

FARM: Fuzzy Action Rule Mining

Zahra Entekhabi

Department of Computer Engineering
Fars Science and Research Branch, Islamic Azad University
Marvdasht, Iran

Pirooz Shamsinejadbabki

Department of Computer Engineering and Information
Technology
Shiraz University of Technology
Shiraz, Iran

Abstract—Action Mining is a sub-field of Data Mining that concerns about finding ready-to-apply action rules. The majority of the patterns discovered by traditional data mining methods require analysis and further work by domain experts to be applicable in target domain while Action Mining methods try to find final cost-effective actions that can be applied immediately in target domain. Current state-of-the-art methods in AM domain only consider discrete attributes for action rule mining. Therefore, one should discretize continuous attributes using traditional discretization methods before using them for action rule mining. In this paper, the concept of Fuzzy Action Rule has been introduced. In this type of action rule, continuous attributes can be presented in fuzzy form. So that they can suggest fuzzy changes for continuous attributes instead of discretizing them. Because the space of all fuzzy action rules can be so huge a Genetic Algorithm-based Fuzzy Action Rule Mining (GA-FARM) method has been devised for finding the most cost-effective fuzzy action rules with tractable complexity. The proposed method has been implemented and tested on different real datasets. Results confirm that the proposed method is successful in finding cost-effective fuzzy action rules in acceptable time.

Keywords—*Action mining; fuzzy action rule mining; genetic algorithm*

I. INTRODUCTION

Nowadays, the increase in computer usage has led to the sharp growth of databases and aggregation of data in the majority of organizations. In recent years, tendency to discovering actionable knowledge from data with the aim of decision making has seen a dramatic increase. Due to the large size of the data, human experts cannot investigate and analyze all data; on the other hand, the capability of fast and effective response to customer needs requires fast enough decisions which cannot be made without accurate evaluation and analysis of information and data. Therefore, discovering knowledge from database in almost real-time is of paramount importance. Data mining is a procedure which uses data analysis tools to find patterns and relationships between data. Data mining methods focus on discovering and extracting descriptive patterns and do not take into account the actionability of these patterns in target domain. For instance, consider the customer loyalty detection system. The conventional data mining systems predict which customer will leave the company in the close future; however, the companies need to know how they can prevent this and the subsequent losses. In fact, companies need profitable actions instead of just frequent patterns.

Action mining is among the proposed feasible solutions which have attracted attention recently. Actions determine for the user what should be done in order to reach high profits. Action is a type of knowledge which aims to maximize profits in desired areas by suggesting a number of alterations to a sample's condition. The methods proposed for action mining can be categorized into transductive and inductive [1]. In inductive methods, such as what was done by Ras et al. [2]-[9], the obtained action rules are general rules which can be applied to a group of things with similar attributes. Ras method, which has been introduced as DEAR [2], initially extracts the conventional classification rules in data mining and later, by combining compatible classification rules, creates action rules (rules that suggest alterations in attributes).

Tsay et al. [10] improved the DEAR algorithm by DEAR2 system which initially classifies the rule table based on the number of existing decision attribute values and consequently creating a sub-tree based on stable attributes. At the end, leaves are compared with each other to extract action rules.

Tsay et al. [11] also proposed DEAR3 algorithm for incomplete databases which can reduce the errors of undetermined, unreliable and outlier data to obtain more reliable action rules.

Another method which was suggested by Ras to extract action rules is called action rule discovery (ARD) [3]. This method extracts action rules without combining classification rules. The advantage of this method is its lower time complexity that makes it proper for large databases.

Meat-actions are actions that should be triggered to activate another action rule. In [12] Meta-actions have been used for mining business action rules and in [13] surgical meta-actions have been mined for medical diagnosis.

Inductive methods in other hand try to find most profitable actions for each case. The most prominent method of this type is presented by Yang et al. [14]-[15] which uses a traditional decision tree to extract actions. In summary, at first, a decision tree is inferred from the data. Next, for each new sample, it searches a leaf node and computes the net profit made from the movement of the new sample to the other leaves. Finally, the leaf that has the highest profit is selected as the destination leaf and the necessary alterations are performed for transferring the sample from the current node to the destination node. In order to extract action rules, Shamsinejadbabki et al. [1] proposed a new action mining method by using causal networks. This

method can obtain practical actions in the real world by considering causal relationships between the attributes.

Traditional decision making trees (used by Yang) in [15] require data with discrete values and the continuous values need to be discretized before being put in the algorithm. Despite conventional in data mining, the continuous to discretized domain transformation is little compatible with the real world and it is therefore difficult to define a precise border for continuous attributes' values. For instance, consider the question of how the border between middle age and old can be determined. If it is 60 years, a person who is 59 years old is considered middle-age and a 60-year-old person is considered old. Obviously this classification is not compatible with the real world actualities. Using fuzzy logic in traditional fuzzy logic capabilities makes it possible to work with continuous data. Kalanat et al. [16] use fuzzy set theory to improve Yang method when we face continuous attributes in data. They have shown that although the final rules are not fuzzy but their actionability is higher than Yang rules.

In this study, the concept of Fuzzy Action Rule (FAR) has been introduced. This type of action suggests fuzzy alterations for continuous attributes and therefore, it is expected that the resulting actions will be more applicable in the real world circumstances. Besides, a method has been devised for derivation of fuzzy action sets with maximum net profit from data using Genetic Algorithm.

The rest of paper is as follows: In Section 2 preliminaries are explained. Fuzzy Action Rule will be introduced in Section 3 and GA_FARM, a GA-based FAR mining method will be presented in Section 4. Experimental results are argued in Section 5 and paper will be concluded in Section 6.

II. PRELIMINARIES

Set and element are the two primary concepts of set theory. In classic set theory (crisp), x either can or cannot be a member of set S. As a result, the membership function of element x to set S (μ_S) can be shown as:

$$\mu_S(x) = \begin{cases} 1 & \text{if } x \in S, \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

However, a large number of the real world sets are inherently fuzzy sets. Fuzzy set theory is a mathematical structure which allows an element to belong to more than one set. The membership degree of element x in fuzzy set S as μ_S can be shown as $x \rightarrow \{0,1\}$ where $0 \leq \mu_S \leq 1$ [17].

Suppose three fuzzy sets (young, middle age, old) which are attributed to age variable. Now if in a classic set, 60 is considered as the border between middle-aged and old people, this is tantamount to considering a 59 years-old person as a middle age person and a 60 years-old person as old. However, fuzzy theory puts narrow borders between the values so that an element can be classified with various degrees. For instance, a 55 years-old person can belong to the middle-age and old sets at the same time with different membership degrees.

A. Fuzzy Decision Tree

A fuzzy decision tree is a combination of fuzzy theory with traditional decision making trees. The generality in learning

from samples, clarity of the knowledge obtained from the decision tree and also the ability of working with fuzzy imprecise and incomplete data are among the benefits of using this combination. It can also increase robustness and applicability in inaccurate and unclear areas. The knowledge obtained from FDT more closely resembles human decision making and therefore, is more practical in real world circumstances. For example, in [18] FDT has been used for software cost estimation to handle imprecise data in describing software.

Fig. 1 depicts an example of FDT from [17] where attribute x_1 is the root and attribute x_2 is the internal node. The v^{th} membership function can be determined from the j^{th} attribute by μ_{jv} . The K^{th} probable class in the L^{th} leaf is determined with β_{KL} . The number of shown classes in tree is Equal to 2. As a result, each leaf includes two fuzzy variables: β_{1L} and β_{2L}

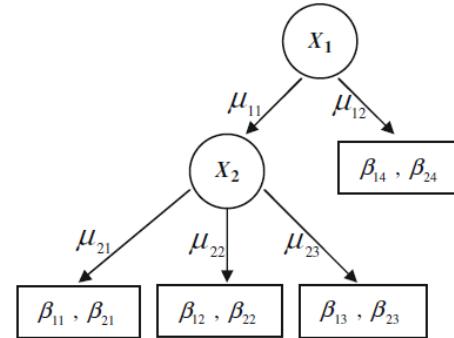


Fig. 1. A sample fuzzy decision tree [17].

B. Action, Action Set and Action Rule

An action has a $(A, a \rightarrow a')$ structure in which A is an attribute except the goal (A is called the action attribute) and $a, a' \in \text{Dom}(A)$. This action indicates change of A from value a to a' by a force outside of the system [1]. Action set is a set of actions. For example, (Service, low \rightarrow high; Bill, notPayed \rightarrow payed) is an action set. Action Rule is type of rule that its antecedent is an action set and its consequent is a change in goal attribute. It is shown an action rule below:

(Service, low \rightarrow high; Bill, notPayed \rightarrow payed) \Rightarrow (loyalty, no \rightarrow yes)

III. FUZZY ACTION RULE

In this work the fuzzy action concept is proposed. In fuzzy action, unlike the conventional actions, fuzzy alterations are suggested. For instance, fuzzy action suggests the reduction in smoking rather than cessation for pulmonary disease patients. The same is true with fuzzy action and soft drinks for patients. The motivation beyond this is that if fuzzy changes is considered it is possible to find more cost-profitable actions. Suppose attribute A has a_i value in a sample. Then, fuzzy action on A is defined as:

$$\varphi_A = (A, a_i \rightarrow a_i + k\Delta t) \quad (2)$$

Where Δt is the change granularity of attribute A and can be computed using the following formula:

$$\Delta t = \frac{A_{\max} - A_{\min}}{n} \quad (3)$$

A_{\max} : maximum value of attribute A in dataset.

A_{\min} : minimum value of attribute A in dataset.

n : determines to what extend Δt will be fine-grained or coarse-grained. Larger n results in smaller Δt and therefore more fine-grained fuzzy action sets.

k : is change magnitude and can be positive, negative or zero.

If k is positive, the value of A increases and if it is negative, it decreases. It is obvious that if k is considered zero, there is no change in attribute value.

The following example can clarify the concept of fuzzy action:

Consider three attributes X_1 , X_2 and X_3 with the values (12, 3, 5). Suppose that Δt and K are 0.6 and -2 for X_1 , respectively. By the fuzzy action, $X_3, 5 \rightarrow 5 + (-2 \times 0.6)$, X_3 will change from 5 to 3.8. A set of fuzzy actions suggest alterations to multiple attributes and we name it Fuzzy Action Set. For example:

$$\begin{bmatrix} (X_1, 12 \rightarrow 12 + (0 \times 0.8)) \\ \wedge (X_2, 3 \rightarrow 3 + (5 \times 0.4)) \\ \wedge (X_3, 5 \rightarrow 5 + (-2 \times 0.6)) \end{bmatrix}$$

Since Δt is constant for each attribute, an action set can be written as:

$$[(X_1, 12, 0) \wedge (X_2, 3, 5) \wedge (X_3, 5, -2)]$$

The above action set suggests that value of attribute X_1 does not change, value of attribute X_3 with an initial value of 5 decreases by $2\Delta t_{x3}$ and value of attribute X_2 with an initial value of 3 increases by $5\Delta t_{x2}$.

Fuzzy Action Rule is type of action rule that its antecedent is a fuzzy action set. E.g. a Fuzzy Action Rule can be like this:

$$[(X_1, 12, 0) \wedge (X_2, 3, 5) \wedge (X_3, 5, -2)] \Rightarrow (G, n \rightarrow y)$$

IV. GA-FARM

The major problem in finding fuzzy action sets is the search space. Even in a dataset with a few attributes and small domain ranges of attributes the space of all possible fuzzy action sets is huge. So that exhaustive searching of this space is intractable. As a meta heuristic solution genetic algorithm has been used that is quite strong in solving such NP-Hard problems. We are looking for such rules that net profit, i.e. profit minus cost, will be maximum by applying them. At the following, different parts of proposed GA algorithm will be described.

A. Chromosome Structure

In GA, each solution is shown as a chromosome and each chromosome contains multiple genes. In our work each chromosome represents a Fuzzy Action Set (FAS) in which each gene is a Fuzzy Action. So that number of genes indicates

number of attributes in FAS and because number of actions in FAS is various chromosomes' length is also variable.

Each gene has three components: attribute name A, current attribute value V and change magnitude k. So that a gene with structure (A, V, k) represents fuzzy action $(A, V \rightarrow V + k\Delta t)$.

B. Fitness Functions

The most important part of a GA is its fitness function. It is defined as:

$$\text{Fitness(ch)} = \sum_{i=1}^m \text{profit}(i) \times (\text{md}_{\text{after}}(\text{ch}, i) - \text{md}_{\text{before}}(\text{ch}, i)) - \text{cost}(\text{ch}) \quad (4)$$

Where m is the number of values of goal attribute, $\text{profit}(i)$ is the profit of being in class i that will be provided by domain experts. $\text{md}_{\text{before}}(\text{ch}, i)$ and $\text{md}_{\text{after}}(\text{ch}, i)$ indicate the membership degree of the instance to class i before and after applying the rule, respectively. For example, consider the following chromosome:

$$\text{ch: } [(A, 1, 3), (B, 2, 0), (C, 3, -1)]$$

Also assume $\Delta t_A = 2$, $\Delta t_B = 3$, $\Delta t_C = 1$ and goal attribute G has two values G_1 and G_2 . $\text{md}_{\text{before}}(\text{ch}, G_i)$ is the membership degree of instance ($A = 1$, $B = 2$, $C = 3$) to class G_1 and $\text{md}_{\text{after}}(\text{ch}, G_i)$ is the membership degree of instance ($A = 1 + 3 * 2$, $B = 2 + 0 * 3$, $C = 3 + (-1) * 1$) to class G_1 . Here fuzzy decision tree will be used for computing these values.

Cost will be computed using the following formula:

$$\text{cost} = \sum_{i=1}^n (k \times \text{cost}_i(\Delta t)) \quad (5)$$

Where $\text{cost}_i(\Delta t)$ is the cost of changing attribute i with value Δt . Cost values will be fed into model by domain experts.

C. Crossover

Crossover is responsible for exploiting search space in GA. It tries to combine two good chromosomes with the hope of generating more fitted offspring. Here a one-point crossover operation has been designed for combining the most fitted chromosomes probabilistically.

For each pair of chromosomes, which are selected using roulette wheel selection strategy, two offspring will be generated by exchanging their genes around crossover point which is determined probabilistically. Designed crossover has been depicted in Fig. 2 for two sample chromosomes.

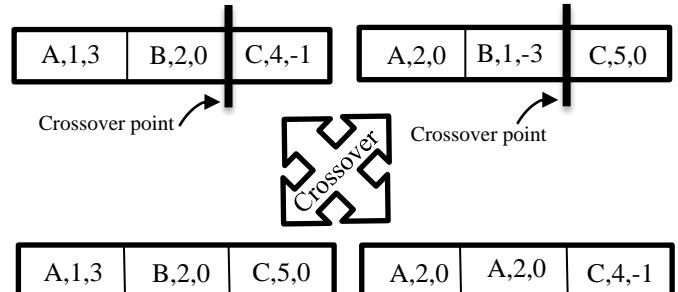


Fig. 2. One-point crossover operation used in GA-FARM.

D. Mutation

Mutation is the core of exploration mechanism of GA. It helps GA exit from local optimums. Mutation operator tries to mutate some genes of a chromosome with the hope of generating a new chromosome from other parts of the search space.

In GA-FARM a predefined percentage of chromosomes are muted by changing value k of some of their genes by a random value. In Fig. 3, the result of mutation on a sample chromosome has been depicted.

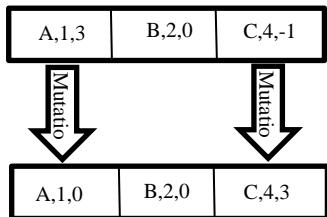


Fig. 3. Mutation operation used in GA-FARM.

V. EXPERIMENTAL RESULTS

For experimental results GA- FARM has been tested on three different datasets: Bank Loan dataset which has gathered from one of the most prominent banks in IRAN (dataset description has been shown in Table II), Pima Indians Diabetes [19] and Turkey Student Evaluation [20] both from UCI Machine Learning Repository. The summary of these datasets can be seen in Table I.

One of the major problems in action mining domain is evaluating the mined actions because it is necessary to apply them in real world. To investigate the effect of proposed method we have designed two research questions: How logical action rules will be mined using FARM and whether GA can converge in acceptable time or not. The former question is subjective while the latter is objective.

Fig. 4 shows the results of applying the GA-FARM on bank data sets. It shows the convergence of GA for this dataset. As it can be seen the GA is successful in improving the overall fitness of population. Elapsed time for this dataset is 281.31 seconds which confirms that GA-FARM has acceptable time complexity.

TABLE I. DATASET DETAILS

Dataset	No. of Instances	No. of Attributes	No. of Continuous Attributes	No. of class attributes
Bank Loan	199	10	7	3
Pima Indians Diabetes	768	8	7	2
Turkey Student Evaluation	5820	33	0	5

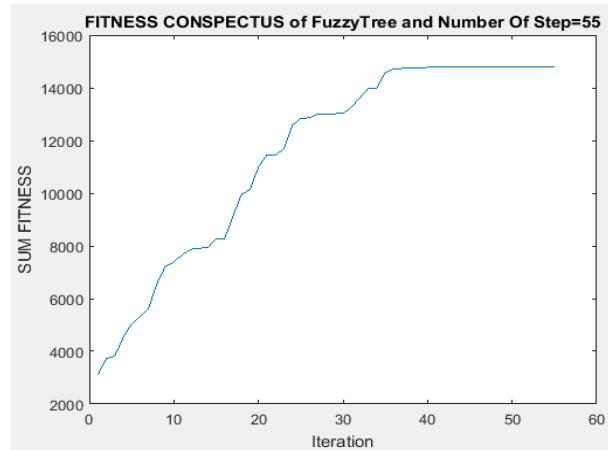


Fig. 4. Results of applying the GA-FARM on bank loan dataset.

One of the mined rules using GA-FARM from bank loan dataset has been depicted below:

$$\left[\begin{array}{l} (X_1,0,0) \wedge (X_2,36,0) \wedge (X_3,0,-2) \\ \wedge (X_4,1,1) \wedge (X_5,2,-1) \wedge (X_6,0,3) \\ \wedge (X_7,0,3) \wedge (X_8,770000,0) \wedge \\ (X_9,770000,0) \wedge (X_{10},32,-2) \end{array} \right] \Rightarrow [\text{Loyal, No} \rightarrow \text{Yes}]$$

This rule shows alterations of the values of two attributes: marital status and the number of deferred payments. There is no change in the attributes where k is zero. For instance, for attribute #10 which is the number of deferred installments, -2 is suggested for k and 0.5 is computed for Δt . Based on this rule, the customer must reduce his or her deferred paid installments by -2×0.5 , i.e. from 32 to 31. By applying this rule, the bank may reach a net profit of 190. This action rule shows that it is possible to achieve good profit but with less changes in attributes. This is the motivation behind fuzzy action rule.

TABLE II. BANK LOAN DATASET ATTRIBUTES

X	Name of Attribute
X1	sex
X2	Age
X3	Marital Status
X4	Degrees of Education
X5	Job
X6	Number of bank transactions during last 3 months
X7	Number of Electronic Services used during last 30 transactions
X8	6-month average balance
X9	12-month average balance
X10	Number of deferred instalments

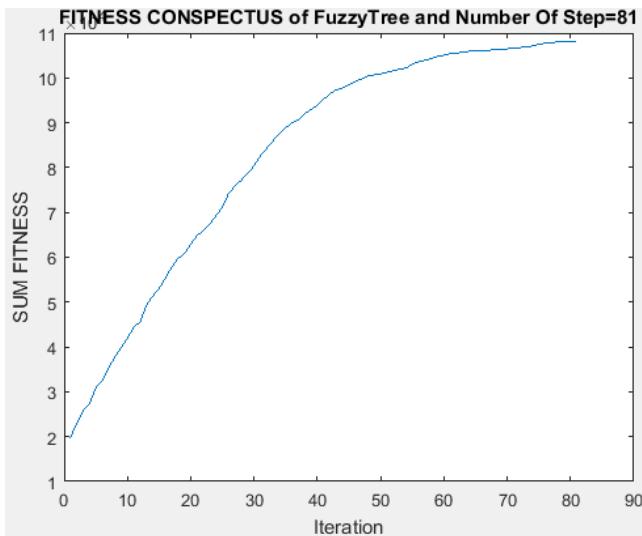


Fig. 5. Results of applying the GA-FARM on Pima Indians Diabetes dataset.

By applying proposed method on Pima Indian Diabetes dataset it is succeeded to find the most profitable fuzzy action Rules. Fig. 5 shows the convergence of GA-FARM on this dataset. One of the mined action rules is:

$$\left[\begin{array}{l} (X_1,4,0) \wedge (X_2,129,0) \wedge (X_3,60,-3) \\ \wedge (X_4,12,-9) \wedge (X_5,231,0) \\ \wedge (X_6,27.5000,-7) \wedge (X_7,0.5270,-1) \\ \wedge (X_8,31,0) \end{array} \right] \Rightarrow [\text{Diabetes, Yes} \rightarrow \text{No}]$$

This rule suggests that the patient can reduce the probability of getting diabetes by losing $7\Delta t(4.697)$ kg weight.

Finally, GA-FARM is applied on Turkey Student Evaluation dataset. The GA convergence graph has been depicted in Fig. 6 which shows it succeeded in finding local optimum fuzzy action sets from this dataset. One of the mined action rules has been depicted below.

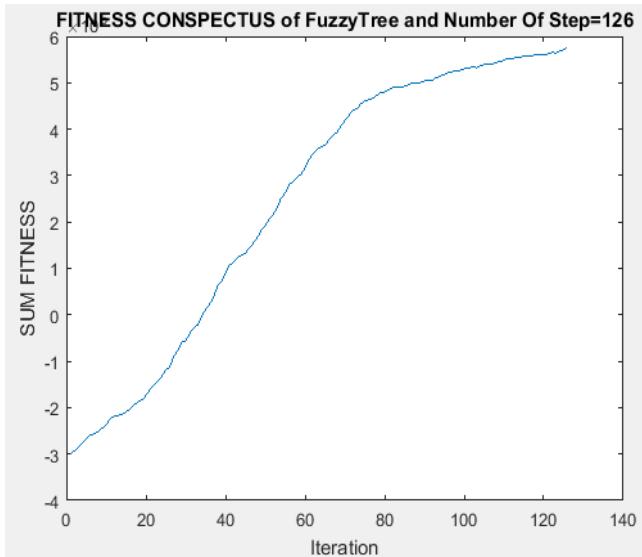


Fig. 6. Results of applying the GA-FARM on Turkey Student Evaluation dataset.

$$\left[\begin{array}{l} (X_1,1,0) \wedge (X_2,1,0) \wedge (X_3,1,0) \\ \wedge (X_4,3,0) \wedge (X_5,3,0) \wedge (X_6,4,0) \\ \wedge (X_7,4,0) \wedge (X_8,4,0) \wedge (X_9,4,0) \\ \wedge (X_{10},4,0) \wedge (X_{11},4,1) \wedge (X_{12},4,0) \\ \wedge (X_{13},4,2) \wedge (X_{14},4,0) \wedge (X_{15},4,2) \\ \wedge (X_{16},4,0) \wedge (X_{17},4,1) \wedge (X_{18},4,0) \\ \wedge (X_{19},4,2) \wedge (X_{20},4,0) \wedge (X_{21},4,0) \\ \wedge (X_{22},4,0) \wedge (X_{23},4,0) \wedge (X_{24},4,3) \\ \wedge (X_{25},4,0) \wedge (X_{26},4,0) \wedge (X_{27},4,0) \\ \wedge (X_{28},4,0) \wedge (X_{29},4,3) \wedge (X_{30},4,0) \\ \wedge (X_{31},4,0) \wedge (X_{32},4,0) \wedge (X_{33},4,0) \end{array} \right] \Rightarrow [\text{raise points, low} \rightarrow \text{high}]$$

Fuzzy action rules mined from datasets show that GA-FARM can extract meaningful rules in corresponding domains. The elapsed time for GA and number of epochs before convergence show that GA operations succeeded.

VI. CONCLUSION AND FUTURE WORKS

Action Rule is a set of change suggestions with the aim of converting an instance from less profitable class to more profitable one. Current state of the art methods in this domain only suggest crisp changes in attribute values. In this paper the concept of Fuzzy Action Rule has been introduced which suggests fuzzy changes. Applying fuzzy changes help us to gain profit with less cost comparing to crisp changes.

A big burden against finding fuzzy action rules is the huge search space. So that, GA-FARM has been devised for extracting fuzzy action rules using GA. The proposed method then has been implemented and tested on different datasets. Results confirmed that GA-FARM can find optimum fuzzy action sets in acceptable time.

The most important limitation in the field of action mining is about evaluating the results because checking the effect of an action needs to apply it in real world. It is necessary to design some frameworks that can evaluate the effect of actions more realistically. For future works we will also try to devise new FARM methods that can find more profitable fuzzy rules. Incorporating fuzzy type-2 into action rules can be another promising research area.

REFERENCES

- [1] P. Shamsinejadbabaki, M. Saraee, and H. Blockeel, "Causality-based Cost-effective Action Mining", International Journal of Intelligent Data Analysis, Volume 17 Issue 6, Pages 1075-1091, 2013.
- [2] Z. W. Ra's and L. S. Tsay, "Discovering Extended Action-Rules (System DEAR)", Intelligent Information Processing and Web Mining. Advances in Soft Computing, Springer Berlin Heidelberg, vol. 22, pp. 293-300, 2003.
- [3] Z. W. Ra's and A. Dardzińska, "Action Rules Discovery, a New Simplified Strategy", Foundations of Intelligent Systems. ISMIS 2006. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, Vol. 4203, pp. 445-453, 2006.
- [4] Z. W. Ra's, A. Dardzińska, L. S. Tsay and H. Wasyluk, "Association Action Rules". IEEE International Conference on Data Mining Workshops, vol. 66, pp. 283-290, 2008.
- [5] Z. W. Ra's, L. S. Tsay and A. Dardzińska, "Mining E-Action Rules", Journal of Mining Complex Data, Springer-Verlag Berlin Heidelberg , 2008.
- [6] Z. W. Ra's, and A. Wieczorkowska, "Action-Rules: How to increase profit of a company", European Conference on Principles of Data

- Mining and Knowledge Discovery, Springer Berlin Heidelberg ,Vol. 1910, pp. 587–592, 2000.
- [7] Z. W. Ra's, and A. Dardzinska, "From Data to Classification Rules and Actions", International Journal of Intelligent Systems, Vol. 26(6), pp. 572–590, 2011.
- [8] Z. W. Ra's, A. Tzacheva, L. S. Tsay and O. Gurdal, "Mining for interesting action rules", international conference on intelligent agent technology, IEEE Computer Society, 187–193, 2005.
- [9] Z. W. Ra's, A. Dardzińska, L.S. Tsay, L.S and H Wasyluk, "Association action rules". In IEEE International conference on data mining workshops, 283–290, 2008.
- [10] L.S. Tsay and Z. W. Ra's, "Action rules discovery: system DEAR2, method and experiments". Journal of Experimental & Theoretical Artificial Intelligence, Vol. 17, No. 1–2, pp. 119–128, 2005.
- [11] L.S. Tsay and Z. W. Ra's, "Action Rules Discovery System DEAR3", International Symposium on Methodologies for Intelligent Systems, Springer, pp. 483–492, 2006.
- [12] J. Kuang and Z. W. Ra's, "In Search for Best Meta-Actions to Boost Businesses Revenue", Flexible Query Answering Systems, Springer, 2016.
- [13] H. Touati, Z. W. Ra's, J. Studnicki, and A. A. Wieczorkowska, "Mining Surgical Meta-actions Effects with Variable Diagnoses' Number", International Symposium on Methodologies for Intelligent Systems Springer, pp. 254–263 , 2014.
- [14] Q.Yang and C. Hong, "Mining Case Bases for Action Recommendation", IEEE International Conference on Data Mining, pp. 522 – 529, 2002.
- [15] Q.Yang , Jie Yin , C.X. Ling and T. Chen, "Postprocessing decision trees to extract actionable knowledge". Third IEEE International Conference on Data Mining, pp. 685–688, 2003.
- [16] N. Kalanat, P. Shamsinejad, and M. Saraei, "A Fuzzy Method for Discovering Cost-effective Actions from Data", Journal of Intelligent and Fuzzy Systems, pp.757-765, 2014.
- [17] H. Sattar and Y. Yang, "Flexible Decision Trees for Mining High Speed Data Streams at Presence of Noise and Uncertainty", Data Mining and Knowledge Discovery, vol. 19, pp. 95–131,2009.
- [18] A. Idri and S. Elyassami, "A Fuzzy Decision Tree to Estimate Development Effort for Web Applications" (IJACSA) International Journal of Advanced Computer Science and Applications, 2011.
- [19] G. Gunduz and E. Fokoue, "UCI Machine Learning Repository", Irvine, CA: University of California, School of Information and Computer Science, 2013.
- [20] V. Sigillit, UCI Machine Learning Repository. Irvine, CA: University of California, School of Information and Computer Science, 2013.