METHOD FOR EVALUATING OF DISCREPANCY BETWEEN REGULARITIES SYSTEMS IN DIFFERENT GROUPS

Oleg Senko, Anna Kuznetsova, Natalia Malygina, Irina Kostomarova

Abstract: *A new method of data analysis is discussed. Goal of represented techniques is complete and statistically valid comparing of regularities existing in two different groups of objects. It is supposed that regularities that tie levels of forecasted and explanatory variables are searched with the help of optimal partitioning technique. The developed technique was applied for analysis of genetic factors impact on severity of discirculatory encephalopathy (DEP). At the first stage computer method for evaluating of DEP severity was developed with the help of pattern recognition techniques. It was revealed that computer estimates of severity rather strongly correlate with DD variant of gene coding angiotensin-converting enzyme (ACE). Systems of regularities that ties various clinical or biochemical factors with computer estimates of severity were found with the help of optimal valid partitioning (OVP) method in groups of patients with three different variants of gene coding ACE. Statistically significant discrepancies were found between revealed regularities systems with the help of developed methods of comparing.*

Keywords: *Optimal partitioning, statistical validity, permutation test, regularities, explanatory variables effect,pattern recognition, discirculatory encephalopathy, genetic factors*

ACM Classification Keywords: *H.2.8 Database Applications - Data mining, G.3 Probability and Statistics - Nonparametric statistics, Probabilistic algorithms*

Introduction

Goal of this paper is development of method for statistically valid comparing of regularities that were found in two sets of objects. It is supposed that objects from both sets are described by the same variables. Such task arises in particular when we try to evaluate effect of gene's variant on biological processes related to severity of some disease. In present paper developed technique is used to evaluate influence polymorphous variants of genes coding angiotensin-converting enzyme (ACE), lipoprotein lipase (LPL) and cholesterol ester transfer protein (CETP) on biological processes related to severity of dyscirculatory encephalopathy [Vodolagina et al,2010]. Previously developed in works [Senko and Kuznetsova,1998], [Senko et al, 2003], [Senko and Kuznetsova, 2006] optimal valid partitioning (OVP) technique was used for regularities searching. OVP technique belongs to family of data analysis methods aimed to study regularities that may be described as subregions in explanatory variables space where levels of forecasted variable Y significantly differ from Y levels in neigbouring subregions . Method of logical regularities searching [Ryazanov,2003] or various models of classification or regression trees may be mentioned thereupon [Kuncheva, 2004].

OVP method. OVP technique implements partitioning of explanatory variables $X_1,\ldots,X_n$ space that allows to achieve possibly separation of forecasted variable Y levels. Forecasted variable Y in particular may be binary variable indicating two dichotomous groups. Optimal partitions are searched by initial dataset

$\tilde{S}_* = \{(y_1, \mathbf{x}_1), \ldots, (y_r, \mathbf{x}_r)\}$ where y_j is value of Y and $\mathbf{x}_j$ is vector of $X-$ variables for object s_j. Partition with the maximal value of special quality functional F_Q is searched inside several partitions families that differ from each other by dimensions, complexities of geometrical forms. Revealed regularities in OVP method are verified with the help of permutation tests [Gorman, 2001]. Permutation test implement verification by comparing maximal value of F_Q that was achieved at initial true dataset with maximal F_Q values that were calculated at variety of artificial datasets. At that artificial datasets are generated from initial dataset by random permutation of Y values relatively fixed position of $\mathbf{x}$ descriptions. Statistical validity of regularity (p-value) is estimated as fraction of random permutations for which maximal F_Q at artificial dataset exceeds maximal F_Q at initial dataset. Besides functional F_Q and p-value additional validity measure F_{2Q} is used that is ratio of maximal value of F_Q that was achieved at random dataset to optimal F_Q value at initial dataset. In case of more complicated regularities validity is evaluated with the help of additional second variant of permutation test. Instead of testing null hypothesis that Y is completely independent on $X-$ variables second variant implement testing of null hypotheses that Y is independent on $X-$ variables inside subregions of $X-$ space related to more simple regularities that were previously revealed for the same variables. Let r^{ij} is two-dimensional regularity with one boundary for each variable from pair (X_i, X_j). Let statistically valid simplest regularities exist for variables related to two-dimensional regularity r^{ij}. Then two p-values and two values of F_{2Q} may be calculated for r^{ij}: p_1 and F_{2Q}^1 are calculated from null hypothesis that is independent on (X_i, X_j) inside two subregions formed by partition of X_i range, p_2 and F_{2Q}^2 are calculated from null hypothesis that is independent on (X_i, X_j) inside two subregions formed by partition of X_j range. In case there is no simple valid regularity for one or for both variables from (X_i, X_j) values p_1, F_{2Q}^1 and/or p_2, F_{2Q}^2 are calculated from null hypothesis that Y is completely independent on (X_i, X_j). Using p-values and functionals p_1, F_{2Q}^1 and p_2, F_{2Q}^2 allows to evaluate contribution of each explanatory variable to two-dimensional regularity

OVP method was rather successfully used in several biomedical tasks [Kuznetsova et al, 2000]. However there is problem arising during OVP using that is related to large complexity of regularities systems in high-dimensional tasks where number of regularities may achieve several hundreds. Some additional mathematical tools are necessary to help researcher to evaluate regularity system. One of possible approaches that is associated with classification of regularities was suggested in [Senko et al, 2010]. Aanother method is suggested here that allows evaluating importance of each explanatory variable by calculating corresponding index. Importance of explanatory variables X is calculated as sum of contributions of X to various regularities. Let $\tilde{R} = \{r^{ij}\}$ is found regularities system. Index $\gamma(X_i)$ characterizing importance of X_i may be calculated as sum

$$\gamma(X_i, \tilde{R}) = \sum_{r_*^{ij} \in \tilde{R}} F_{2Q}^1(r_*^{ij}) + \sum_{r_*^{ji} \in \tilde{R}} F_{2Q}^2(r_*^{ji}).$$

Another problem that arises particularly in genetic researches is necessity of comparing of regularities that are found for the same explanatory variables in different groups of patients.

Evaluating of genetic factors effect in patients with discirculatory encephalopathy

Discirculatory encephalopathy is form of cerebral and vascular chronic insufficiency that is caused by worsening of blood supply in brain tissue [Maksudov, 1975].

Clinical data. Database developed at neurology department of "Scientific and clinical center (SCC) gerontology" was used in represented researches. This database includes information about variety of anamnestic, clinical, laboratory indicators in population of 358 patients with range in age from 35 to 102. Database also includes information about results of instrumental method (Magnetic Resonance Tomography, Ultrasonic Imaging, Electrocardiography and others) and results of molecular and genetic screening of genes coding ACE, LPL and CETP that was received in laboratory of age specific population genetics of "SCC gerontology". Levels of many indicators in group with age over 89 deviate significantly from mean levels in group under 89. So group over 89 was deleted from analysis. Population under 89 includes three groups of patients with different severity stages: 43 patients with first stage, 145 patients with second stage and 47 patients with third stage.

Computer method of severity evaluating. Influence of explanatory variables on severity may be estimated by using or standard statistical tests or OVP technique in groups with first or third severity stages. However full number of patients with first and third stages (90) is too small to evaluate effect of explanatory variables inside groups with fixed gene variant**.** At that number of patients with second intermediate severity stage significantly exceeds number of patients with first and third stages. It is possible to calculate computer estimates of severity by valid clinical or laboratory indicators for all patients using algorithm previously trained at dataset including first and third severity groups.

Performance of several pattern recognition methods from program system «RECOGNITION» was evaluated with the help of cross validation technique. «RECOGNITION» is program system for intellectual data analysis including collection of pattern recognition techniques that are based on different approaches: standard statistical methods, methods based on voting by systems of regularities, logical and combinatorial methods, neural networks, linear divisors, support vector machine, classifying trees and others. Program system also includes tools for calculating collective solutions. Goal of our researches was evaluating of relationship between estimates of severity and genetic factors. So only those explanatory variables were used that corresponds to anamnestic, clinical, laboratory indicators and results of instrumental methods. Results of molecular and genetic screening were not used as predicting factor at this stage of studies. It must be noted that one of explanatory variable indicating that patient underwent acute stroke strongly correlate with cerebral encephalopathy severity and its contribution to solving rules may significantly exceed contributions of another variables. So it is interesting to evaluate performance of pattern recognition methods without indicator of acute cerebral disorders. Correct recognition rates by full set of variables and by set without indicator of acute cerebral disorders (ACD) are given in table 1.

It is seen that forecasting ability of recognition algorithms without indicator of acute cerebral disorders is somewhat lower forecasting ability of whole set of variable. Algorithms trained at groups with first and third severity stages were then used for computer diagnostics of second group. Besides single algorithms collective rules were also studied that ascribe recognized object to a class where it is put by majority of algorithms. Calculated estimate for groups with first or third severity stages will be further referred to as calculated severity estimates. Calculated severity estimates were compared with genetic factors indicating variants of genes coding

ACE, LPL or CETP . Statistical validity of relationship between genetic factors and calculated severity estimates was evaluated with the help of standard statistical criteria Wilcoxon-Mann-Whitney (WMW), Chi-square(X^2).

Table1. Diagnostic abilites of different pattern recognition models by full set of explanatory variable and by set of explanatory variables without indicator of acute stroke evaluated with the help of cross validation technique.

Pattern recognition models	Full set 125 explanatory variables	Without indicator of acute stroke
Linear Fisher discriminant (LDF)	77,00%	74,50%
k-nearest neighbors (k-NN)	67,80%	74,30%
Statistically weighted syndromes (SWS)	85,50%	76,50%
Support vector machine (SVM)	83,3 %,	80,00%
Multiplicative neural network (MNN)	81,00%	67,70%
Method based on voting by Logical regularities (LR)	89,90%	77,80%

Besides for evaluating of relationship uni-dimensional OVP model was used. It was shown that for many algorithms calculated severity estimates rather strongly correlate with indicators of DD variant of ACE coding gene. At that very good correlation exists for ensemble of 4 algorithms: LDF, k-NN, SVM,SWS. Results of studies are given in Table 2.

Table 2. Statistical validity of relationship between calculated severity estimates and genetic factors

Recognition model	WMW		X^2		OVP	
	Full set	Without AS	Full set	Without AS	Full set	Without AS
LDF	0.007	0,01	0.07	0,05	0.08	0,04
SVM	0.024	0,02	0.017	0,01	0.05	0,01
SWS	0.06	0,02	0.07	No	0.07	0,05
Collective solution	0,014	0,05	0.014	0.0038	0.007	0,01

Distributions of calculated severity stages in group with DD variant and in group with alternative variants are given in table 3. It is seen that that in group with DD variant number of cases with calculated third stage of severity slightly exceeds number of cases with the calculated first stage of severity. At the same time number of cases with the first stage more than two times exceeds number of cases with the third stage.

Tabl. 3. Distributions of calculated severity stages in group with DD variant and in group with DI or II variants

	ID or II variants of gene coding ACE number of cases	DD variant of gene coding ACE number of cases
First severity stage	82	22
Third severity stage	39	28

So it was shown that computer estimates of DE severity in group of patients with variant DD of ACE significantly differ from DE severity in group of patients with variants ID and II. It is interesting to unveil causes of existing effect.

Regularities systems comparing

Possible approach is comparing of regularities that tie *DE* severity and corresponding levels of clinical, biochemical or genetic indicators in two groups of patients. Let $\tilde{R}_{dd}$ is set of valid regularities that were found for group of patients $\tilde{S}_{dd}$ with variants *DD*. Then effect of variant *DD* may be compared with effect of another variant zz by evaluating accordance between $\tilde{R}_{dd}$ and regularities characterizing dependence of *DE* severity on the same explanatory variables in group of patients $\tilde{S}_{zz}$ with variant zz. This problem may be also referred to as problem evaluating discrepancy between regularities in two groups of objects. Different approaches may be used. For example regularities system $\tilde{R}_{zz}$ may be found for group of patients $\tilde{S}_{zz}$. Then we may compare systems $\tilde{R}_{dd}$ and $\tilde{R}_{zz}$. Effect of variant may be described by selecting following regularities from $\tilde{R}_{dd}$:

a) in case regularity $r_{dd}^{ij} \in \tilde{R}_{dd}$ exists for pair of explanatory variables (X_i, X_j) and there no regularities for (X_i, X_j) in $\tilde{R}_{zz}$;

b) in case regularities $r_{dd}^{ij} \in \tilde{R}_{dd}$ and $r_{zz}^{ij} \in \tilde{R}_{zz}$ exist for pair of explanatory variables (X_i, X_j) but difference between regularities is sufficiently great .

Difference between regularities may be evaluated as difference between forecasting functions associated with regularities. Union of regularities sets satisfying conditions (a) and (b) will be referred to as $\tilde{R}_{dd\setminus zz}$

There is important drawback in described approach. In case (a), absence of regularity in $\tilde{R}_{zz}$ may be related not to quite different effect of zz variant relatively *DD* variant. Regularity for zz variant may be less expressed. Thus it is not revealed at corresponding significance level. So another approach was developed that allows describing difference between two types of effects more exactly.

It is supposed that regularity $r_{dd}^{ij} \in \tilde{R}_{dd}$ was found with the help of OVP technique by and r_{dd}^{ij} consist of k subregions $q_1, \ldots, q_k$. Let $n_i^1(r_{dd}^{ij}, \tilde{S}_*)$, $n_i^3(r_{dd}^{ij}, \tilde{S}_*)$ are numbers of cases with calculated first and third severity

stages in dataset $\tilde{S}_* \cap q_i$, $fr_i^1(r_{dd}^{ij}, \tilde{S}_*) = n_i^1(r_{dd}^{ij}, \tilde{S}_*)/[n_i^1(r_{dd}^{ij}, \tilde{S}_*) + n_i^3(r_{dd}^{ij}, \tilde{S}_*)]$. Set of ratios $\{fr_1^1(r_{dd}^{ij}, \tilde{S}_*), \ldots, fr_k^1(r_{dd}^{ij}, \tilde{S}_*)\}$ describes empirical dependence of DE on (X_i, X_j) in dataset $\tilde{S}_* \cap q_i$

Deviation between dependences may be evaluated with the help of functional

$$F_Q^{\Delta}(r_{dd}^{ij}, \tilde{S}_{dd}, \tilde{S}_{zz}) = \sum_{i=1}^{k} \{[fr_i^1(r_{dd}^{ij}, \tilde{S}_{dd}) - fr_i^1(r_{dd}^{ij}, \tilde{S}_{zz})]^2 \sqrt{n_i^1(r_{dd}^{ij}, \tilde{S}_{dd}) * n_i^1(r_{dd}^{ij}, \tilde{S}_{zz})}\}$$

So optimal partitions are searched by $\tilde{S}_{dd}$ and functional $F_Q^{\Delta}(r_{dd}^{ij}, \tilde{S}_{dd}, \tilde{S}_{zz})$ is calculated by $\tilde{S}_{dd}$ and $\tilde{S}_{zz}$ Important issue is validation of differences evaluated with the help of $F_Q^{\Delta}(r_{dd}^{ij}, \tilde{S}_{dd}, \tilde{S}_{zz})$. The same variants of permutation tests that were used in previous main version of OVP technique may be also used for comparing of two set of regularities. Pairs of artificial datasets $(\tilde{S}_{dd}^r, \tilde{S}_{zz}^r)$ are generated from $\tilde{S}_{dd}$ and $\tilde{S}_{zz}$ by random permutations of Y values relatively fixed position of $\mathbf{x}$ descriptions. Then again optimal partitions are found by $\tilde{S}_{dd}^r$ and functional F_Q^{Δ} is calculated by $(\tilde{S}_{dd}^r, \tilde{S}_{zz}^r)$. The same types of permutation test that was used in previous OVP technique are used in new developed method for evaluating discrepancy between regularities in two groups. Values of functional F_Q^{Δ} calculated by $(\tilde{S}_{dd}^r, \tilde{S}_{zz}^r)$ are compared with functional F_Q^{Δ} value for initial pair $(\tilde{S}_{dd}, \tilde{S}_{zz})$ and p-values are evaluated as fractions of $(\tilde{S}_{dd}^r, \tilde{S}_{zz}^r)$ pairs, for which F_Q^{Δ} value exceeds F_Q^{Δ} value for pair $(\tilde{S}_{dd}, \tilde{S}_{zz})$. Functional F_{2Q} and p_1, F_{2Q}^1, p_2, F_{2Q}^2 values are calculated according the same scheme that was used in previous variant of OVP technique. System of regularities that is received using validation with the help of functional F_Q^{Δ} and characterizes deviation between dependences existing in groups $\tilde{S}_{dd}$ and $\tilde{S}_{zz}$ will be referred to as $\tilde{R}_{dd\backslash zz}^{\Delta}$. Table 4 represents discrepancy between effect of gender on calculated severity estimated in groups with DD and II variants of ACE.

Table 4. Discrepancy between effect of gender on calculated severity estimated in $\tilde{S}_{dd}$ and $\tilde{S}_{ii}$

	DD		*II*	
	First stage	*Third stage*	*First stage*	*Third stage*
Female	*23*	*13*	*18*	*5*
Male	*3*	*12*	*5*	*0*

It is seen from table that there are 5 cases with calculated first stage of severity and no cases with calculated third stage in males with II variant of ACE. In males with DD variant of ACE there are 3 cases with first calculated severity stage and 12 cases with third calculated severity stage. So, rate of calculated third stage increases dramatically for DD variant. Statistical validity of difference between dependences of calculated severity stage on gender in group with DD and II variant of ACE was evaluated at $p<0.01$ with the help of permutation test using functional F_Q^{Δ}.

Figure 1 represents discrepancy between effect of cholesterol and thyrocsyn containments on calculated severity in group $\tilde{S}_{dd}$ with DD variant of gene coding ACE and in group $\tilde{S}_{id}$ with ID variant of gene coding *ACE*. At the left side of figure regularity is represented that ties calculated *DE* severity and two abovementioned explanatory

variables in group $\tilde{S}_{dd}$ and at the right side $\tilde{S}_{id}$ group empirical distribution is represented for the same pair of explanatory variables. It is seen that quadrant II at left part of figure contains 4 cases with calculated third severity stage and the same quadrant II at right part of figure contains 10 cases with calculated first severity stage. Statistical validity of difference between distributions represented al left and right parts of figure was evaluated at $p<0.01$ with the help of permutation test using functional F_Q^{Δ} .

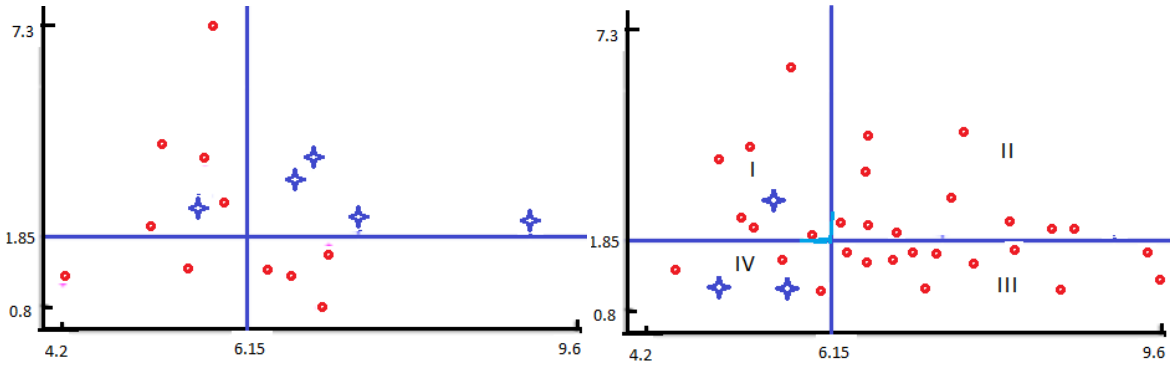

Fig 1. At the left part of figure regularity for Axis X corresponds containment of cholesterol in blood (mmmol/l), Y-containment of T3 (mmol/l)

- case with calculated third stage of severity,

- case with calculated first stage of severity.

Index γ characterizing importance of explanatory variables also may be used in method for evaluating discrepancy between regularities in two groups. It may be calculated as $\gamma(X_i,\ \tilde{R}^{*}_{dd\backslash zz}) = \sum_{r_*^{ij}\in\tilde{R}^{*}_{dd\backslash zz}} F^{1}_{2Q}(r_*^{ij}) + \sum_{r_*^{ji}\in\tilde{R}^{*}_{dd\backslash zz}} F^{2}_{2Q}(r_*^{ji})$, where $\tilde{R}^{*}_{dd\backslash zz}$ is equal $\tilde{R}^{\Delta}_{dd\backslash zz}$ or $\tilde{R}_{dd\backslash zz}$

Let by $\tilde{R}_{dd}$ and $\tilde{R}_{id}$ systems of two-dimensional regularities that were found in group $\tilde{S}_{dd}$ and in group $\tilde{S}_{id}$ with correspondingly. System $\tilde{R}^{\Delta}_{dd\backslash id}$ includes two-dimensional regularities characterizing deviation between dependences existing in groups $\tilde{S}_{dd}$ and $\tilde{S}_{id}$. Ten explanatory variables with greatest values of γ indices in discussed systems of regularities are represented in table 5.

It is seen that γ indices differently evaluate significance of explanatory variables for regularities from system $\tilde{R}_{dd}$ and $\tilde{R}_{id}$. In group $\tilde{S}_{id}$ the most informative variable is indicator of acute stroke. The γ index for acute stroke is more than 1.5 times than γ index for $\beta-$lipoproteins levels that is the most important variable in $\tilde{R}_{id}$ system besides indicator of acute stroke. In group $\tilde{S}_{dd}$ the most informative variable is indicator of thyroid ganglions that are detected by ultrasonic scanning. The γ index for this variable is more than 1.5 times greater than γ index for acute stroke. The only variables that belong to ten most important in $\tilde{R}_{dd}$ and $\tilde{R}_{id}$ are indicator of acute stroke and indicator heart conduction abnormality detected by ECG. It is seen that discrepancy between $\tilde{R}_{dd}$ and $\tilde{R}_{dd/id}$ is even greater. The only variable that is important in both sets of regularities is indicator of thyroid ganglions that are detected by ultrasonic scanning.

Table 5. Ten explanatory variables with greatest values of γ indices calculated are represented in table 5.

$\tilde{R}_{dd}$		$\tilde{R}_{id}$		$\tilde{R}^{\Delta}_{dd \setminus id}$	
Variable	γ indices	*Variable*	γ indices	*Variable*	γ indices
Thyroid ganglions by US scanning	13,175	Acute stroke	24,7728	Tortuous vessels by US diagnostics	15,1705
Alcohol	9,0115	β – lipoproteins	16,3555	Fibrosis gradations	7,9377
Acute stroke	8,2958	Abnormality of conduction by ECG	15,4514	Thyroid ganglions by US scanning	7,5158
Abnormality of conduction by ECG	7,5347	Hepatosis	15,2711	Alanin-amin-transferaze gradations	7,1757
Glucose	6,9559	Obesity stages	14,6469	Ultrasonic scanning-Sz	6,343
Hydrocephaly by MRT	6,8861	High density lipoproteins	13,2375	Br	5,7808
Nodular goiter	6,6777	H-H- variant of LPL gene	11,3931	Ultrasonic scanning-EC	5,6907
Red(blood)cells	6,4063	Type II diabetes	10,1628	Very low density lipoproteins	5,6695
Pancreas fibrosis	6,2065	Systolic pressure	10,1422	Free cholesterol gradations	4,9893
Prostate cancer	6,1434	Age gradations	9,8108	Ultrasonic scanning-atherosclerosis	4,6961

Conclusion

Results that are represented in paper may be summarized as follows. Method for computer evaluating of severity in patients with discirculatory encephalopathy was developed. It is appeared that severity estimates calculated with the help of computer recognition methods by variety of clinical, biochemical or instrumental indicators rather strongly correlates with indicator of DD variant of gene coding for angiotensin-converting enzyme. At that severity of DE is higher in patients with DD variant. To study the effect of ACE coding gene variants it was suggested to compare dependence of DE severity on variety of factors in groups with different variants of ACE gene. Method for evaluating of discrepancy between dependencies existing in two independent groups. At the first step optimal partitioning is searched by one of the groups. At the second step $F^{\Delta}_{Q}(r^{ij}_{dd}, \tilde{S}_{dd}, \tilde{S}_{zz})$ functional is calculated by regularity and two compared groups. Validity of discrepancy found at these two steps is evaluated with the help of special variant of permutation test using numerous recalculation of both steps for pairs of artificial random datasets. New method was developed that allows to calculate coefficients characterizing importance of single explanatory variables. So each system of regularities may be evaluated by vector of such coefficients. It was shown that there is statistically significant deviation between dependences of DE on variety of clinical and biochemical indicators in groups with DD and ID variant of gene coding ACE.

Acknowledgements

The paper is published with financial support by the project ITHEA XXI of the Institute of Information Theories and Applications FOI ITHEA (www.ithea.org) and the Association of Developers and Users of Intelligent Systems ADUIS Ukraine (www.aduis.com.ua).

Bibliography

[V.V.Ryazanov ,2003] Ryazanov V.V. About some approach for automatic knowledge extraction from precendent data // Proceedings of the 7th international conference "Pattern recognition and image processing", Minsk, May 21-23, 2003, vol. 2, pp.35-40.

[Gorman, 2001] T.W. O'Gorman An adaptive permutation test procedure for several common test of significance. Computational Statistics & Data Analysis. 35(2001) 265-281.

[Kuncheva, 2004] L.I. Kuncheva, Combining Pattern Classifiers. Methods and Algorithms. Wiley Interscience, New Jersey, 2004.

[Senko et al, 2003] Senko O.V., Kuznetsova A.V., Kropotov D.A. (2003). The Methods of Dependencies Description with the Help of Optimal Multistage Partitioning. Proceedings of the 18-th International Workshop on Statistical Modelling Leuven, Belgium, 2003, pp. 397-401.

[Sen'ko et al, 1998] Sen'ko O.V., Kuznetsova A.V. (1998). The use of partitions constructions for stochastic dependencies approximation. Proceedings of the International conference on systems and signals in intelligent technologies. Minsk (Belarus), pp. 291-297.

[Kuznetsova et al, 2000] Kuznetsova A.V., Sen'ko O.V., Matchak G.N., Vakhotsky V.V., Zabotina T.N., Korotkova O.V. The Prognosis of Survivance in Solid Tumor Patients Based on Optimal Partitions of Immunological Parameters Ranges //J. Theor. Med., 2000, Vol. 2, pp.317-327.

[Sen'ko et al, 2006] Oleg V.Senko and Anna V. Kuznetsova, The Optimal Valid Partitioning Procedures . Statistics on the Internet ***http://statjournals.net/***, April, 2006

[Senko et al, 2010] Senko O, Kuznetsova A., Malygina N, Kostomarova I.,2010 Methods for evaluating of regularities systems structure. In book «New Trends in Classification and Data Mining.» ITHEA. Sofia. Bulgaria. P. 40-46, 2010.

[Vodolagina et al,2010] Vodolagina N.N., Kostomarova I.V., Kuznetsova A.V., Malygina N.A. Senko O.V. Study of relationship between severity discirculatory encephalopathy and polymorphism of some genetic factors with the help of pattern recognition methods. Reports of 3-rd international conference «Mathematical biology and bioinformatics», Puschino,Russia, October 2010, p.238-239.

[Maksudov, 1975] Maksudov G.A. Discirculatory encephalopathy. In book «Vascular diseases of nervous system» under editorship of acad. AMS USSR E.V. Shmidt.

Authors' Information

Senko Oleg Valentinovich *– Leading researcher in Dorodnicyn Computer Center of Russian Academy of Sciences, Russia, 119991, Moscow, Vavilova, 40,* senkoov@mail.ru

Kuznetsova Anna *– senior researcher in Institute of Biochemical Physics of Russian Academy of Sciences, Russia, 117997, Moscow, Kosygina, 4,* azfor@narod.ru

Malygina Galina *–chief of laboratory in Russian Research Institute of Gerontology, Moscow, Russia*

Kostomarova Irina *- Russian Research Institute of Gerontology, Moscow, Russia*