavior, Product Features, and Medium into account, with Heineken as a decision-making variable. The rule sets are shown below.

Table 12. Rule set (Channels, Consumer Behavior, Product Feature, and Medium)

	Match	Decision Rules
1	177	{(Channels= Convenience Stores)&(Product Feature= Price)&(Medium= Advertising)=>(Heineken)}
2	128	{(Consumer Behavior= Purchase by promotions)&(Channels= Hypermarkets)&(Product Feature= Price)&(Medium= Advertising)=>(Heineken)}
3	76	{(Consumer Behavior= Purchase by promotions) &(Product Feature= Flavor)&(Medium= Advertising)=>(Heineken)}

5 Conclusion and Future Works

Traditional association rules should be adjusted to avoid retaining only trivial rules, or discarding interesting rules. In fact, situations using relative comparisons were more complete than those using absolute comparisons. In this paper, a new approach was used to determine rules of association possessing the ability to handle uncertainty in the process of classification suitable for ratio scale data. We established a brand trust evaluation model $E_c = \alpha \times \beta(a_{ij} / x_i / d_{ij})$ in the first step of data processing, weighing them according to the upper and lower bounds set for the attributes. The system had to be readjusted to optimize the new rules, while the conditions of the traditional association rules changed. In the study, an extension of the concept of utility function used to establish the demand for users to adjust the brand image with brand trust evaluation model. The purpose and benefits for adjusting rules:

- Traditional rules of association can only generate rules, no function to amend rules. Through adjusted by the weight α can increase the convenience for using association rules.
- Expert information can be used to adjust the weighting, in order to increase the credibility of the rules.

Knowledge engineers can acquire and represent knowledge in the form of decision tables, which are easily transformed to rules for use in knowledge-based systems [6]. It is our hope that in the future, we will be able to build a decision-making resources system based on rough set theory, which more closely matches the needs of users.

Acknowledgements. This research was funded by the National Science Council, Taiwan, Republic of China, under contract No. NSC 98-2410-H-032 -038–MY2.

References

1. Liao, S.H., Ho, H.H., Yang, F.C.: Ontology-based data mining approach implemented on exploring product and brand spectrum. Expert Systems with Applications 36, 11730–11744 (2009)
2. Mikhailitchenko, A., Javalgi Rajshekhar (Raj) G., Mikhailitchenko, G., Laroche, M.: Cross-cultural advertising communication: Visual imagery, brand familiarity, and brand recall. Journal of Business Research 62, 931–938 (2009)
3. Kent, R.J., Kellaris, J.J.: Competitive interference effects in memory for advertising: are familiar brands exempt? J. Mark Commun. 7, 59–69 (2001)
4. Liu, G., Zhu, Y.: Credit Assessment of Contractors: A Rough Set Method. Tsinghua Science & Technology 11, 357–363 (2006)
5. Jiayuan, G., Chankong, V.: Rough set-based approach to rule generation and rule induction. International Journal of General Systems 31, 601–617 (2002)
6. Hewett, R., Leuchner, J.: Restructuring decision tables for elucidation of knowledge. Data and Knowledge Engineering 46, 271–290 (2003)
7. Hong, T.-P., Tseng, L.-H., Chien, B.-C.: Mining from incomplete quantitative data by fuzzy rough sets. Expert Systems with Applications 37, 2644–2653 (2010)
8. Štrumbelj, E., Kononenko, I., Robnik Šikonja, M.: Explaining instance classifications with interactions of subsets of feature values. Data & Knowledge Engineering 68, 886–904 (2009)
9. Brohman, M.K.: Knowledge creation opportunities in the data mining process. In: HICSS 2006: Proceedings of the 39th Annual Hawaii International Conference on System Sciences, p. 170. IEE Computer Society, Washington (2006)
10. Nemati, H.R., Steiger, D.M., Iyer, L.S., Herschel, R.T.: Knowledge warehouse: an architectural integration of knowledge management, decision support, artificial intelligence and data warehousing. Decision Support Systems 33(2), 143–161 (2002)
11. Chi, C.-L., Street, W.N., Ward, M.M.: Building a hospital referral expert system with a Prediction and Optimization-Based Decision Support System algorithm. Journal of Biomedical Informatics 41, 371–386 (2008)
12. Fortes, I., Mora-López, L., Morales, R., Triguero, F.: Inductive learning models with missing values. Mathematical and Computer Modelling 44, 790–806 (2006)