https://doi.org/10.26552/com.C.2013.1.68-73

COMMUNICATIONS

Tomasz Kanik *

HEPATITIS B DISEASE DIAGNOSIS USING ROUGH SET

This paper describes processing of the medical data by means of the prediction system based on Rough Set Theory (RST). The Rough Sets proved to be very useful for the analysis of the decision problems concerning objects described in a data table by a set of condition attributes as well as a set of decision attributes. In order to make efficient data analysis and suggestive predictions in a case of the data of patients suffering from viral hepatitis were used to predict a probability of their death or serious disability. This paper also demonstrates an extension of the Rough Set methodology for reducing number of input data in order to increase prediction accuracy without loss of knowledge.

Keywords: *Index Terms – Rough Set, hepatitis, machine learning, prediction system.*

1. Introduction

The Rough Sets and their theory have been developed as a way of dealing with incomplete sets of information in the early eighties by Zdzislaw Pawlak. The Rough Set Theory has led to many interesting applications and extensions. The theory is in a wide spread used in a scientific world and now it is one of the fastest growing methods of artificial intelligence. As the author of the theory stated [1] it seems that the Rough Set approach is fundamentally important in artificial intelligence and machine learning, especially in research areas such as pattern recognition, cognitive sciences, mereology, decision analysis, intelligent systems, expert systems, inductive reasoning and knowledge discovery [2].

The *Rough Sets*, as the name suggests, are the sets defined on the discrete split place. The space is discretised by the definition of the elementary set and its size depends on the level of space approximation. The items in the elementary set have interesting features; they are indiscernible among themselves and each of them has all characteristic properties typical for the whole set. The membership function takes a set of values corresponding to the number of groups to which the item is added: 1 – if the element belongs to class 1, 2– if the item belongs to class 2, and so on; the value 0 is assigned to those items which are not classified, that is for those ones we cannot determine the group they belong to.

Basic operations on the Rough Set [1] are the same as the operations on classical sets, for example:

The *information system* is defined as $I = (U, A)$, where U is a finite, non-empty set of objects called *a universum* and A is a finite, non-empty set of attributes such that $\forall a \in A : U \rightarrow V_a$. V_a is the set of values that attribute a may take. The information table assigns a value $a(x)$ from V_a to each attribute a and object x in the universum U.

The *indiscernibility* relationship of the x and y is written in the form (x is in indiscernibility relation to y in the set of B-attributes), which means the elements x and y have the same values of attributes in B. In other words, owing to the set of attributes in B, the elements x and y cannot be distinguished between each other.

For each sub-set of features $B \subseteq A$ there is association with an indiscernibility relation:

$$\mathrm{I}ND(B) = \{(x,y) \in U^2 \mid \forall a \in \mathrm{B}, a(x) = a(y)\}$$

Lower approximation $\underline{B(X)}$ is the complete set of objects in U which can be certainly classified as the elements in X by using the set of attributes B. It is the largest subset of B contained in X.

$$\underline{B}(X) = \{x \in U \mid [x]_{IND(B)} \subseteq X\}$$

Upper approximation $\overline{B(X)}$ is the set of elements in S that can be possibly classified as the elements in X.

$$\bar{B}(X) = \{x \in U \mid [x]_{IND(B)} \cap X\}$$

The B-*boundary* of X in the information system I, is defined as:

$$BND(X) = \bar{B}(X) - \underline{B(X)}$$

The most important properties of the Rough Set are shown in Fig. 1.

The *reduct* presents minimum attributes subset that keeps the degree of dependencies attributes to the conditional attributes. It is subset $R \subseteq B \subseteq A$ such that $X_B(X) = X_R(X)$ and is noted by $Red_X(B)$.

* **Tomasz Kanik**
Department of Mathematical Methods, Faculty of Management Science and Informatics, University of Zilina, Slovakia,
E-mail: Tomasz.Kanik@fri.uniza.sk

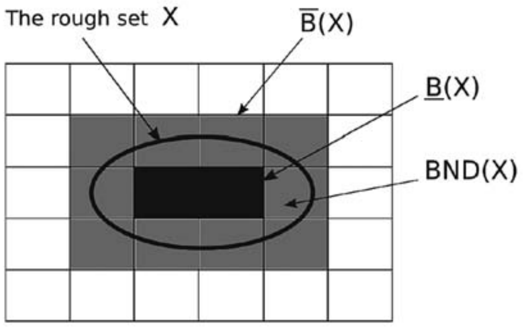

Fig. 1 A graphical representation of a Rough Set environment

The intersection of all reducts is called a core. It cannot be removed from the information system without deteriorating the basic knowledge of the system. Thus, none of its elements can be removed without affecting the classification power of attributes. The set of all indispensable attributes of B is called the X-core. Formally,

$$Core_X(B) = \cap\ Red_X(B)$$

The parameter characterizing the Rough Set numerically is *the accuracy of the approximation* and it measures how much the set is rough. If a set has $\underline{B}(X) = \overline{B}(X) = X$, the set is precisely called *the crisp* and for its every element the relationship: $x \in X \in U$ is valid. It is represented by the formula:

$$\mu_B(X) = \frac{Card\left|\underline{B(X)}\right|}{Card\left|\underline{B(X)}\right|}$$

where $Card|X|$ denotes the cardinality of $X \notin \emptyset$.

When $0 \le \mu_B \le 1$, and if $\mu_B = 1$ then X is a crisp in respect to B.

Additionally, several new concepts were introduced by Ziarko and Shan [3], [4]. They distinguish in an information system two disjoint classes of attributes, called *condition* and *decision attributes*. An information system is then called the *decision table*, respectively. The decision table is denoted by $S = (U, C, D)$, where C and D are disjoint sets of condition and decision attributes.

Every $x \in U$ determines a *sequence* $c_1(x), \ldots, c_n(x), d_1(x), \ldots, d_m(x)$, where $\{c_1, \ldots, c_n\} = C$ and $\{d_1, \ldots, d_n\} = D$. The sequence is called a *decision rule induced by* x (in S) and is denoted by $c_1(x), \ldots, c_n(x) \rightarrow d_1(x), \ldots, d_m(x)$ or in short $C \rightarrow {}_x D$.

The number $supp_x(C, D) = |A(x)| = |C(x) \cap D(x)|$ is called *support* of the decision rule $C \rightarrow {}_x D$ and the number

$$\sigma_x(C,D) = \frac{supp_x(C,D)}{|U|}$$

is referred to as the *strength* of the decision rule $C \rightarrow {}_x D$, where $|X|$ denotes the cardinality of X.

The *certainty factor* of the decision rule is denoted $cer_x(C, D)$ and defined as follows:

$$cet_x(C,D) = \frac{|C(x) \cap D(x)|}{|C(x)|} = \frac{supp_x(C,D)}{|C(x)|} = $$
$$= \frac{\sigma_x(C,D)}{\pi_x(C(x))}$$

where $\pi_x(C(x)) = \frac{|C(x)|}{|U|}$. The certainty factor may be interpreted as a conditional probability that y belongs to $D(x)$ given y belongs to $C(x)$, symbolically $\pi_x(C|D)$. If $cer_x(C, D) = 1$, then $C \rightarrow {}_x D$ is called a *certain decision rule*; if $0 < cer_x(C, D) < 1$ the decision rule is refered to as an *uncertain decision rule*.

The *coverage factor* of decision rule is denoted $cov_x(C, D)$ and defined as follows:

$$\text{cov}_x(C,D) = \frac{|C(x) \cap D(x)|}{|D(x)|} = \frac{supp_x(C,D)}{|D(x)|} = $$
$$= \frac{\sigma_x(C,D)}{\pi_x(D(x))}$$

where $\pi_x(D(x)) = \frac{|D(x)|}{|U|}$. Similarly $cov_x(C, D) = \pi_x(C|D)$.

The *inverse decision rule* is denoted $D \rightarrow {}_x C$ and it is inversion of decision rule $C \rightarrow {}_x D$. It can be used to give *explanation (reason)* for a decision.

2. Data Pre-processing

The attribute reduction is very important in the rough set-based data analysis – according to Smolinski [5], [6] it improves the efficiency of the predictor itself and cuts down the time needed for the future data processing.

A. Attributes filtering

The way to improve the predictor is to select the most important attributes. Following Aboul ella Hassanien [7], in common practices a domain expert's opinion is required to set the data importance, but sometimes the problem is too complicated for a single expert. In that case, according to Slezak et al. [8], a *filter* (selecting algorithm) is needed to check each attribute significances and influences on the whole result, and then to create a new set of attributes as a linear combination of the weighted sum of selected ones. The original aim of the presented method is to create another attributes to adjust the classification result. Here I suggest using my own algorithm, which filters the attribute set and leaves only those attributes which possess the weight over a specified level – similar to

Wroblewski's Classification Algorithms [9]. The algorithm repeatedly check accuracy and coverage rules by the calculation of the different cuts result. The rules are created by *LEM2* algorithm which proved to have the best result in the experimental data. As stated by Polkowsky [10], the *exhaustive selection algorithm* test shows significantly lower accuracy of the created rules and therefore was omitted in further experiments.

B. Attributes reduction

Attributes reduction is done in 2-steps. The first one creates a reconstruction of a decision table. The second one computes the optimal reduct for data analysis. Thus the knowledge is exquisite by continuous dataset discretisation. Discretisation of the dataset is the process of reducing the domain of a continuous attribute with an irreducible and optimal set of cuts, while preserving the consistency of the dataset classification. The basic idea of the *Quick Reduct Algorithm* (QRA) is based on the fact that the discernibility matrix (table) *DM* and the reduct *Red* cannot be empty for any items intersection. The object of matrix *i* and *j* would be indiscernible to the reduct, if there are any empty intersections between items c_{ij} with reduct, this contradicts the definition that reduct is the minimal attribute set discerning all objects. As X.Hu notices [11], the frequency of attribute is used as heuristic and makes it applicable to the optimal rule generation. A QRA reduct set $Red = \phi$, then sort the discernibility matrix $|c_{ij}|$ and examine every items of discernibility matrix $|c_{ij}|$. In case that their intersection is empty the shorter and frequent attribute is picked and inserted in *Red*, otherwise the entry is skipping. A shorter and frequent attributes contribute more classification power to the reduct. If there is only one element in $|c_{ij}|$, it must be a member of reduct. The procedure is repeated until all entries of discernibility matrix are examined. Finally, QRA get the optimal reduct in *Red*. According to Thangavel et al. [12], the discretisation improves classification of unseen objects. The algorithm used for a data reduct is presented below.

The input is $I = (U, B \cup \{d\})$, $B = \cup b_i$, $i = 1 \dots n$. The $count(b_i)$ sums up frequency of the attribute computing by $f(b_i)$, *DM* is decision matrix, $|c|$ is cardinality of *c*, *d* is the decision. The output is the optimal reduct *Red*.

```
$Red = \phi$, $count(b_i) = 0$,   $i = 1 \dots n$;
$DM = \cup c_{ij}$,   $i,j = 1 \dots n$;
Count frequency
for $\forall b_i$ in $DM$ do {
    $f(b_i) = f(b_i) + \frac{n}{|c|}$
}
Merge and sort $DM$
for $\forall c_{ij}$ in $DM$ do {
    if ($c_{ij} \cap Red$ equal $\phi$) then {
        $Red = Red \cup Max(\{f(b_i)\})$
    }
}
Return $Red$
```

C. Attributes decomposition

Attributes decomposition is a process of discretisation of numerical attributes or grouping (quantisation) of nominal ones. The decomposition algorithm indicates how to divide (or join) attribute values. After division of all attributes domain the new decision rules are created. The new rule set should cover the most of cases now.

3. Experiment

A. Data

In order to present the proposed method of data processing, let me consider an example of UCI repository [13] – dataset of patients records. It was donated by Josef Stefan Institute in Ljubljana. Hepatitis (in Greek) means 'liver' and the suffix -itis denotes 'inflammation' of the liver and may be due to infectious or non-infectious causes. As stated by Worman [14], the five types of hepatitis viruses are common infectious causes of the liver inflammation, and some of them such as hepatitis A (*HAV*), B (*HBV*) and C (*HCV*) are more frequently seen as the infectious agents. The inflammation may lead to death of the liver cells (hepatocytes) which severely compromises the normal liver function. An acute HBV Infection (less than 6 months) may resemble the fever, flu, muscle aches, joint pains and general being unwell. The symptoms specifying those states are: dark urine, loss of appetite, nausea, vomiting, jaundice, pain up the liver. Chronic hepatitis B is the infection persisting more than 6 months, the clinical features of that state correspond to the liver dysfunction, so the following signs may be noticed: enlarged liver, splenomegaly, hepatosplenomegaly, jaundice, weakness, abdominal pain, confusion and abdominal swelling.

The dataset of patient's probability of survival is used in the given example. The dataset contains 155 records of which 32 patients die and 123 survive. There are 20 attributes (including the class attribute) – 14 nominal and 6 numerical. All the symptoms found in the patient's record are the following:

1. CLASS: {DIE, LIVE}
2. AGE: {10, 20, 30, 40, 50, 60, 70, 80}
3. SEX: {male, female}
4. STEROID: {no, yes}
5. ANTIVIRALS: {no, yes}
6. FATIGUE: {no, yes}
7. MALAISE: {no, yes}
8. ANOREXIA: {no, yes}
9. LIVER BIG: {no, yes}
10. LIVER FIRM: {no, yes}
11. SPLEEN PALPABLE: {no, yes}
12. SPIDERS: {no, yes}
13. ASCITES: {no, yes}
14. VARICES: {no, yes}
15. BILIRUBIN: {0.39, 0.80, 1.20, 2.00, 3.00, 4.00}
16. ALK PHOSPHATE: {33, 80, 120, 160, 200, 250}
17. SGOT: {13, 100, 200, 300, 400, 500}
18. ALBUMIN: {2.1, 3.0, 3.8, 4.5, 5.0, 6.0}
19. PROTEINE: {10, 20, 30, 40, 50, 60, 70, 80, 90}

20. HISTOLOGY: {no, yes}

B. Application of the Rough Set Theory

During the experiment, the data was divided randomly into two datasets in the rate of 50 : 50 % by *Orthogonal Array-Based Latin Hypercubes* (OABLH) method [15]. In Orthogonal sampling, the sample space is divided into equally probable subspaces. All sample points are then chosen simultaneously making sure that the total ensemble of sample points is a Latin Hypercube sample [16] and that each subspace is sampled with the same density. The first dataset (T) is applied for train the algorithms and the second one (C) is used for the classification and the rules estimation. The results do not depend on the dataset division. The test dataset (C) has 77 records, of which 12 patients died and 65 survived. The train dataset (T) has 78 records, of which 20 patients died and 58 survived.

C. Number of attributes reduction

The result of the approximate reduct is {BILIRUBIN, ALK_ PHOSPHATE, SGOT, ALBUMIN, PROTEIN}. In this case we have numeric attributes only, but the approximate reduct can be a combination of any available attributes. The attribute reduction is $(20 - 5) / 20 = 75\%$. The train dataset (T) was used for selection of classification rule and then the rules were applied to the test dataset (C). The result is shown in Table II. We can observe the increase of accuracy after reducing the number of attributes. The same is shown in the confusion matrix in Table 1.

The *confusion matrix* [17] is a specific table layout that allows visualization of the performance of an algorithm. Each column of the matrix represents the instances in a predicted class, while each row represents the instances in an actual class. The matrix also shows the overall *accuracy* of the classifier as the percentage of correctly classified patterns in a given class divided by the total number of classified patterns. The overall *coverage* is the percentage of whole classified patterns divided by the total number of patterns. The *specificity* measures the proportion of messages that are negative of all the messages that are actually negative. The *sensitivity* is the proportion of messages that are positive of all the messages that are actually positive. In general here, Sensitivity means the accuracy on the class Negative, and Specificity means the accuracy on the class Positive.

Confusion matrix for training dataset (T) initially, after filtering and after reduction of attributes Table 1

	Initially			After filtering			After reduction		
	LIVE	DIE	NC	LIVE	DIE	NC	LIVE	DIE	NC
LIVE	37	2	19	36	0	22	46	4	8
DIE	3	0	17	3	11	6	2	14	4

NC – not classified

D. Decomposition of attributes value

After the subset of attributes was created, another algorithm is used to generate decompositions of attribute value sets. As Bazan et al. suggested [18] the decomposition may be done by discretisation of numerical attributes or by grouping (quantisation) of nominal attributes. The decomposition algorithm indicates the following division:
BILIRUBIN into intervals:

$$\langle -\infty;1.8] \cup \langle 1.8;2.6] \cup \langle 2.6;3.645] \cup \langle 3.645;\infty\rangle;$$

ALK PHOSPHATE into intervals:

$$\langle -\infty;149.0] \cup \langle 149.0;236.5] \cup \langle 236.5;\infty\rangle;$$

SGOT into intervals:

$$\langle -\infty;68.5] \cup \langle 68.5;\infty\rangle;$$

PROTEIN into intervals:

$$\langle -\infty;26.0] \cup \langle 26.0;44.5] \cup \langle 44.5;\infty\rangle;$$

The result of the test of the decision rules created after discretisation can be found in Table III. The global accuracy increased by about 2.85 % but however, the global coverage decreases by about 2.57%. It means that the rules can better classified unseen cases. The confusion matrix after decomposition is shown in Table 2.

Confusion matrix for training dataset (T) after attributes decomposition Table 2

	LIVE	DIE	NC
LIVE	46	2	10
DIE	2	14	4

Result of application training dataset (T) classification rules to test dataset (C) Table 3

	Initially [%]	After filtering [%]	After reduction [%]	After decomposition [%]
Global accuracy	88.10	94.00	90.90	93.75
Global coverage	53.85	64.10	84.62	82.05
sensitivity	92.50	92.31	95.83	95.83
specificity	0	100.00	77.78	87.50

E. Results and discussion

In this study, hepatitis disease diagnosis was conducted by the use of a novel medical decision support system based on the Rough Set Theory and modification of filtering / reducing algorithm. The obtained maximal diagnostic accuracy is 94% and effective 93.75% using LEM2 algorithm to generate decision rules. Sensitivity and specificity for the hepatitis disease dataset were obtained as 0, 100, 77.78, and 87.50%, respectively.

Only half of the data was used as a training set, which shows the underestimated power of that solution for the future use in medical data analysis. Moreover, the sensitivity and specificity values for the hepatitis disease dataset were obtained as 92.50%, 92.31% and 95.83%, respectively. These obtained values are shown in Table 3. It contains four columns indicating accuracy and coverage of four different steps of the algorithm: initially, after filtering, after reduction and after decomposition. Similar misclassification occurs after filtering (94%) and after decomposition (93.75%). Table 2 presents confusion matrix, which shows the misclassification of the rules. In the confusion matrix, each cell contains the raw number of examples classified for the corresponding combination of desired and actual network outputs. By combining the Rough Set and filter / reduce algorithm modification, the obtained classification accuracy is the highest among classifier reports found by Polat and Gune [19] in literature. In the view of classification accuracy, Table IV shows my accuracy classification methods with comparison to other methods.

A new medical diagnosis system gives very promising results in classifying the healthy and ill patients suffering from the hepatitis disease. I propose a complimentary system that can be implemented into the medical diagnostic devices. The benefit of the system is to assist the physician to make the final decision without hesitation.

Literature example of classification accuracies for hepatitis disease classification problem Table 4

Author	Method	Classification accuracy (%)
Gudzinski	Weighted 9NN	92.90
Gudzinski	18NN, stand. Manhattan	90.20
Gudzinski	15NN, stand. Euclidean	89.00
Adamczak	FSM with rotations	89.70
Adamczak	RBF (Tooldiag)	79.00
Adamczak	MLP+BP (Tooldiag)	77.40
Bradley, Diaconis	Bootstrap	84.00
Stern and Dobnikar	LDA, linear discriminant analysis	86.40
Stern and Dobnikar	ASI	82.00
Stern and Dobnikar	LFC	81.90
Norbert Jankowski	IncNet	86.00
Ozyıldırım, Yıldırım	MLP	74.37
Ozyıldırım, Yıldırım	RBF	83.75
Ozyıldırım, Yıldırım	GRNN	80.00
Polat, Gune	FS & AIR	92.59
Kanik	Rough Set	93.75

4. Conclusion

It was proved that the proposed algorithm is capable of confirm the people suffering from viral hepatitis – based on the real biometric data. Further work can lead into increasing overall algorithm accuracy and deeper data analysis as well. Combining the Rough Set Theory and modified pre-processing algorithm revealed some possibilities of their use in many other domains.

References

[1] PAWLAK, Z.: Rough Sets. *Intern. J. of Computer and Information Sciences,* 1982, 11, 341–346

[2] DOMINO, M., KANIK, T.: WEKA – *Empirical Study on Applications of Data Mining Techniques in Veterinary Medicine,* papers delivered at the Open Source Software in Education, Research and IT Solutions conference, Zilina, June 2011, Bratislava : SOIT, 2010, pp. 75–86. (doi:http://dx.doi.org/10.5300/2011-OSSConf/75)

[3] ZIARKO, W., SHAN, N.: *Discovering Attribute Relationships, Dependencies and Rules by Using Rough Sets:* Proc. of the 28th Annual Hawaii Intern. Conference on System Sciences, Wailea, Hawaii, 3–6 Jan 1995. Hawaii, 1995, pp. 293–299. (doi: 10.1109/HICSS.1995.375608)

[4] PAWLAK, Z.: Rough Set Theory and Its Applications. *J. of Telecommunications and Information Technology,* 2002, 3, 7–10.

[5] SMOLINSKI, T., CHENOWETH, D., ZURADA, J.: *Application of Rough Sets and Neural Networks to Forecasting University Facility and Administrative Cost Recovery,* Artificial Intelligence and Soft Computing – ICAISC 2004, Lecture Notes in Computer Science. Berlin: Springer, 2004, 80(3070), pp. 538–543. (doi: 10.1007/978-3-540-24844-6_80)

[6] SMOLINSKI, T., CHENOWETH, D., ZURADA, J.: *Time Series Prediction Using Rough Sets and Neural Networks Hybrid Approach:* Proc. of the Intern. Conference on Neural Networks and Computational Intelligence (NCI 2003), Cancun, May 2003, Cancun : ACTA Press, 2003, pp. 108–111.

[7] HASSANIEN, A. E., ABDELHAFEZ, M. E., OWN, H. S.: Rough Sets Data Analysis in Knowledge Discovery: A Case of Kuwaiti Diabetic Children Patients. *Advances in Fuzzy Systems,* 2008, 528461, pp. 13. (doi:10.1155/2008/528461)

[8] SLEZAK, D., WROBLEWSKI, J.: *Classification Algorithms Based on Linear Combinations of Features,* Principles of Data Mining and Knowledge Discovery, Lecture Notes in Computer Science, Prague : Springer-Verlag, 1999, pp. 548–553. (doi: 10.1007/978-3-540-48247-5_72)

[9] WROBLEWSKI, J.: Genetic Algorithms in Decomposition and Classification Problem. In: L. Polkowski, A. Skowron (eds.). *Rough Sets in Knowledge Discovery.* vol. 3. Berlin : Physica–Verlag, Heidelberg, 1998, pp. 471–487.

[10] POLKOWSKI, L., ATIEMJEW, P.: *On Granular Rough Computing: Factoring Classifiers Through Granulated Decision System. Proc. of the Intern. Conference RSEISP '2007.* Warsaw : Springer, 2007, pp. 271. (doi: 10.1.1.104.6475)

[11] HU, X.: *Knowledge discovery in databases: An Attribute-oriented Rough Set Approach.* Ph.D. thesis, University of Regina, 1995. (doi: 10.1.1.21.1988)

[12] THANGAVEL, K., JAGANATHAN, P., PETHALAKSHMI, A., KARNAN, M.: Effective Classification with Improved Quick Reduct for Medical Database Using Rough System. *J. of Bioinformatics and Medical Engineering,* 2006, 1(5), pp. 7–14.

[13] UCI machine learning repository [online]. Hepatitis Data Set [viewed 12 December 2012] Available from: http://www.ics.uci.edu/~mlearn/MLRepository.html

[14] WORMAN, H.: *The Liver Disorders and Hepatitis Sourcebook.* 2nd ed. United States : McGraw-Hill, 2006, pp. 17–36.

[15] TANG, B.: Orthogonal Array-Based Latin Hypercubes. *J. of the American Statistical Association,* 1993, 88(424), pp. 1392–1397.

[16] MCKAY, M. D., BECKMAN, R. J., CONOVER, W. J.: A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics,* WSC '05: Proc. of the 37th conference on Winter simulation, 2000, 42(1), pp. 202–208. (doi: 10.2307/1271432)

[17] STEHMAN, STEPHEN, V.: Selecting and Interpreting Measures of Thematic Classification Accuracy. *Remote Sensing of Environment,* 1997, 62(1), pp. 77–89.

[18] BAZAN, J., NGUYEN, H. S., SYNAK, P., WROBLEWSKI, J.: Rough Set Algorithms in Classification Problem, *Rough Set Methods and Applications.* Heidelberg : Physica-Verlag, 2000, pp. 49–88.

[19] POLAT, K., GUNE, S.: Hepatitis Disease Diagnosis using a new Hybrid System Based on Feature Selection (FS) and Artificial Immune Recognition System with Fuzzy Resource Allocation. *Digital Signal Processing,* 2006, 16, pp. 889-901. (doi: 10.1016/j.dsp.2006.07.005).