UDC 681.518:004.93

MACHINE LEARNING DECISION SUPPORT SYSTEMS FOR ADAPTATION OF EDUCATIONAL CONTENT TO THE LABOR MARKET REQUIREMENTS

Shelehov I. V. – Associate Professor, PhD, Head of the Department of Computer Science, Sumy State University, Sumy, Ukraine.

Prylepa D. V. – Assistant, Department of Computer Science, Sumy State University, Sumy, Ukraine.

Khibovska Yu. O. – Postgraduate Student of the Department of Computer Science, Sumy State University, Sumy, Ukraine.

Otroshcenko M. S. – Postgraduate Student of the Department of Computer Science, Sumy State University, Sumy, Ukraine.

ABSTRACT

Context. The urgent task of increasing the functional efficiency of machine learning of decision support system (DSS) for assessing compliance with content modern requirements of the educational disciplines of the graduation department based on the results of the employer survey has been solved.

Objective. Increasing the functional efficiency of machine learning of DSS for assessing compliance with modern requirements of the educational disciplines content of the first (bachelor's) level specialty educational and professional program based on machine learning and pattern recognition.

Method. The method of machine learning of DSS is proposed for adapting the educational content of the graduation department to the labor market requirements. The idea of the method is to maximize the information capacity of the DSS in the machine learning process, which allows in the monitoring mode to guarantee a high full probability of making the correct classification decisions. The method was developed as part of a functional approach to modeling cognitive processes of natural intelligence, which makes it possible to provide DSS with flexibility when retraining the system due to increasing the power of the recognition classes alphabet. The method is based on the principle of maximizing the amount of information in the machine learning process. The modified Kullback information measure, which is a functional of the accuracy characteristics of classification solutions, is considered as a criterion for optimizing machine learning parameters. According to the proposed functional category model, an information-extreme machine learning algorithm was developed based on the hierarchical data structure in the form of a binary decursive tree. The use of such a data structure allows you to automatically divide a large number of recognition classes into pairs of nearest neighbors, for which optimization of machine learning parameters is carried out according to a linear algorithm of the required depth. The geometric parameters of hyperspherical containers of recognition classes were considered as optimization parameters, which were restored in the radial basis of the binary space of Hamming features in the machine learning process. At the same time, the input traning matrix was transformed into a working binary training matrix, which was changed in the machine learning process through admissible transformations in order to adapt the input information description of the DSS to the maximum reliability of classification decisions.

Results. The informational, algorithmic, and software of the DSS was developed to assess the educational content quality based on the machine analysis results of respondents' answers. Within the framework of the geometric approach, based on the information-extreme machine learning results, highly reliable decisive rules, practically invariant to the multidimensionality of the recognition features space, were constructed based on the hierarchical data structure in the form of a binary decursive tree. The influence of machine learning parameters on the functional effectiveness of machine learning of the DSS was studied on the evaluation example of the educational content of the educational and professional bachelor's program of the specialty 122 Computer Science.

Conclusions. The computer modeling results confirm the high functional efficiency of the proposed method of information-extreme hierarchical machine learning and can be recommended for practical use in institutions of higher education to assess compliance with modern requirements of the educational content of graduation departments.

KEYWORDS: information-extreme machine learning, functional categorical model, information criterion, hierarchical data structure, decursive tree, educational content.

ABBREVIATIONS

IEIT is an information-extreme intellectual technology;

DSS is decision support system;

CNN is convolutional neural network.

NOMENCLATURE

M is a set of recognition classes;

m is a number of the recognition class;

N is a set of recognition features in the structured vector;

i is a number of the recognition feature;

J is a set of structured vectors of recognition features;

J is a number of the structured vector;

H is a set of tiers of decursive tree;

h is a number of the tier of decursive tree;

S is a set of strata of decursive tree;

s is a number of the stratum of decursive tree;

m_s is a serial number of the recognition class in the s-th stratum;

$\delta_{h,s}$ is a parameter that is equal to half of the control tolerances field of the recognition feature.

DOI 10.15588/1607-3274-2023-1-6

$x_{h,s,c}$ is the averaged feature vector of the recognition class neighboring in the stratum

δ_H is the field of normalized tolerances, which specifies the values range of control tolerances

$E_{h\,s,m_s}(d)$ is the informational criterion for optimization of machine learning parameters of DSS to recognize feature vectors of recognition class X^o_{h,s,m_s};

d is a distance measure that is equal to the value of the the recognition class container radius;

G_E is the working area of defining the information criterion;

G_d is the permissible area for changing the radii of the recognition class containers;

P is a set of thematic modules of the educational content, which are evaluated by the respondents;

T is a set of reading information time moments;

Ω is the recognition features space;

A is the recognition classes alphabet;

Y is the input Euclidean training matrix;

$Y^{|S|}$ is the set of input Euclidean traning matrices of recognition classes for all levels of the decursive tree;

$X^{|S|}$ is the set of working binary traning matrices for all strata of the decursive tree;

f_1 is the training matrix formation operator Y;

f_2 is the operator for constructing a decursive binary tree H;

f_3 is the training matrix formation operator $Y^{|S|}$;

f_4 is the training matrix formation operator $X^{|S|}$;

$\{k\}$ is a set of machine learning steps;

$\tilde{\Re}^{|M|}$ is a fuzzy division of the feature space into M recognition classes;

$Y^{|S|}$ is an input training matrix of recognition classes S strata of decursive tree;

$X^{|S|}$ is a binary training matrix of recognition classes S strata of decursive tree;

E is a term set of values of the information criterion;

R is an operator of construction of division $\tilde{\Re}^{|2|}$ of space of signs on recognition classes;

ψ is an operator for testing the basic statistical hypothesis about the affiliation of the vector x_{h,s,m_s} of the recognition class X^o_{h,s,m_s};

γ is an operator of the set formation of exact characteristics for the set system of decisions estimations;

φ is an operator for calculating the information criterion for optimizing the machine learning parameters;

U is an operator that regulates the machine learning process;

$K^{(1)}_{h,s,m_s}(d)$ is the number of events in which "own" feature vectors do not belong to recognition class to the container X^o_{h,s,m_s};

$K^{(2)}_{h,s,m_s}(d)$ is the number of events in which "foreign" feature vectors belong to the recognition class container X^o_{h,s,m_s};

10^{-p} is a sufficiently small number, which is entered to avoid division by zero (in practice, $p = 2$ was accepted);

$x^{(j)}$ is a structured feature vector that can be recognized;

$\mu_{m\,s}$ is the function of belongingness of feature vector $x^{(j)}$ to the recognition class container X^o_{h,s,m_s};

d^*_{h,s,m_s} is the optimal radius of the recognition class container X^o_{h,s,m_s}.

INTRODUCTION

Establishing a stable connection between the graduation department of a higher education institution and employers is a necessary condition for improving the quality of the educational process. The main way to establish such a connection is to survey employers and specialists in the relevant field of knowledge. The development of modern information and communication technologies makes it possible to automate data collection. At the same time, in practice, the results of the respondents survey are processed using the multidimensional statistical analysis methods. Such methods to ensure statistical stability and homogeneity require large volumes of input data, which requires organizers to spend significant time and money on processing and analyzing survey results.

The object of research is the process of sustainable monitoring of the education quality by creation of the DSS for adapting the educational content of the graduation department to the labor market requirements on the machine learning and pattern recognition basis.

The subject of research is the method of hierarchical information-extreme machine learning of DSS for adapting the educational content of the graduation department to the labor market requirements.

The purpose of the work is to increase the functional efficiency of the DSS machine learning during automatically forming an input traning matrix based on the respondents survey results, building decisive rules according to the optimal (here and in the text in the text in the informational sense) machine learning parameters, and in the monitoring mode assessing the compliance of the educational content of the graduation department with modern requirements.

DOI 10.15588/1607-3274-2023-1-6

1 LITERATURE REVIEW

The article considers the method of hierarchical information-extreme machine learning of DSS for adapting the educational content of the graduation department to the labor market requirements. According to European educational standards, the quality of education is determined by the benefits that both employers and graduates of a higher education institution will receive [1, 2]. One of the main ways of organizing sustainable monitoring of the education quality is the creation of the DSS for adapting the educational content of the graduation department to the labor market requirements [3]. A traditional approach to processing and analyzing the stakeholders survey results and specialists in the knowledge relevant fields is the application of decision-making systems existing in social communications using multidimensional statistical analysis methods [4–6]. At the same time, the disadvantages of statistical decision-making methods are the need for large volumes of data, ensuring their statistical stability and homogeneity. A promising direction for increasing the functional efficiency of computerized systems for assessing the quality of education is the use of intelligent information technologies for data analysis [7–9]. Works [10, 11] give examples of creating expert systems for evaluating the quality of the educational process based on fuzzy logic. The main disadvantages of such systems are that they are inflexible and do not provide feedback in the monitoring mode between the graduation department, employers and students of higher education of various forms of education. The further development of computerized systems for assessing the education quality is the development of the scientific and methodological foundations of the information synthesis of DSS, capable of automatically forming an input information description and identifying regularities based on machine learning and pattern recognition [12]. CNN [13, 14] is the most common among known intelligent information technologies for data analysis, but the main drawback of CNN is its sensitivity to the multidimensionality of the feature space and the recognition classes alphabet. Works [15, 16] consider the use of extractors based on artificial neural networks to reduce the impact of input data multidimensionality, but this approach is associated with the information loss possibility.

Works [17–19] consider the use of fuzzy neural networks for functional diagnosis, but at the same time there is also the multidimensionality problem, which significantly limits the capabilities of the fuzzy logic apparatus.

The main scientific and methodological reasons that complicate the use of CNN for the information synthesis of the DSS for the education quality assessment are:

– arbitrary initial conditions of the evaluation process;

– intersection in the recognition classes features space that characterize the corresponding levels of educational content quality;

– multidimensionality of the signs dictionary;

– the impact on the machine assessment of the education quality of uncontrollable disturbing factors, for example, man-made disasters, an unfavorable epidemiological situation or the introduction of martial law.

A promising way to reduce the multidimensionality impact of the recognition features dictionary is the use of machine learning methods, based on the results of which decisive rules are built within the framework of a geometric approach [20]. Among such methods, the information technologies of the support vector method [21, 22] deserve attention, but the functional effectiveness of the algorithms of this method significantly depends on the degree of recognition classes intersection in the feature space. This shortcoming is not present in machine learning methods, which are developed within the framework of the so-called information-extreme intelligent data analysis technology (IEIT) created at Sumy State University (Ukraine) [23–25], the methods of intelligent data analysis proposed within this technology are based on the principle maximizing the information capacity of the system in the process of its machine learning. The work [25] considered information-extreme machine learning based on a hierarchical data structure. But in the case of hierarchical machine learning DSS to assess the educational content quality, it is necessary to take into account such a feature as the presence of an ordered alphabet of recognition classes that characterize different levels of educational content quality when building a hierarchical data structure in the form of a decursive tree.

The purpose of the article is to develop an algorithm of hierarchical information-extreme machine learning of DSS and to verify it on the example of assessing compliance with modern requirements of the educational content of the educational and professional bachelor's degree program of the "Computer Science" specialty.

2 PROBLEM STATEMENT

Let's consider within the framework of IEIT the formalized setting of the information synthesis task of DSS to assess of the educational content quality and professional program of the specialty, which is used to train students of higher education at the graduation department.

It is necessary to build a hierarchical data structure in the form of a decursive binary tree $\{X^o_{h,s,m} \mid h=\overline{1,H}; s=\overline{1,S_h}; m_s=\overline{1,2}\}$ for the given alphabet $\{X^o_m \mid m=\overline{1,M}\}$ of recognition classes that characterize the levels of educational content quality assessment according to the appropriate system.

According to the concept of IEIT, transform the input training matrices of the recognition classes of each layer into the corresponding working binary matrices specified in the Hamming space. At the same time, let the DSS machine learning parameters be set, which, for example,

DOI 10.15588/1607-3274-2023-1-6

for recognition class X^o_{h,s,m_s} are represented in the form of a structured vector

$$g_{h,s} =< x^o_{h,s,m_s}, d_{h,s,m_s}, \delta_{h,s} > . \quad (1)$$

The number of optimization parameters in expression (1) sets the second depth level of information-extreme machine learning, since vector x_{h,s,m_s} depends on parametera $\delta_{h,s}$ depends on parameter.

Machine learning parameters are limited:

a) $d_{h,s,m_s} \in [0; d(x_{h,s,m_s} \oplus x_{h,s,c})]$;

b) $\delta_{h,s} \in [0; \delta_H / 2]$.

In the machine learning process of DSS, it is necessary to:

1) optimize the parameters of the vector (1) by finding the global maximum of the alphabetically averaged recognition classes of the s-th stratum h-th tier of the information criterion:

$$\overline{E}_{h,s} = \frac{1}{2}\sum_{m_s=1}^{2} \max_{G_E \cap G_d} E_{h,s,m_s}(d), \quad (2)$$

2) in the information-extreme machine learning process, build highly reliable decision rules based on the optimal (here and in the text in the informational sense) geometric parameters of the recognition classes containers;

3) during the operation of the DSS in the monitoring mode, determine whether the recognized feature vector belongs to the corresponding class from the given alphabet.

3 MATERIALS AND METHODS

Information-extreme machine learning of DSS was carried out using a hierarchical data structure in the form of a decursive binary tree [24]. With such a structure, the attribute from the top of the stratum of the higher tier is transferred to one of the vertices of the child stratum s, the lower tier. At the same time, the training matrix of the corresponding recognition class is considered as an attribute of the vertex.

Construction of a decursive binary tree was carried out according to the scheme:

1) alphabet $\{X^o_m \mid m = \overline{1, M}\}$ of ordered recognition classes is divided into two groups, which define two branches of the decursive tree, respectively;

2) as attributes of the vertices of the upper (first according to dendrographic classification) tier of the decursive tree, the training matrices of the boundaries for each of the recognition classes groups are selected;

3) the attributes of the strata of the upper tier are transferred to the vertices of the corresponding strata of the lower tier;

4) the strata of the lower tiers of each branch of the tree contain, in addition to the traning matrix transported from the upper tier, also the traning matrix of the nearest neighboring recognition class in its group;

5) the construction of the tree continues until the final strata are formed, which contain the training matrices of all recognition classes.

Thus, the binary decursive tree built according to the above scheme divides the given recognition classes alphabet into strata, each of which contains two nearest neighboring classes, which allows applying a linear information-extreme machine learning algorithm for each final stratum. At the same time, the construction of error-free decisive rules based on the training matrix is achieved by optimizing additional parameters of the DSS functioning.

The functional categorical model of the information-extreme machine learning of DSS is considered in the form of a directed graph, the edges of which are the mapping operators of the corresponding sets. At the same time, the input information description of the DSS is given by the structure

$$I =< P, T, \Omega, A, Y, H, Y^{|S|}, X^{|S|}; f_1, f_2, f_3, f_4 > .$$

At the same time, the Cartesian product $P \times T \times \Omega \times Z$ is considered as a source of information.

The functional categorical model of information-extreme machine learning of DSS with optimization of machine learning parameters (1) is shown in Fig. 1.

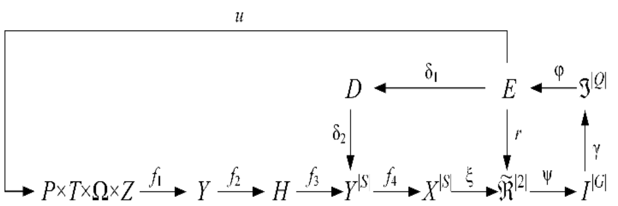


Figure 1 – Functional categorical model of machine learning of DSS

In Figure 1 operator ξ maps the structured binary feature vectors of training matrices $X^{|S|}$ to a fuzzy, in the general case, partition of $\tilde{\Re}^{|2|}_{h,s}$ recognition classes of each stratum of the decursive tree. The classification operator $\psi : \tilde{\Re}^{|2|} \to I^{|G|}$, where $I^{|G|}$ – is a set of G statistical hypotheses, checks the basic statistical hypothesis that the vector $x^{(j)}_{h,s,m_s}$ belongs to the fuzzy recognition class X^o_{h,s,m_s}. By evaluating statistical hypotheses, operator γ forms a set of accuracy characteristics $\Im^{|Q|}$, where $Q = G^2$ is the number of accuracy characteristics. Operator φ calculates a set of values of information criterion E, which is a functional of accuracy

DOI 10.15588/1607-3274-2023-1-6

characteristics. The optimization contour of the geometric parameters of partition $\Re^{|M|}$. The optimization contour of the geometric parameters of partition r, which at each step of machine learning restores the recognition classes containers in the radial basis of the Hamming feature space. The optimization contour of parameter $\delta_{h,s}$ of the control tolerances field includes the term set D, the elements of which are the allowable values of control tolerances for recognition features. At the same time, operator δ_1 changes the control tolerances, and operator δ_2 changes the quantization levels of the vector recognition features of the training matrices $Y^{|S|}$. Operator u regulates the machine learning process.

According to the functional category model (Fig. 1). We will present the algorithm of information-extreme machine learning of DSS in the form of a two-cycle iterative procedure for finding the global maximum of an information criterion in the working area of determining its function:

$$\delta^*_{h,s} = \arg\max_{G_\delta}\{\max_{G_E \cap G_d} \overline{E}_{h,s}(d)\}. \qquad (3)$$

The internal loop of the procedure (3) implements the basic machine learning algorithm, the purpose of which is to calculate at each step of learning the optimization information criterion $\overline{E}^{(k)}_{h,s}$ and search for the maximum value in the working area of determining its function.

Since the values of recognition features have the same measurement scale, a machine learning algorithm was implemented with parallel optimization of the control tolerances system, according to which control tolerances are changed for all recognition features simultaneously at each step of machine learning.

The input information for the machine learning algorithm is arrays of training matrices $\{y^{(j)}_{h,s,m_s}\}$ and a system of fields of normalized tolerances $\{\delta_{H,i}\}$ for recognition features, which sets the values range of the corresponding control tolerances.

Consider the optimization scheme of machine learning parameters for the recognition classes of the s-th stratum of the h-th tier of the decursive tree:

1) the recognition class counter of the s-th stratum of the h-th tier of the decursive tree is reset: $m_s := 0$;

2) $m_s := m_s + 1$;

3) the parameter change step counter is reset to zero $\delta_{h,s}$: $l := 0$;

4) the counter $l := 1+1$ is started and the lower $A_{HK,i}$ and upper $A_{BK,i}$ control tolerances are calculated for all recognition features:

$$\{A_{HK,i}[l] = y_{h,s,i} - l\};\ \{A_{BK,i}[l] = y_{h,s,i} + l\};\ i = \overline{1,N},$$

5) the basic algorithm of information-extreme machine learning is implemented, the tasks of which are to calculate at each step of learning the value of the information criterion (2) and search for the maximum value of the criterion in the working area of determining its function;

6) the optimal radius of the recognition class container X^o_{h,s,m_s} is determined

$$d^*_{h,s,m_s} = \arg\max_{G_E \cap G_d} E_{h,s,m_s}(d),$$

7) if $l \le \delta_H/2$, then point 4 is fulfilled, otherwise – point 8;

8) if $m_s \le 2$, then point 2 is fulfilled, otherwise point 9;

9) the average value of information criterion $\overline{E}_{h,s}$ and the optimal parameter of the control tolerances field $\delta^*_{h,s}$ are calculated;

10) STOP.

Since the machine learning of DSS takes place according to the hierarchical data structure, an algorithm was implemented, the main stages of which are:

1) zeroing of the data structure tiers counter: $h := 0$;

2) initialization of the data structure tier counter: $h := h + 1$;

3) zeroing of the sratum counter: $s := 0$;

4) initialization of the tier counter: $s := s + 1$;

5) for each s-th stratum of the h-th tier of the decursive tree, an information-extreme algorithm of machine learning with parallel optimization of control tolerances for recognition features is implemented, which calculates:

a) the maximum value of the information criterion $\overline{E}_{\bar{h},s}$ averaged for the stratum;

b) averaged for the stratum $\{x^*_{h,s,m_s}\}$ recognition classes of the s-th stratum of the h-th tier;

c) optimal radii of $\{d^*_{h,s,m_s}\}$ recognition classes of the s-th stratum of the h-th tier;

d) optimal paramete $\delta^*_{h,s}$ of the control tolerances field for recognition classes features of the s-th stratum of the h-th tier of the decursive tree;

6) if $s \le S_h$, where S_h is the number of strata on the h-th tier, then point 4 is fulfilled, otherwise point 7;

7) if $h \le h_{max}$, where h_{max} is the number of the decursive tree levels, then point 2 is fulfilled, otherwise point 8;

8) the maximum value of information criterion $\overline{E}_H$ averaged over all strata of the decursive tree is calculated;

DOI 10.15588/1607-3274-2023-1-6

9) a structured set of decisive rules is formed for all recognition classes;

10) STOP.

As a criterion for optimizing machine learning parameters, the modified Kullback information measure was considered, which for two-alternative solutions and equally likely hypotheses has the form [25]

$$E_{h,s,m_s}(d)=\frac{\left|n-(K^{(1)}_{h,sw,m_s}(d)+K^{(2)}_{h,s,m_s}(d))\right|}{n}\times$$
$$\times\log_2\left\{\frac{2n+10^{-p}-\left[(K^{(1)}_{h,sw,m_s}(d)+K^{(2)}_{h,s,m_s}(d)\right]}{\left[K^{(1)}_{h,s,m_s}(d)+K^{(2)}_{h,s,m_s}(d)\right]+10^{-p}}\right\}. \qquad (4)$$

Based on the optimal geometric parameters of the recognition classes containers obtained in the machine learning process, decisive rules are constructed, which will be presented in the form

$$(\forall X^{o}_{m,h,s}\in\Re^{|M|})(\forall x^{(j)}\in\Re^{|M|})\{if\ [(\mu_{m_s}>0)\ \&$$
$$\&\,(\mu_m=\max_{\{m\}}\{\mu_{m_s}\mid m_s=\overline{1,2}\}]\ then\ x^{(j)}\in X^{o}_{h,s,m_s} \qquad (5)$$
$$else\ x^{(j)}\notin X^{o}_{h,s,m_s}\}.$$

In expression (5), the membership function for the hyperspherical container of recognition class X^{o}_{h,s,m_s} is determined by the formula

$$\mu_{m_s}=1-\frac{d(x^{*}_{h,s,m_s}\oplus x^{(j)})}{d^{*}_{h,s,m_s}}. \qquad (6)$$

In expression (6), the code distance between the optimal averaged feature vector x^{*}_{h,s,m_s} and the recognized feature vector $x^{(j)}$, is denoted as $d(x^{*}_{h,s,m_s}\oplus x^{(j)})$.

Thus, the geometric decisive rules (5) built in the information-extreme machine learning process of DSS are characterized by low computational complexity and are practically invariant to the multidimensionality of the recognition features dictionary.

4 EXPERIMENTS

The implementation of the above information-extreme machine learning algorithm of DSS was carried out on the example of assessing the compliance with the labor market requirements of the educational content of the bachelor's level educational and professional program in the specialty "Computer Science", which is taught to students of Sumy State University. The input training matrix was formed by simulating the evaluations of the educational disciplines thematic modules, which were randomly generated according to the normal distribution of probabilities and displayed on a stobal scale. Since the thematic modules were evaluated according to the European rating scale, the stobal scale was previously divided into six fuzzy intervals, which specified the areas of the recognition classes:

1) recognition class X^{o}_{A} included feature vectors that fell within the range from 88 to 100 points;

2) for recognition class X^{o}_{B} – from 81 to 91 points;

3) for recognition class X^{o}_{C} – from 74 to 84 points;

4) for recognition class X^{o}_{D} – from 67 to 77 points;

5) for recognition class X^{o}_{E} – from 59 to 69 points;

6) for recognition class X^{o}_{F} – from 50 to 62 points.

The simulation training matrix for each recognition class consisted of 40 syllabuses of educational disciplines of the educational and professional program structured by thematic modules. Each vector consisted of 144 features, the number of which was equal to the thematic modules number. In addition, according to the educational and professional program of the specialty, educational disciplines were divided into seven blocks, which included: general scientific, fundamental, design, technological organizational, humanitarian and selective educational disciplines.

For a given recognition classes alphabet, a hierarchical data structure in the form of a decursive tree was built according to the above algorithm (Fig. 2).

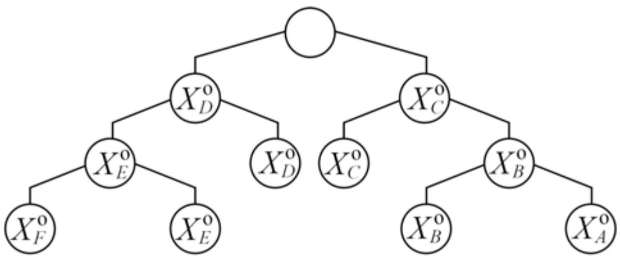


Figure 2 – Hierarchical data structure in the form of a decursive binary tree

According to the decursive tree (Fig. 2), information-extreme machine learning was implemented according to procedure (3). Optimization of machine learning parameters specified by vector (1) was carried out for the two nearest neighboring recognition classes of each stratum. Based on the optimal geometric parameters of the recognition class containers obtained in the machine learning process, decisive rules (5) were constructed. Since each recognition class belongs to two strata, according to the minimum-distance principle of pattern recognition theory, the geometric parameters of the recognition class whose container had the minimum optimal radius were taken as optimal.

According to the machine learning results of DSS, the values of the averaged information criteria (4) were

DOI 10.15588/1607-3274-2023-1-6

analyzed for each stratum in order to determine the need to increase the depth of machine learning by optimizing additional parameters of the DSS operation, including the formation parameters of the input information description.

5 RESULTS

In the programmatic implementation process of information-extreme machine learning of DSS according to the iterative procedure (3), the optimal parameters of the recognition classes for each of the strata of the decursive binary tree were determined (Fig. 2). As an example, Figure 3 shows the dependency graph of the averaged information criterion (4) for the stratum of the first (upper according to dendrographic classification) tier of the decursive tree, obtained by the machine learning results.

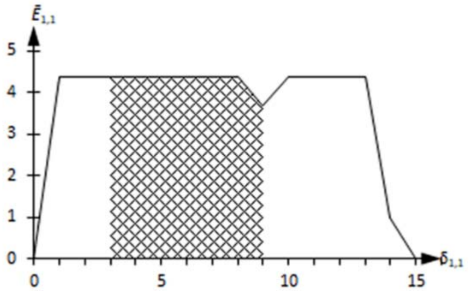


Figure 3 – Dependency graph of the information criterion on the parameter of the control tolerances field for the strata of the first tier

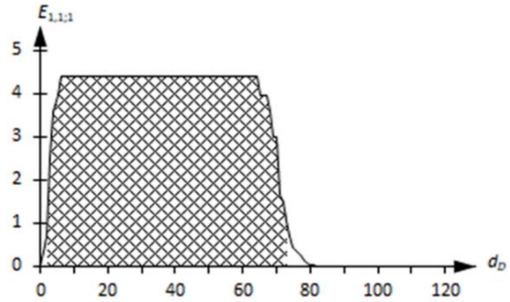
In Fig. 3 double hatching indicates the working area of the definition of the criterion function (4), in which the errors of the first and second kind are less than the first and second reliabilities, respectively. The analysis of Figure 3 shows that the global maximum of the optimization criterion is reached in the area of the plateau-type graph. In this case, the optimal parameter is determined under the condition of the minimum value of the so-called coefficient of recognition classes intersection [23]

$$\eta = \frac{d^*_{h,s,m_s}}{d(x^*_{h,s,1} \oplus x^*_{h,s,2})} \to \min_{\{m_s\}}. \qquad (7)$$

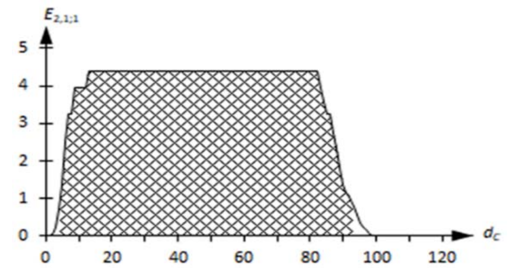
In expression (7), the intercenter code distance between the recognition classes and the $X^o_{h,s,1}$ and $X^o_{h,s,2}$ s-th stratum of the h-th tier of the decursive tree is denoted as $d(x^*_{h,s,1} \oplus x^*_{h,s,2})$.

After checking the fulfillment of condition (7), the optimal values interval of the control tolerance field parameter $\delta^*_{1,1} = 5$ was determined, from which the value $\overline{E}^*_{1,1} = 4.50$ was selected. At the same time, the averaged information optimization criterion reaches its maximum value at parameters $n = 40$ and $p = 2$ specified in formula (4).

Since decisive rules (5) are built within the framework of a geometric approach, they require knowledge of the optimal geometric parameters of recognition class containers obtained in the machine learning process. Figure 4 shows dependency graphs of criterion (4) on the radii of the recognition classes containers of the first tier of the decursive tree (Fig. 2).



a



b

Figure 4 – Dependency graphs of the information criterion (4) on the radii of the recognition classes containers of the first stratum of the second tier: a – class X^o_D; b – class X^o_C

The optimal radii of the recognition classes containers of the first stratum according to the condition (7) are equal to any number from the intervals: $d^*_D \in [38,39]$ (hereafter in code units) for class X^o_D and $d^*_C \in [24,27]$ for class X^o_C. According to the minimum-distance principle of the pattern recognition theory, the minimum radii were chosen: $d^*_D = 38$ and $d^*_C = 24$.

Fig. 5 shows dependency graph of the averaged information criterion (4) on parameter δ, obtained during machine learning based on the training matrix of recognition classes of the first stratum of the second tier of the decursive tree (Fig. 2).

DOI 10.15588/1607-3274-2023-1-6

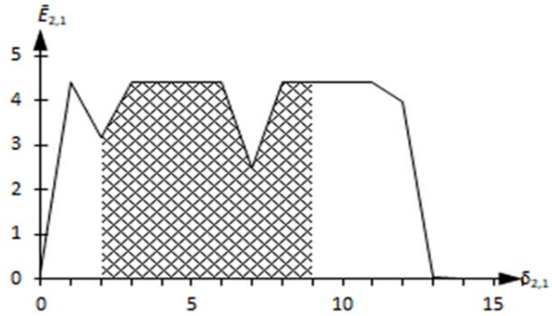


Figure 5 – Dependency graph of the information optimal parameter of machine learning of the criterion on the parameter of the control tolerances field for the first stratum of the second tier

When condition (7) is fulfilled, the optimal parameter of the control tolerances field for recognition classes of the first stratum of the second tier of the decursive tree (Fig. 2) is equal to $\delta^*_{2,1} = 4$ at the maximum value of the criterion $\overline{E}^*_{2,1} = 4.50$.

Figure 6 shows dependency graphs of criterion (4) on the radii of the recognition classes containers of the first stratum of the second tier of the decursive tree (Fig. 2).

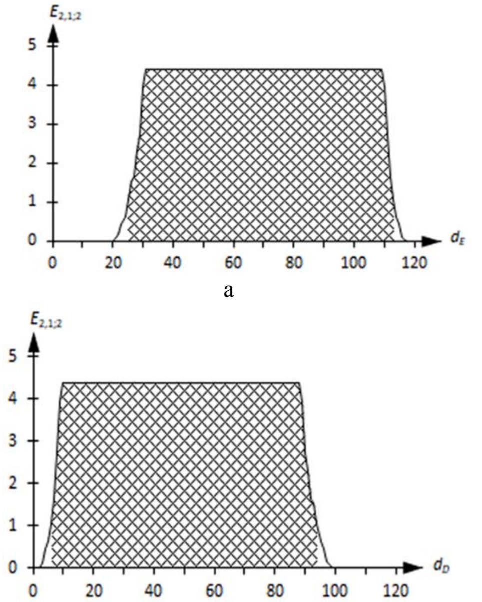


Figure 6 – Dependency graphs of the information criterion (4) on the radii of the recognition classes containers of the first stratum of the second tier: a – class $X^o_{2,1,1}$ (X^o_E); b – class $X^o_{2,1,2}(X^o_C)$

The optimal radii of the recognition classes containers of the first stratum of the second tier according to condition (7) are equal to $d^*_{2,1,1} = 46$ for class X^o_E and $d^*_{2,1,2} = 32$ for class X^o_D. Similarly, the optimal machine learning parameters for other recognition classes were determined. Table 1 shows the results of information-extreme machine learning of DSS according to the hierarchical structure (Fig. 2).

Table 1 – Optimal machine learning parameters

Strate number (h,s)	$\delta^*_{h,s}$	$d^*_{h,s,,1}$	$d^*_{h,s,,2}$
1,1	$\delta^*_{1,1} = 5$	$d^*_{1,1,1} = d^*_D = 38$	$d^*_{1,1,,2} = d^*_C = 24$
2,1	$\delta^*_{2,1} = 4$	$d^*_{h,s,,1} = d^*_E = 46$	$d^*_{2,1,2} = d^*_D = 32$
2,2	$\delta^*_{2,2} = 5$	$d^*_{2,2,1} = d^*_C = 52$	$d^*_{2,2,2} = d^*_B = 16$
3,1	$\delta^*_{3,1} = 5$	$d^*_{3,1,1} = d^*_F = 64$	$d^*_{3,2,2} = d^*_E = 12$
3,2	$\delta^*_{3,2} = 4$	$d^*_{3,2,1} = d^*_B = 57$	$d^*_{3,2,2} = d^*_A = 16$

Comparing in Table 1 the radii of the recognition classes containers X^o_D for the strata of the upper and lower tiers, a smaller value of $d^*_D = 32$, should be taken as the optimal one, since class X^o_E is the nearest neighbor for it. Similarly, for recognition class X^o_C the permissible optimal radius is equal to $d^*_C = 24$, for class X^o_B – $d^*_B = 16$ and for class X^o_E – $d^*_E = 12$.

When calculating the membership function (6), vectors x^*_{h,s,m_s} were determined by optimal control tolerances for recognition features.

6 DISCUSSION

The advantages of building a hierarchical data structure in the form of a decursive binary tree are shown on the example of information-extreme machine learning of DSS for adapting the educational content of the educational and professional program of the second (bachelor) level to the labor market requirements. Thanks to the proposed structure, it was possible to automatically divide a set of recognition classes into pairs of nearest neighbors. In addition, the a priori structuring of the recognition classes due to the corresponding knowledge level assessment system is taken into account. As a result, there is no need to form a variational series of recognition classes, which is an advantage of the training required to build a decursive binary tree. The use of a hierarchical data structure in the form of a decursive binary tree, in addition, allows for each stratum to implement information-extreme machine learning according to a linear algorithm with the required level of depth. Unlike neuro-like structures in IEIT methods, the depth level of information-extreme machine learning is determined by the number of optimization parameters that affect the functional efficiency of the intelligent system. This approach corresponds to a greater extent to the mechanism of making classification decisions by natural intelligence. It is also justified to use information criterion (4) in classification tasks as an optimization criterion,

DOI 10.15588/1607-3274-2023-1-6

which in the logical and epistemological aspect is considered as a measure of the diversity of recognition classes.

The analysis of the obtained results of the information-extreme machine learning of DSS for adapting the educational content to the labor market requirements shows that for each layer of the decursive tree it was possible to build error-free decisive rules according to the training matrix. At the same time, in accordance with the minimum-distance principle of pattern recognition theory, the radii of the recognition classes containers determined relative to the nearest neighboring class were considered optimal.

The specifics of the development of algorithmic and software of the DSS for adapting educational content to the labor market requirements is the use of simulation modeling for the formation of the input traning matrix. Such an approach is due to the large material and time costs of obtaining the results of the respondents survey in the absence of the DSS. However in the future, when the DSS is functioning in the monitoring mode, as archival representative data accumulates, it becomes possible to retrain the system based on real data. Further research will be aimed at expanding the functional capabilities of the DSS in order to evaluate blocks and individual academic disciplines. As a result, graduation departments will be able to quickly adjust the educational content to modern requirements.

CONCLUSIONS

The actual problem of improving the quality of education is being solved by building an information and communication system for adapting the educational content of the graduation department of the university to the requirements of the labor market.

The scientific novelty of the obtained results is that the method of information-extreme machine learning based on the hierarchical data structure in the form of a decursive binary tree is proposed for the first time. The method automatically divides the alphabet of high-power recognition classes into pairs of nearest neighbors. It performs two-class machine learning of DSS, ensuring high reliability with a minimum depth of machine learning. As a result, error-free decision rules based on the training sample were built in the monitoring mode of DSS. It allows the total probability of correct classification decisions to be close to the maximum limit.

The practical significance of the obtained results is that the developed DSS software allows for quickly adjusting the educational content of the graduation department of the university, taking into account the requirements of the labor market, with minimal material and time costs.

Prospects for further research consist in simplifying the formation of the input information description of the DSS by automatically reading the thematic modules from the syllabuses of the educational disciplines of the specialty of the corresponding level of training of higher education applicants.

ACKNOWLEDGEMENTS

The work was carried out in the laboratory of intelligent systems of the Department of Computer Sciences of Sumy State University with the funds of the state budget scientific research theme “Information technology of autonomous navigation of unmanned aerial vehicles by terrestrial natural and infrastructural landmarks” (№ DR 0122U000786).

REFERENCES

1. Bates T. National strategies for e-learning in post-secondary education and training. UNESCO, 2001, 132 p.
2. Sharable Content Object Reference Model (SCORM) Version 1.2 / Advanced Distributed Learning, 2001, 170 p.
3. Agapova M. O. Pro problemu optymizatsii navchalnoho materialu, *Teorija i praktyka upravlinnja social'nymy systemamy,* 2012, № 1, pp. 34–39.
4. Nagy A., Bruck P., Buchholz A., Karssen Z., Zerfass A. eds E-Content: Technologies and Perspectives for the European Market. Berlin, Springer-Verlag, 2005, pp. 79–96.
5. Dovgyj S. O., Bidyuk P. I., Trofymchuk O. M. Decision support systems based on statistical and probabilistic procedures. Kyiv, Logos, 2014, 419 p.
6. Boyd S., Parikh N., Chu E. et al. Distributed optimization and statistical leaming via the altemating direction method of multipliers, *Foundations and Trends R in Machine Leaming,* 2011, Vol. 3(1), pp. 1–122.
7. Xu G., Zong Y., Yang Y. Z. Applied Data Mining, CRC Press, 2013, 284 p.
8. Dua S., Du X. Data Mining and Machine Learning in Cybersecurity, Auerbach Publications, 2011, 256 p.
9. Sklyar V. V., Kharchenko V. S. Case-Based and Project-Based Methods for Effective E-learning in ICT Safety and Security, *ICTERI Workshops,* 2020, pp. 157–170.
10. Stuart Russell, Peter Norvig Intelligence artificielle, 2e éd. (French translation of Artificial Intelligence: A Modern Approach, second edition). Paris, Pearson Education France, 2006.
11. Sivanandam S. N., Sumathi S., Deepa S. N. Introduction to fuzzy logic using MATLAB. Springer, 2007, 430 p.
12. Watt J., Borhani R., Katsaggelos A. Machine Learning Refined: Foundations, Algorithms, and Applications. 2nd Edition, Cambridge University Press, 2020, 594 p.
13. Wu S.-J., Gebraeel N., Lawley M. A. et al. A neural network integrated decision support system for condition-based optimal predictive maintenance policy. Systems, *Man and Cybernetics, Part A: Systems and Humans. IEEE Transactions*, 2007, Vol. 37, №. 2, pp. 226–236.
14. Radovic M., Adarkwa O., Wang Q. Object recognition in aerial images using convolutional neural networks, *J. Imaging,* 2017, Vol. 3, № 2. DOI: 10.3390/jimaging3020021.
15. Alpaydin Ethem. Introduction to Machine Learning. London, The MIT Press, 2010, 110 p.
16. Moskalenko V. V., Korobov A. G. Information-extreme algorithm of the system for recognition of objects on the terrain with optimization parameter feature extraction, *Radio Electronics, Computer Science, Control*, 2017, №2, pp. 38–45.
17. Shi Y., Mizumoto M. An improvement of neuro-fuzzy learning algorithm for tuning fuzzy rules, *Fuzzy sets and systems*, 2001, Vol. 118, № 2, pp. 339–350.
18. Subbotin S. A. The neuro-fuzzy network synthesis and simplification on precedents in problems of diagnosis and

DOI 10.15588/1607-3274-2023-1-6

pattern recognition, *Optical Memory and Neural Networks (Information Optics),* 2013, Vol. 22, № 2, pp. 97–103. DOI: 10.3103/s1060992x13020082.
19. Efendigil T., Önüt S., Kahraman C. A decision support system for demand forecasting with artificial neural networks and neuro-fuzzy models: a comparative analysis, *Expert Systems with Applications*, 2009, Vol. 36, № 3, pp. 6697–6707.
20. Rahimi A., Recht B. Random features for large-scale kernel machines, *Advancesin Neural Information Processing Systems,* 2008, pp. 1177–1184.
21. Chapelle O. Training a support vector machine in the primal, *Neural Computation*, 2007, Vol. 19, №5, pp. 1155–1178.
22. Ivanciuc O. Applications of Support Vector Machines in Chemistry, *Reviews in Computational Chemistry,* 2007, Vol. 23, pp. 291–400.
23. Dovbysh A. S., Budnyk M. M., Piatachenko V. Yu. et al. Information-Extreme Machine Learning of On-Board Vehicle Recognition System, *Cybernetics and Systems Analysis*, 2020, Vol 56(4), pp. 534–543. DOI:10.1007/s10559-020-00269-y
24. Dovbysh A. S., Budnyk M. M., Moskalenko V. V. Information-Extreme Algorithm for Optimizing Parameters of Hyperellipsoidal Containers of Recognition Classes, *Journal of automation and information sciences.* New York, Begell House Inc., 2012, Vol. 44(10), pp. 35–44. DOI: 10.1615/jautomatInfScien.v44.i10.30.
25. Shelehov I. V., Barchenko N. L., Prylepa D. V. et al. Information-extreme machine training system of functional diagnosis system with hierarchical data structure, *Radio Electronics, Computer Science, Control,* 2022, № 2, pp. 189-200. DOI:10.15588/1607-3274-2022-18.

Received 25.11.2022.
Accepted 22.01.2023.

УДК 681.518:004.93

ІНФОРМАЦІЙНО-ЕКСТРЕМАЛЬНЕ МАШИННЕ НАВЧАННЯ СИСТЕМИ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ ДЛЯ АДАПТАЦІЇ НАВЧАЛЬНОГО КОНТЕНТУ ДО ВИМОГ РИНКУ ПРАЦІ

Шелехов І. В. – доцент, канд. техн. наук, доцент кафедри комп’ютерних наук, Сумський державний університет, Суми, Україна.
Прилепа Д. В. – асистент кафедри комп’ютерних наук, Сумський державний університет, Суми, Україна.
Хібовська Ю. О. – аспірант кафедри комп’ютерних наук, Сумський державний університет, Суми, Україна;
Отрощенко М. С. – аспірант кафедри комп’ютерних наук, Сумський державний університет, Суми, Україна.

АНОТАЦІЯ

Актуальність. Розв’язана актуальна задача підвищення функціональної ефективності машинного навчання системи підтримки прийняття рішень (СППР) для оцінки відповідності сучасним вимогам контенту навчальних дисциплін випускової кафедри за результатами опитування роботодавців.

Мета. Підвищення функціональної ефективності машинного навчання СППР для оцінки відповідності сучасним вимогам контенту навчальних дисциплін освітньо-професійної програми спеціальності першого (бакалаврського) рівня на основі машинного навчання та розпізнавання образів.

Метод. Запропоновано метод інформаційно-екстремального машинного навчання СППР для адаптації навчального контенту випускової кафедри до вимог ринку праці. Ідея методу полягає у максимізації інформаційної спроможності СППР в процесі машинного навчання, що дозволяє в режимі моніторингу гарантувати високу повну ймовірність прийняття правильних класифікаційних рішень. Метод розроблено в рамках функціонального підходу до моделювання когнітивних процесів природнього інтелекту, що дозволяє надати СППР гнучкість при перенавчанні системи через збільшення потужності алфавіту класів розпізнавання. В основу методу покладено принцип максимізації кількості інформації в процесі машинного навчання. Як критерій оптимізації параметрів машинного навчання розглядається модифікована інформаційна міра Кульбака, яка є функціоналом від точнісних характеристик класифікаційних рішень. Згідно із запропонованою функціональною категорійною моделлю розроблено алгоритм інформаційно-екстремального машинного навчання за ієрархічною структурою даних у вигляді бінарного декурсивного дерева. Застосування такої структури даних дозволяє автоматично розбивати велику кількість класів розпізнавання на пари найближчих сусідів, для яких оптимізація параметрів машинного навчання здійснюється за лінійним алгоритмом необхідної глибини. Як параметри оптимізації розглядалися геометричні параметри гіперсферичних контейнерів класів розпізнавання, які в процесі машинного навчання відновлювалися в радіальному базисі бінарного простору ознак Геммінга. При цьому вхідна навчальна матриця трансформувалася в робочу бінарну навчальну матрицю, яка змінювалася в процесі машинного навчання шляхом допустимих перетворень з метою адаптації вхідного інформаційного опису СППР до максимальної достовірності класифікаційних рішень.

Результати. Розроблено інформаційне, алгоритмічне і програмне забезпечення СППР для оцінки якості навчального контенту за результатами машинного аналізу відповідей респондентів. В рамках геометричного підходу за результатами інформаційно-екстремального машинного навчання за ієрархічною структурою даних у вигляді бінарного декурсивного дерева побудовано високо достовірні вирішальні правила, практично інваріантні до багато вимірності простору ознак розпізнавання. Досліджено вплив параметрів машинного навчання на функціональну ефективність машинного навчання СППР на прикладі оцінки навчального контенту освітньо-професійної програми бакалаврського рівня спеціальності «Комп’ютерні науки».

Висновки. Результати комп’ютерного моделювання підтверджують високу функціональну ефективність запропонованого методу інформаційно-екстремального ієрархічного машинного навчання і можуть бути рекомендовані до

DOI 10.15588/1607-3274-2023-1-6

практичного використання у закладах вищої освіти для оцінки відповідності сучасним вимогам навчального контенту випускових кафедр

КЛЮЧОВІ СЛОВА: інформаційно-екстремальне машинне навчання, функціональна категорійна модель, інформаційний критерій, ієрархічна структура даних, декурсивне дерево, навчальний контент.

ЛІТЕРАТУРА

1. Bates T. National strategies for e-learning in post-secondary education and training / Bates Tony – UNESCO, 2001. – 132 p.
2. Sharable Content Object Reference Model (SCORM) Version 1.2 / Advanced Distributed Learning, 2001. – 55 p.
3. Агарова М. О. Про проблему оптимізації навчального матеріалу / М. О. Агарова // Теорія і практика управління соціальними системами. – 2012. – № 1. – С. 34–39.
4. E-Content: Technologies and Perspectives for the European Market / [eds: A. Nagy, P. Bruck, A. Buchholz, Z. Karssen, A. Zerfass]. – Berlin : Springer-Verlag, 2005. – P. 79–96.
5. Dovgyj S. O. Decision support systems based on statistical and probabilistic procedures / S. O. Dovgyj, P. I. Bidyuk, O. M. Trofymchuk. – Kyiv : Logos, 2014. – 419 p.
6. Distributed optimization and statistical leaming via the altemating direction method of multipliers / [S. Boyd, N. Parikh, E. Chu et al.] // Foundations and Trends R in Machine Leaming. – 2011. – Vol. 3(1). – P. 1–122.
7. Xu G. Applied Data Mining / G. Xu, Y. Zong, Y. Z. Yang. – CRC Press, 2013. – 284 p.
8. Dua S. Data Mining and Machine Learning in Cybersecurity / S. Dua, X. Du. – Auerbach Publications, 2011. – 256 p.
9. Sklyar V. V. Case-Based and Project-Based Methods for Effective E-learning in ICT Safety and Security / V. V. Sklyar, V. S. Kharchenko // ICTERI Workshops. – 2020: – P. 157–170.
10. Stuart Russell. Intelligence artificielle, 2e éd. (French translation of Artificial Intelligence: A Modern Approach, second edition) / Stuart Russell, Peter Norvig – Paris: Pearson Education France, 2006.
11. Sivanandam S. N. Introduction to fuzzy logic using MATLAB / S. N. Sivanandam, S. Sumathi, S. N. Deepa – Springer, 2007. – 430 p.
12. Watt J. Machine Learning Refined: Foundations, Algorithms, and Applications. – 2nd Edition / J. Watt, R. Borhani, A. Katsaggelos. – Cambridge University Press, 2020. – 594 p.
13. A neural network integrated decision support system for condition-based optimal predictive maintenance policy. Systems / [S.-J. Wu, N. Gebraeel, M. A. Lawley et al.] // Man and Cybernetics, Part A: Systems and Humans. IEEE Transactions. – 2007. – Vol. 37, № 2. – P. 226–236.
14. Radovic M Object recognition in aerial images using convolutional neural networks / M. Radovic, O. Adarkwa, Q. Wang // J. Imaging. – 2017. – Vol. 3, № 2. DOI: 10.3390/jimaging3020021.
15. Alpaydin Ethem. Introduction to Machine Learning / Ethem Alpaydin – London : The MIT Press, 2010. – 110 p.
16. Moskalenko V. V. Information-extreme algorithm of the system for recognition of objects on the terrain with optimization parameter feature extraction / V. V. Moskalenko, A. G. Korobov, // Radio Electronics, Computer Science, Control, 2017. – №2. – P. 38–45.
17. Shi Y. An improvement of neuro-fuzzy learning algorithm for tuning fuzzy rules / Y. Shi, M. Mizumoto // Fuzzy sets and systems. – 2001. – Vol. 118, № 2. – P. 339–350.
18. Subbotin S. A. The neuro-fuzzy network synthesis and simplification on precedents in problems of diagnosis and pattern recognition / S. A. Subbotin // Optical Memory and Neural Networks (Information Optics). – 2013. – Vol. 22, № 2. – P. 97–103. DOI: 10.3103/s1060992x13020082.
19. Efendigil T. A decision support system for demand forecasting with artificial neural networks and neuro-fuzzy models: a comparative analysis / T. Efendigil, S. Önüt, C. Kahraman // Expert Systems with Applications. – 2009. – Vol. 36, № 3. – P. 6697–6707.
20. Rahimi A. Random features for large-scale kernel machines / A. Rahimi, B. Recht // Advaпcesiп Neural Iпformation Processing Systems. – 2008. – P. 1177–1184.
21. Chapelle O. Training a support vector machine in the primal / O. Chapelle // Neural Computation. – 2007. – Vol. 19, №5. – P.1155–1178.
22. Ivanciuc O. Applications of Support Vector Machines in Chemistry / O. Ivanciuc // Reviews in Computational Chemistry. – 2007. – Vol. 23. – P. 291–400.
23. Information-Extreme Machine Learning of On-Board Vehicle Recognition System / [A. S. Dovbysh, M. M. Budnyk, V. Yu Piatachenko et al.] // Cybernetics and Systems Analysis. – 2020. – Vol 56(4). – P 534–543. DOI:10.1007/s10559-020-00269-y
24. Dovbysh A. S. Information-Extreme Algorithm for Optimizing Parameters of Hyperellipsoidal Containers of Recognition Classes / A. S. Dovbysh, M. M. Budnyk, V. V. Moskalenko // Journal of automation and information sciences.− New York: Begell House Inc. – 2012. – Vol. 44(10) – P. 35–44. DOI: 10.1615/jautomatInfScien.v44.i10.30.
25. Information-extreme machine training system of functional diagnosis system with hierarchical data structure / [I. V. Shelehov, N. L. Barchenko, D. V. Prylepa et al.] // Radio Electronics, Computer Science, Control. – 2022. – № 2. – P. 189–200. DOI:10.15588/1607-3274-2022-18.

DOI 10.15588/1607-3274-2023-1-6