Rough Set Theory and Rule Induction Techniques for Discovery of Attribute Dependencies in Medical Information Systems

Jerzy Stefanowski[1] and Krzysztof Słowiński[2]

[1] Institute of Computing Science, Poznań University of Technology, 3A Piotrowo, 60-965 Poznań, Poland, E-mail: Jerzy.Stefanowski@cs.put.poznan.pl
[2] Clinic of Traumatology, K.Marcinkowski University of Medical Sciences in Poznań, 1/2 Długa Street, 61-848 Poznań, Poland

Abstract. Problems connected with applications of the rough set theory to identify the most important attributes and with induction of decision rules from the medical data set are discussed in this paper. The medical data set concerns patients with multiple injuries. The direct use of the original rough set model leads to finding too many possibilities of reducing the input data. To solve this difficulty, a new approach integrating rough set theory, rule induction and statistical techniques is introduced. First, the Chi–square test is additionally performed in order to reject non-significant attributes. Then, starting from remaining attributes we try to construct such definitions of new attributes that improve finally discovered decision rules. The results have shown that the proposed approach integrating all methods has given better results than those obtained by applying the original rough set method.

1 Introduction

One of the essential issues in the analysis of experience represented in data sets is looking for dependencies among attributes describing objects of our interest. It is usually assumed that objects are described by two kinds of attributes: condition (features characterizing object properties by means of numerical and symbolic values) and decision (referring to decision, diagnosis and, in general, expressing classification of objects) attributes.

In this paper, we consider the discovery of attribute dependencies as two subtasks:

– discovery of redundant attributes in the representation of dependencies between condition and decision attributes and identification of the most important condition attributes,
– discovery of decision rules characterizing the dependency between values of condition attributes and decision attributes.

In recent years the authors tried to solve the above tasks in the analysis of various data sets, mainly medical ones, using several data mining approaches (see e.g. [8, 9, 11]). These approaches were mostly based on the *rough set theory* and *rule induction* techniques [7, 4, 12, 14].

Our experiences show that in some of the analysed cases it was not possible to solve the considered tasks in one-step process by the direct use of a general-purpose technique. These applications required that the use of the chosen techniques had to be supported by human expertise, in particular concerning the changes of attribute definitions in the input data, e.g. choosing the proper attributes for the given problem, rescaling and discretizing attribute domains. We think that the proper choice and definition (or redefinition) of input attributes is one of the most important factor influencing the quality of the final results of the discovery process.

By good final results in the case of considered two subtasks and chosen techniques we understand:

- unambiguous identification of the most significant attributes (the direct use of the rough set theory may lead to finding too many possible reducts what makes the final choice quite difficult),
- discovery of the smallest set of decision rules having good discrimination properties.

The aim of this paper is to present a new approach for machine construction of such a definition of condition attributes that leads to good final results of applying the rough set and rule induction techniques. In the current study decision rules are induced by means of the LEM2 algorithm (originally introduced by Grzymala, see e.g. [4]).

The proposed approach integrates the rough set theory, rule induction and some statistical techniques. It performs two additional stages of data analysis when the preliminary use of rough set and rule induction techniques leads to unsatisfactory results. The first stage consists in using statistical tests for supporting rejection of irrelevant attributes. The second stage consists in constructing possible definitions of new attributes which are logical combinations of remaining ones. The aim of this stage is to get such a combination of new and old attributes that leads to the most satisfactory final results.

The approach integrating all of the considered techniques has been applied to analyse the medical information system containing patients with *multiple injuries*. This is a real life problem coming from the Clinic of Traumatology, K.Marcinkowski University of Medical Sciences in Poznan, and concerns patients of the emergency unit whose life is directly threatened due to severe multiple injuries. Thus, the discovered dependencies between attributes expressing clinical status of patients in time of admission to the hospital and results of treatment (live or dead) are particularly important for fast diagnosing, proper ordering of medical procedures and reducing the threat of the patients' life [1, 5].

The paper is organised as follows. Section 2 gives the brief description of the rough set and LEM2 rule induction algorithm; then the proposed approach is presented. The results of the analysis of the medical information system are given in section 3. Final remarks are discussed in section 4.

2 Brief description of the rough sets and rule induction techniques

2.1 Rough set theory

Rough set theory was introduced by Z.Pawlak [7] as a tool to deal with uncertainty and vagueness in the analysis of information systems. An information system is a formal representation of the analysed data set and is defined as a pair $S = (U, A)$ where U is a finite set of *objects* and A is a finite set of *attributes*. With every attribute $a \in A$, set of its values V_a is associated. Each attribute a determines a function f_a: $U \to V_a$. In practice, we are mostly interested in discovering dependencies in a special case of information system called *decision table.* It is a pair $(U, A \cup \{d\})$, where $d \notin A$ is a distinguished decision attribute. The decision attribute determines the partition of U into k disjoint classes $X_1, X_2, \ldots, X_k$ (where k is a number of different values of attribute d) called *decision classes.*

The rough set theory is based on an observation that objects may be *indiscernible* (indistinguishable) due to limited available information. For a subset of attributes $B \subseteq A$ the indiscernibility relation is defined by $I(B) = \{(x,y) \in U \times U : f_a(x) = f_b(y), \forall a \in B\}$. The indiscernibility relation defined in this way is an equivalence relation. The classes of this relation are called *B-elementary sets.* An elementary equivalence class (i.e. single block of the partition $U/I(B)$) containing element x is denoted by $I_B(x)$.

The precise definition of the subset of objects $X \subseteq U$ in terms of elementary sets is not always possible because some elementary sets may be *inconsistent* (i.e. containing examples described by the same values of condition attributes and assigned to different decision classes). The idea of a *rough set* consists in approximating a set by a pair of sets called its *lower* and *upper* approximations. Formally, if $B \subseteq A$ and $X \subseteq U$ then the sets: $\{x \in U : I_B(x) \subseteq X\}$, $\{x \in U : I_B(x) \cap X \neq \emptyset\}$ are called B-lower and B-upper approximations of X, denoted by $\underline{B}X$ and $\bar{B}X$, respectively. The difference between upper and lower approximation is called a *boundary region.*

Next, we are interested in an approximation of the partition of U by d into decision classes $X_1, X_2, \ldots, X_k$. The set $\underline{B}X_1 \cup \underline{B}X_2 \cup \ldots \cup \underline{B}X_k$ is called *B-lower approximation of classification* induced by d. The dependency between condition attributes B and decision attribute d is characterized by coefficient called *quality of approximation of classification*

$$\gamma_B(d) = \sum_{i=1}^{k} card(\underline{B}X_i)/card(U)$$

This coefficient is sometimes called degree of dependency (see e.g. [7, 13]). If $\gamma_B(d) = 1$ then d totally depends on B otherwise $(0 < \gamma_B(d) < 1)$ d partly depends on B.

Discovering whether the information system can be simplified or not is performed on the basis of reducts and a core of attributes. A *reduct* is a minimal subset of attributes that ensures the same quality of classification as the entire set of attributes. In general, the information system may contain more reducts than one. A *core* is an intersection of all reducts in the information system.

Complexity of computing all reducts in the information system is rather high. However we are often interested in finding one or a few reducts instead of computing all ones. This could be done by using the strategy based on adding to the core the attributes of the highest increase of discriminatory power (for more details see [9]). In this strategy, the core of attributes is chosen as a starting reduced subset of attributes. A single remaining attribute is temporarily added to the core and the influence of this adding on the change of the quality is examined. Such an examination is repeated for all attributes. The attribute with the highest increase of the quality of classification is chosen to be added to the reduced subset of attributes. Then, the procedure is repeated until an acceptable quality of the classification is obtained.

2.2 Inducing decision rules

In the current study it is assumed that decision rules are represented in the following form:

$$IF\ (a_1 = v_1)\&(a_2 = v_2)\&\ \ldots\&(a_n = v_n)\ \ THEN\ class_j$$

where a_i is the ith condition attribute, v_i is its value and $class_j$ is formula referring to ith decision classes or its approximation. Formula $(a_i = v_i)$ is called an elementary condition.

Induction of decision rules from the multiple-injured patient data is performed by authors' implementation of LEM2 algorithm. The algorithm induces, so called, *discriminating rules* from lower approximations of decision classes. These rules distinguish positive examples, i.e. objects belonging to the lower approximation of the decision class, from other objects. Additionally *approximate rules* (i.e. non-discriminating ones) are induced from boundary regions of decision classes. Each rule has a non–redundant condition part and moreover the number of rules is minimized. The algorithm is precisely described in [4, 3].

To evaluate the discovered rules the measure of the rule *strength* is used. It is the number of objects in the information system whose description satisfy the condition part of the rule. Generally, one is interested in discovering the strongest rules.

Additionally we estimate the *classification accuracy.* The estimation of the classification accuracy is done by using the standard 'leaving-one-out' reclassification technique. While performing the classification, possible ambiguity in matching the testing example to decision rules is solved by VCR approach [10], where the decision is chosen on the basis of multiple or partly matched rules.

2.3 Multistrategic approach to determining the relevant attributes

Let us notice that the direct application of the rough set theory to analyse the information system may results in finding too many reducts what makes the task of identifying the most important attributes very difficult. To support the analysis of such data sets we propose a multistrategic approach.

The first stage of this approach integrates rule learning techniques and statistical tests. The goal of this stage is to simplify the input data by focusing attention on the attributes relevant to the objects' classification. This integration has been inspired by the approach originally presented in [6], where Chi-square test and analysis of condition parts of induced rules were considered.

Following this inspiration the proposed integration includes two strategies: (1) performing Chi-square test to check the dependency between each condition attributes and the decision attribute, (2) computing occurrence ratio for each attribute used to construct condition parts of the induced decision rules.

In the first strategy, the standard Chi-square test for qualitative attributes is used to determine the dependency between the decision attribute and all of the other condition attributes. The value of the Chi-square statistics is calculated. If it is lower than the critical value for an assumed confidence level, the given attribute is treated as irrelevant.

In the second strategy, it is assumed that irrelevant condition attributes are those which do not occur in any of decision rules induced by LEM2 algorithm or occur in rules that match less than a specified percentage of the maximum number of matched objects. It is expressed by the occurrence ratio which is defined as $O(a_j) = n(a_j)/(max(n(a_j))$ where $n(a_j)$ is the total number of objects that are covered by any rule containing attribute a_j in one of its elementary conditions; $max(n(a_j))$ is a maximum number of covered examples for one of the considered attributes.

One can notice that the second strategy corresponds to positive examples only while the Chi-square test uses both positive and negative examples for the given classification. We decided to take the result of Chi-square as the main criterion for identifying relevant and irrelevant attributes. The result of the second strategy is taken into account as a second criterion in case when too many attributes are treated by the first strategy as irrelevant ones. So, taking into account results of both strategies we can try to determine which attributes are irrelevant and then for the reduced information system, we can perform the analysis once again.

If the results are not satisfactory we start the second stage of the approach. This stage consists in an attempt of creating the definition of new attributes and substituting some original attributes by new ones. It is somehow connected with postulates of *a constructive induction* to the change of the input data (cf. [2]). This stage can be summarized in the following points:

1. Construct the set of possible definitions of new attributes being the logical combinations of original attributes; this construction is guided by medical background knowledge, i.e. for each new defined attribute there are given necessary logical rules transforming values of old attributes into new ones,
2. Create a list of possible and allowed (taking into account medical point of view) combinations of new defined attributes with remaining original ones,

3. For all of these combinations compute reducts, induce the set of decision rules, evaluate their strength and classification accuracy,
4. Choose this combination giving the most satisfactory results.

3 Discovery of attribute dependencies in multiple injuries data set

3.1 Analysis of original data set

The analysed data set concerned the multiple injured patients admitted to the Clinic of Traumatology of K.Marcinkowski University of Medical Sciences in Poznań. The data set consisted of 80 patients completely described by results of standard examinations performed in the time of the patient's admission to the hospital. For each patient the diagnostic peritoneal lavage (DPL) examination was additionally performed. All these patients were under the standard medical treatment for multiple injuries and we knew the results of this treatment (lived or died) for each of them.

The information system representing the clinical experience with treatment of the multiple injured patients was finally constructed using 22 *condition attributes* expressing the patient's status in the time of his admission to the hospital and one *decision attribute* referring to results of treatment. These were the following attributes: 1 -age, 2 -sex, 3 -cause of accident, 4 -other victims of accident, 5 -pulse, 6 -arterial pressure, 7 -skin paleness, 8 -abdominal pain, 9 -muscular defence, 10 -Blumberg sign, 11 -peristalsis, 12 -state of consciousness, 13 -presence of chest bruises, 14 -rib fractures 15 -pneumothorax, 16 -pleural haematoma, 17 -bruises of limbs, 18 -fracture of long bones, 19 -amputation of limb after injury, 20 -alcohol odour from mouth, 21 -time from the admission to the hospital to the execution of DPL, 22 -DPL result.

The *decision attribute* was very important from the clinical point of view. However, it defined patients' classification into two decision classes which were highly unbalanced taking into account their cardinality: first class - *lived* (65 patients), second class - *died* (15 patients). Unfortunately, we could not increase the number of patients, in particular for the second decision class because the DPL examination was performed for the limited number of patients admitted to this hospital emergency unit. The DPL result was very important from the medical point of view.

We are aware of the fact that the chosen data set was 'difficult' for the analysis (e.g. taking into number of objects and definition of decision classes). However, such available data seems to be typical for the analysis of real clinical problems.

The main aims of performed analysis from the medical point of view were the following: identifying the most significant attributes for the results of patients' treatment; discovering decision rules representing dependencies between values of attributes and prediction of final results of treatment.

First we used the computer program RoughDAS to perform basic operations of the rough set theory and to induce decision rules. The obtained results are presented in Table 1.

Table 1. Results of the analysis of original (22 attributes) information system by means of rough set operations and LEM2 induction algorithm. Numbers in brackets refer to decision classes.

quality of classification	the core	number of reducts	number of rules	average rule strength	classification accuracy
1.00	attribute 21	1295	20 (12/8)	6.65 (9.67/2.13)	70.00%

The decision attribute totally depended on chosen condition attributes (quality of approximation of classification was equal to 1.00). However, we had to consider too many reducts, i.e. 1295 ones. The core contained one element, i.e. attribute 21 (time from admission to execution of DPL). The strategy of adding the most discriminatory attributes to the core could not be applied because there were too many (over 10) attributes which could be added at the same time giving as a result the same (in fact, very small) increase of the quality. To sum up, it was not possible to identify the most important attributes. Although some of decision rules were evaluated by the practitioners as interesting and consistent with clinical experience, the values of measures evaluating these rules could be questionable and discussed. The number of rules was limited but besides a few strong rules for first decision class other rules were rather week (especially for 'died' decision class). The classification accuracy was not high and rules for 'died' did not perform well. On the other hand, an attempt of using another classification system, i.e. Quinlan's C4.5 system, did not also led to higher values of average classification accuracy.

3.2 Determining the relevant attributes for the discovery process using statistical tests

The unsatisfactory results of the direct application of the rough set theory encouraged us to use the proposed multistrategic approach.

First, we performed the standard Chi-square test for each condition attribute against the decision one. We assumed a confidence level equal to 10 %. As a result we could identify that the following attributes were the most significant: 3 (0.0379), 4 (0.0584), 6 (0.0787), 12 (0.0326), 19 (0.0361), 22 (0.0794); where the number in brackets is the probability from the Chi-square distribution for the calculated value of the statistics and the appropriate degrees of freedom. On the other hand some attributes seemed to be strongly irrelevant according to the calculated value of statistics, in particular 9 (0.9410), or 13 (0.7715).

As the first strategy indicated few significant attributes only while others seemed to have an intermediate character we tried to examine other irrelevant attributes basing on the analysis of decision rules. So, we calculated the occurrence ratio for each attribute in the decision rules induced by LEM2 algorithm. The attributes with the highest occurrence ratio were the following: 5 (1.0), 7 (0.74), 18 (0.77), 20 (0.43), 22 (0.86); numbers in brackets are the occurrence rations. On the other hand some attributes were characterized by very low occurrence ratio, i.e. 17 (0.0 - did not occur in any of the rule), and 11 (0.035), 19 (0.035). Other attributes had also values of occurrence ration quite low - below 0.5. As the occurrence ratio is treated by us as secondary criterion, we decided to extend the subset of irrelevant attributes by taking attribute 17 only.

To sum up, we removed attributes 9,13 and 17 from the information system and performed the rough set and rule induction analysis once again. However, the results were far from being satisfactory - the number of reducts still high (over 700), 21 decision rules of average strength 6.48 and classification accuracy 73.8%.

3.3 Redefinition of condition attributes

The second stage of the analysis consisted in an attempt of constructing the new attributes being a logical combination of 'weaker' attributed remained as a result of the previous stage. We assumed that we would like to get such a definition of new attributes which could lead to:

- obtaining no more than 40 reducts (with a non-empty core)

- inducing the number of rules no greater than 25, having the averaged strength and the classification accuracy higher than for the originally defined information system.

The construction of possible new attributes was guided by medical experts' background knowledge. Each of proposed attributes is a logical combination of old attributes having the similar clinical meaning. For instance, two original attributes: no. 3 - cause of accident (having five values: traffic, fall from height, blunt blow in abdomen, stabbed wound, unknown) and no. 4 - other victims of the same accident (having two possible values: none and present) could be transformed into a new attribute a3 referring to general cause of accident. The new attribute a3 had the following values: traffic with other victims, traffic without other victims, fall from height, blunt blow in abdomen, stabbed wound, unknown. Other proposed attributes are the following:

a4 - signs of hiporolenic shock (combination of original attributes 5 and 6),

a6 - abdominal signs (combination of original attributes 8, 10 and 11),

a7 - state of consciousness and alcohol (combination of original attributes 12 and 20),

a8 - chest injuries (combination of original attributes 14, 15 and 16),
Original attributes 9, 13 and 17 were already removed in the previous stage. Other attributes remained unchanged (their combination was not consistent with medical experience), i.e. they were a1 - age, a2 - sex, a7 - skin paleness,

a9 - fracture of long bones, a10 - amputation of limb, a11 - time to DPL, a12 - DPL result.

We performed the analysis of new defined information system (the above subset of attributes is denoted by S_1). The results are presented in Table 2. One can notice that these results are quite satisfactory according to our assumptions. On the other hand, in subset S_1 we used all possible redefinitions of attributes. So, one could be interested in checking whether all possible redefinitions of attributes are necessary or it would be sufficient to introduce few of them only. We checked some other possible combinations of attributes that contained part of new redefined attributes. These combinations were created from subset S_1 by temporary skipping one of possible new definition and restoring original attributes used in its definition. This was, in fact, a kind of backward elimination procedure performed for each of new attribute definition.

Proceeding in this way, we discovered two interesting subsets S_2 and S_3 (see Table 2). The subset S_2 was obtained by resigning from the definition of new attribute a7 - state of consciousness and restoring original attributes 12 and 20 instead of it. The subset S_3 resulted from skipping the definition of attribute of a4 - signs of hiporolenic shock and coming back to original attributes 5 and 6. Then, we checked systematically if we could stay with three new attributes only (instead of four). We discovered such a possibility - i.e. resigning from introducing two new attributes a4 and a7. This was denoted as a subset S_4. As one can notice from Table 2, this subset S_4 gave good evaluation of induced decision rules while still giving satisfactory number of reducts. The core was the same for all subsets and contained the following attributes chest injuries, fracture of long bones, time from the admission to the execution of DPL and ensured the quality of classification equal to 0.37.

Table 2. Results of analysis for different combinations of newly constructed attributes. Numbers in brackets refer to decision classes in both classifications.

subset of attributes	quality of classification	no. of attr. in the core	number of reducts	number of rules	average rule strength	classification accuracy
S_1	1.00	3	11	22 (13/9)	5.14(7.54/1.67)	76.25%
S_2	1.00	3	26	23 (13/10)	4.93(7.62/1.5)	75.00%
S_3	1.00	3	21	21 (13/8)	5.76(8.92/1.8)	72.50%
S_4	1.00	3	39	18 (11/7)	7.00(10.09/2.14)	77.50%

The sets of discovered rules characterized in Table 1 and Table 2 could be compared taking into account different criteria. Considering also medical point of view, we were mainly focused on: the number of rules, the rule strength, correspondence of the condition parts to real clinical experience. This led us to choosing the subset S_4. The set of 18 decision rules was characterized by the best properties and according to medical experts these rules were the closest to the clinical experience. Although we did not put main stress on the classification accuracy, it should be notices the subset S_4 gave the highest one. In particular, for decision class 'died' we observed the increase of classification accuracy in comparison to preliminary set of 22 attributes. Some doubts can be stated to obtaining 'weak' rules supported by few examples. However, such rules could have some medical importance for diagnosing aims as they focus practitioners' attention on specific or difficult cases. Detailed consideration of such cases also enriches clinical experience.

The information system based on the subset S_4 was chosen to the more detailed analysis. As application of the rough set theory gave 39 reducts, we decided to evaluate the significance of particular attributes by using the strategy of adding the most discriminatory attributes to the core. The following subset of attributes was found: a3 - cause of accident, a4 - pulse, a6 - skin, a8 - state of consciousness, a9 - chest injuries, a10 - fracture of long bones, a13 - time to DPL, a14 - DPL result. We can notice that many of these attributes were also indicated as significant ones in statistical tests performed in section 3.2.

4 Concluding remarks

The approach for discovery of attribute dependencies in data sets which integrates rough set theory, statistical and induction techniques has been considered in this paper. The goal of this integration is to simplify the input data by finding the definition of condition attributes relevant to the objects' classification and as a result improving the discovered decision rules.

In this paper, the proposed approach was applied to analyse the medical data sets containing patients with multiple injuries. These data could be treated as 'difficult' for a simple use of standard rough set approach (too many reducts, difficulties in identifying the most significant attributes). Using the Chi-square test and analysing conditions of rules discovered by LEM2 algorithm helped in determining a group of the most significant attributes and to rejecting some irrelevant ones. However, the number of reducts was still quite high. Then, the stage of constructing new attributes resulted in finding such attributes that improved the finally discovered knowledge. Let us notice (see Table 1 and 2) that the number of reducts decreased from 1295 to 38, and the number of discovered rules decreased to 18 more general ones. They were also more consistent with clinical experience.

The clinical usefulness of obtained results, discovered significant attributes and suggestions coming from the analysis of decision rules should be still discussed and investigated. However, it is possible that these experiences could be used to construct the medical support procedure which should be verified with treatment of new multiple injured patients admitted to emergency units of the hospital.

References

1. L.B. Bone, P. Stegemann, Babikian G. Management of the multiple injuries patient with fracture. Instr. Course Lect, **44** (1995) 477–485.
2. E. Bloedorn, R.S. Michalski. Data-driven constructive induction in AQ17-PRE: a method and experiments. In Proc. of the Third Conference on Tools for AI, San Jose. CA (1991) 25–35.
3. C.C. Chan, J.W. Grzymala-Busse. On the two local inductive algorithms: PRISM and LEM2. Foundations of Computing and Decision Sciences, vol. 19/4, (1994) 185–204.
4. J.W. Grzymala-Busse. LERS - a system for learning from examples based on rough sets. In R. Słowiński, (ed.) Intelligent Decision Support, Kluwer Academic Publishers (1992) 3–18.
5. A. Hirshberg, Kl. Mattox. Planned reoperation for severe trauma. Annals of Surgery vol. 222 no. 1 (1995) 3–8.
6. I.F.Imam, R.S. Michalski, L.Kerschberg. Discovering attribute dependence in databases by integrating symbolic learning and statistical analysis techniques. In G. Piatetsky-Shappiro (ed.), Proc. of Knowledge Discovery in Databases Workshop, AAAI-93, (1993) 264–275.
7. Z. Pawlak. Rough sets. Theoretical aspects of reasoning about data. Kluwer Academic Publishers, Dordrecht, 1991.
8. Z. Pawlak, K. Slowinski, R. Slowinski. Rough classification of patients after highly selective vagotomy for duodenal ulcer. Int. Journal of Man-Machine Studies, **24** (1988) 413–433.
9. K. Slowinski, R. Slowinski, J. Stefanowski. Rough sets approach to analysis of data from peritoneal lavage in acute pancreatitis. Medical Informatics **13** (1988) 143–159.
10. R. Słowiński, J.Stefanowski. Rough classification with valued closeness relation. In Didey E. et al. (eds.), New Approaches in Classification and Data Analysis, Springer Verlag, Studies in Classification, Data Analysis and Knowledge Organization, (1993) 482–489.
11. J. Stefanowski, K. Slowinski. Rough sets s a tool for studying attribute dependencies in the urinary stones treatment data set. In T.Y. Lin, N. Cecrone (ed.), Rough Sets and Data Mining, Kluwer Academic Publishers (1997) 177–198.
12. J. Stefanowski, and D. Vanderpooten. A general two stage approach to rule induction from examples. In W. Ziarko (ed.) Rough Sets, Fuzzy Sets and Knowledge Discovery, Springer-Verlag, (1994) 317–325.
13. W. Ziarko and N. Shan. KDD-R: A comprehensive system for knowledge discovery in databases using rough sets. In T.Y. Lin, A.M. Wildberg (eds.) Soft Computing, November 1994, San Jose, CA., Simulation Council Inc. Press, San Diego, (1994) 93–96.
14. W. Ziarko. Review of basics of rough sets in the context of data mining. In S.Tsumoto (ed.) Proceedings of the Fourth International Workshop on Rough Sets, Fuzzy Sets and Machine Discovery, Tokyo Nov. 6-8 1996, 447–457.