The American Journal of Engineering and Technology
(ISSN – 2689-0984
Published: April 22, 2021 | **Pages:** 7-16
Doi: https://doi.org/10.37547/tajet/Volume03Issue04-02

IMPACT FACTOR
2021: 5. 705
OCLC - 1121105677

About One Algorithm For Generating Reference Data In Pattern Recognition

Gulomjon Primovich Juraev
Independent Researcher, Information Technologies Center, Tashkent University Of Information Technologies Named After Muhammad Al-Khwarizmi, Uzbekistan

ABSTRACT

In this paper the issues like preprocessing of medical data, reclassification of the training sets and determining the importance of classes, formation of reference tables, selection of an informative features set that differentiate between class objects, formed by medical professionals are discussed and solved. Mainly in the most studied references [5-8, 11-13] the Fisher's criterion is used to obtain solutions to problems/tasks. Also for solving problems, the algorithms for an estimate calculation as well as the related software programs are used. For all cases, algorithms and software programs are suggested.

The study consists of two important steps. The first step is to build a reference table, based on the importance of the features and objects as well as their contribution to the classes [1-4, 9, 10]; the second step is concerned with the choice of the most useful characteristic features set to be investigated. This corresponds to solving the issue of selection of set of informative features from a given table, their visualization, and the determination of the contribution of the features set to the formation of classes [1-13].

KEYWORDS

Reference data; classification; pattern recognition; algorithms for an estimate calculation; preprocessing of medical data.

The American Journal of Engineering and Technology
(ISSN – 2689-0984
Published: April 22, 2021 | **Pages:** 7-16
Doi: https://doi.org/10.37547/tajet/Volume03Issue04-02

IMPACT FACTOR
2021: 5. 705
OCLC - 1121105677

INTRODUCTION

Modern analysis of biomedical data requires feature selection methods that can be applied to large-scale feature spaces, function in noisy tasks, discover complex association patterns, flexibly adapted to different problem areas and data types (for example, genetic variants, gene expression and clinical data) and can be computationally computed. To this end, in [13], a set of algorithms for selecting elements in the filter style based on the Relief algorithm, that is, Relief (RBA) -based algorithms, is considered in the study. [This paper is mainly considered about analyzing algorithms for selecting elements in the filter style Relief based algorithm (RBA).] RBA is being introduced and expanded in an open source environment called ReBATE (a learning environment based on bump algorithms). A comprehensive study of genetic modeling is offered which compares existing RBAs, proposed by RBA under the name MultiSURF and other established methods for selecting features, for a number of problems.

In [12], the problems of diagnosis and treatment of cardiovascular diseases, which are often encountered when making diagnostic decisions in the processing of medical data, are considered. The classification of heart diseases and the identification of informative features are resolved on the basis of algorithms for an estimate calculation. The main purpose of the study is to solve such issues as the construction of inter-object remoteness in an informative features set that distinguish objects of diagnostic classes, the allocation of a set of features characterizing the mutual differences of objects, as well as the identification of the proximity function while diagnosing an unknown object [17-22]. Identification of the level of significance or presentation, which are the main stages of the algorithms for an estimate calculation relative to classes [22]. An algorithm for diagnosing an unknown object in the space of informative features is proposed. The proposed theoretical ideas were confirmed in practice. In addition, the decision-making rules in this space and their software were developed [21, 22].

Multidimensional data analysis is a challenge for researchers and engineers in the field of machine learning and data mining. Selection of functions provides an effective way to solve this problem by removing unnecessary and redundant data that can reduce computation time, improve training accuracy and facilitate understanding of the training model or data. In [14] several commonly used evaluation indicators were studied for selecting features, and then methods for selecting controlled, uncontrolled, and semi-serviced features that are widely used in machine learning problems, such as classification and clustering were investigated.

When a feature-object set contains highly correlated features, the SVM-RFE ranking criterion will be biased, making it difficult to apply the SVM-RFE to gas sensor data. The article [15] considers linear and nonlinear SVM-RFE algorithms. After investigating the correlation bias, an improved SVM-RFE + CBR algorithm is proposed that includes a strategy for reducing correlation bias (CBR) in the feature elimination procedure. The ensemble method is additionally studied to increase the stability of the proposed method.

The selection of features is an important stage of data pre-processing, which increases the

The American Journal of Engineering and Technology
(ISSN – 2689-0984
Published: April 22, 2021 | **Pages:** 7-16
Doi: https://doi.org/10.37547/tajet/Volume03Issue04-02

IMPACT FACTOR
2021: 5. 705
OCLC - 1121105677

performance of training algorithms by removing unnecessary and redundant features. In [16] a method for feature selection using the Forest Optimization Algorithm (FSFOA) is proposed. To select more informative features from data sets, the FSFOA method is proposed and implemented on several real data sets and compared with several other methods, including HGAFS, PSO and SVM-FuzCoc. Experimental results show that FSFOA can improve the classification accuracy of classifiers in some selected data sets.

Feature selection is an important task in data mining and machine learning applications that eliminate unnecessary and redundant functions and increase learning productivity. In many real-world applications, collecting marked data is difficult, while plentiful unlabeled data is easily accessible. This allows researchers to develop methods for selecting objects that use both marked and unlabeled data to assess the relevance of the object. However, to date, no comprehensive survey has been conducted covering the methods of selecting objects under observation. In [23] the objects selection methods under observation are completely studied and two taxonomies of these methods are presented on the basis of two different points of view, which represent the hierarchical structure of objects selection methods under observation. The first point of view is based on the taxonomy of feature selection methods, and the second is based on the taxonomy of observation methods. This question can be useful for the researcher to get a deep background during observation and choose the right method for objects selection based on their hierarchical structure.

In this article, issues such as preprocessing of medical data generated by medical professionals in the intellectual analysis of medical data, reclassification of training sets and identification of importance level of classes, the formation of reference tables, and the selection of informative features that differentiate between class objects are solved based on Fisher criteria using the algorithms for an estimate calculation.

MAIN PART

In this section:

1) The data includes preliminary data preprocessing issues and are given in the case of medical issues. The first of the 4 issues in this section is to determine the feasibility of the characteristic features of the objects classified, the second is to convert characters from classed objects into continuous numbers from 0 to 1, the third is the formation of a reference table by determining whether a class object belongs to its class or to another class, the fourth focuses on the definition of an informative features set that clearly differentiate themselves from the objects in the class;
2) Methods for solving these problems are described, which include the **proximity function** that provides the similarity of objects in the space of these informative features and uses the algorithms for an estimate calculation based on Fisher criteria;
3) The steps to solve practical problems based on the proposed theoretical data are developed. It describes step-by-step solutions for the class of "ischemic heart disease" on the basis of symptoms and associated objects.

1. Problem statement

Let's assume that the curriculum formed on the basis of primary data is in the problem of pattern recognition divided into the training set classes and expressed as follows:

The American Journal of Engineering and Technology
(ISSN – 2689-0984
Published: April 22, 2021 | **Pages:** 7-16
Doi: https://doi.org/10.37547/tajet/Volume03Issue04-02

IMPACT FACTOR
2021: 5. 705
OCLC - 1121105677

$$K_1 = \begin{bmatrix} x_{11}^1 & x_{11}^2 & \dots & x_{11}^N \\ x_{12}^1 & x_{12}^2 & \dots & x_{12}^N \\ \vdots & \vdots & \vdots & \vdots \\ x_{1m_1}^1 & x_{1m_1}^2 & \dots & x_{1m_1}^N \end{bmatrix} \dots$$

$$K_r = \begin{bmatrix} x_{r1}^1 & x_{r1}^2 & \dots & x_{r1}^N \\ x_{r2}^1 & x_{r2}^2 & \dots & x_{r2}^N \\ \vdots & \vdots & \vdots & \vdots \\ x_{rm_r}^1 & x_{rm_r}^2 & \dots & x_{rm_r}^N \end{bmatrix}.$$

This can be summarized into a general form as follows:

$$K_p = \begin{bmatrix} x_{p1}^1 & x_{p1}^2 & \dots & x_{p1}^N \\ x_{p2}^1 & x_{p2}^2 & \dots & x_{p2}^N \\ \vdots & \vdots & \vdots & \vdots \\ x_{pm_p}^1 & x_{pm_p}^2 & \dots & x_{pm_p}^N \end{bmatrix}$$

Here $p = \overline{1,r;}$ and the training set is expressed in the form $K = \bigcup_{p=1}^{r} K_p$. This training set may be represented by classes which do not intersect. This corresponds to conditions $K_p \cap K_q = \emptyset, (p \neq q, p = \overline{1,r;} q = \overline{1,r;})$.

Similarly, the components of the object x_{pi} are x_{pi}^j real numbers, which are read as follows: j correspond to feature of i patients and p is the class; here $p = \overline{1,r;}, i = \overline{1,m_p;}, j = \overline{1,N}$; and r represents the total number of classes, m_p is the total number of patients in the class and N is the total number of features.

The overall concept consists of looking at classes each of which corresponds to a specific type of disease: Class K_1 "Unstable angina pectoris", class K_2 "Acute myocardial infarction", class K_3 "Arithmic form". At the same time, the characteristic feature of each class (type of disease) is formed by experts in the field and consists of 62. The overall procedure is organized around the following tasks:

Task1. Determine the feasibility of the features that characterize the objects of the classes we are considering.

Task 2. The features that characterize the objects of the classes we are considering should be continuously converted into binary (0/1).

Task 3. We solve the classification problem of the K_p class objects. Here the main concern is to define whether the class objects belong to one class or another.

Task 4. We select the informative features in class K_p $(p = \overline{1,3})$. $K_p$requires the selection of $\ell \ll 62$ informative features that can clearly distinguish the objects in the class. Here ℓ is a predetermined small number and is read from 1 to 62.

2. Practical problems:

Practical problems are described into four phases.
Phase 1. Determining the feasibility of characteristic features of each object belonging to the above K_p class. This is done separately for each class in the following order:

a). Let's make the following determinations: $\bar{x}_p = (\overline{x}_p^1, \overline{x}_p^2, \dots, \overline{x}_p^N)$ vector, $X_p$the average representative objects of classes, and $p = \overline{1,r}$. We compute the $\bar{x}_p$ components by the following formula:

$$\overline{x}_p^j = \frac{1}{m_p} \sum_{i=1}^{m_p} x_{pi}^j, p = \overline{1,3}; \; j = \overline{1,62}; \; i = \overline{1,m_p}. (1)$$

The results calculated in the cross section of each class are shown in figure 1.

The American Journal of Engineering and Technology
(ISSN – 2689-0984
Published: April 22, 2021 | **Pages:** 7-16
Doi: https://doi.org/10.37547/tajet/Volume03Issue04-02

IMPACT FACTOR
2021: 5. 705
OCLC - 1121105677

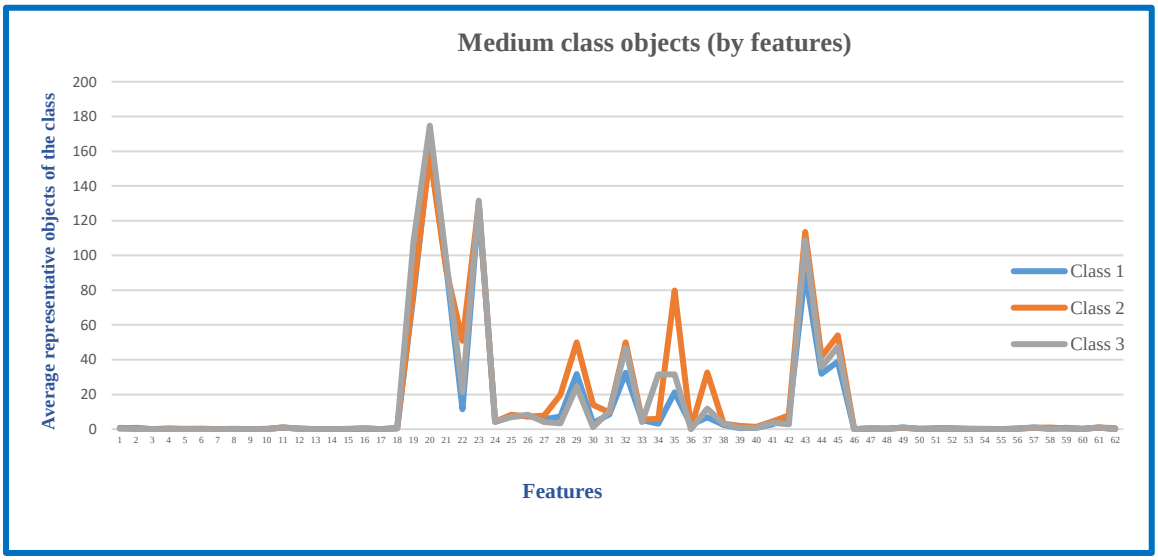

Figure1. Medium class objects (by features)

Figure 1 shows the average representative objects ($\bar{x}_p^j$) of three classes and its 62 features. For small numbers of features the average representative objects are same/common for all three classes. The divergence between the classes is observed for the features located in the interval/range from e.g., 26 to 38. When the number of features increases further (e.g., greater 42) average representative objects are same/common for all three classes.

b). Let's calculate the distance between the objects x_{pi} and $\bar{x}_p$ in the X_p class. This is obtained by the following formula:

$$|x_{pi} - \bar{x}_p| = \sqrt{\sum_{j=1}^{N}(\bar{x}_p^j - x_{pi}^j)^2}, p = \overline{1,3};\ j = \overline{1,62};\ i = \overline{1,m_p}. \quad (2)$$

Using (2) the distance between the object classes is calculated and shown in Figure 2. The Figure shows the variation of the distance between class objects in terms of number of features selected. The feature selection is performed for three different classes. As it appears in Fig.2 the distance between class

object is very big for small number of features in the case of Class 2. In contrast this distance is very big in the case of Class 1 for high number of features. o

The American Journal of Engineering and Technology
(ISSN – 2689-0984
Published: April 22, 2021 | **Pages:** 7-16
Doi: https://doi.org/10.37547/tajet/Volume03Issue04-02

IMPACT FACTOR
2021: 5. 705
OCLC - 1121105677

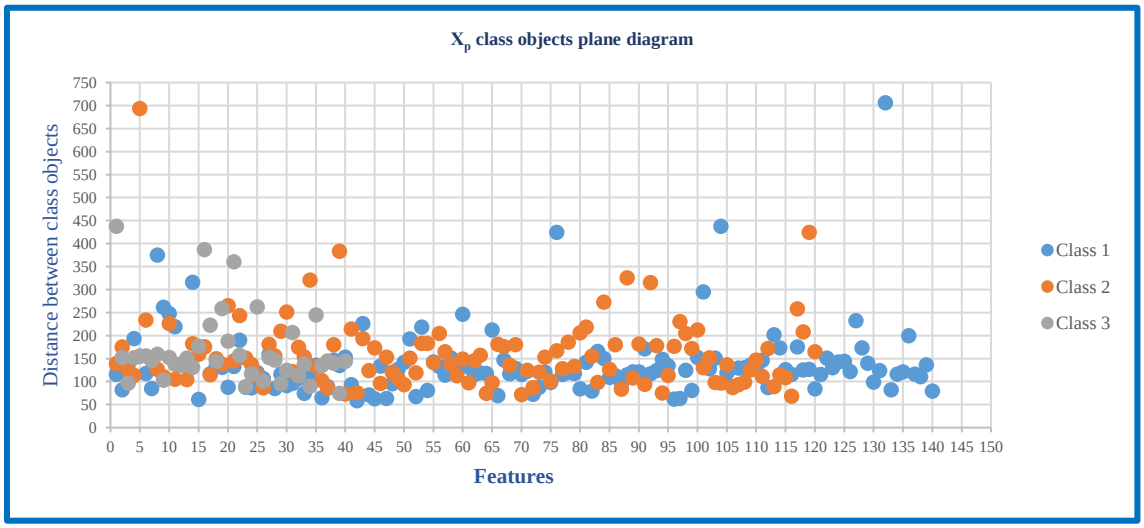

Figure 2. X_p class objects plane diagram

Figure 2 is obtain by calculating the distance $|x_{pi} - \bar{x}_p|$ between objects in each of the three classes. The dots represent the location of the object for a specific number of features selected for each class.

b). The upper limit $D(\bar{x}_p)$ of squares taking into account the objects of class X_p is calculated by the following formula:

$$D(\bar{x}_p) = \sqrt{\frac{1}{m_p}\sum_{j=1}^{m_p}|x_{pi} - \bar{x}_p|^2} = \sqrt{\frac{1}{m_p}\sum_{i=1}^{m_p}\sum_{j=1}^{N}(\bar{x}_p^j - x_{pi}^j)^2}\ . \quad (3)$$

$$p = \overline{1,3};\ j = \overline{1,62};\ i = \overline{1,m_p}.$$

The results of the mean squared deviation of each class $D(\bar{x}_p)$ are shown in the table below (Table 1).

Table 1. The results of the mean squared deviation of each class

Class 1 $(D(\bar{x}_1))$	Class 2 $D(\bar{x}_2)$	Class 3 $D(\bar{x}_3)$
159,6286	177,4676	184,4779

As it appear in table 1 the number of significant features is less in Class 1 compared Classes 2 and 3. Further the number of significant features is less in Class 2 compare Class 3.

c). Let's consider the inequality (4). The related parameters can be calculated as a percentage of class objects:

$$|x_{pi} - \bar{x}_p| \le D(\bar{x}_p), p = \overline{1,r};\ i = \overline{1,m_p}. \quad (4)$$

When the calculation is completed, the feasibility level of characteristic features of each object that belongs to the K_p class are 82.14% (Class 1), 71.67% (Class 2), and 77.50% (Class 3). According to these results, the percentage of performance (expressed by inequality (4)) is obtained/checked in relation to three classes. The mean squared deviation between the objects of each class is determined by the distance $|x_{pi} - \bar{x}_p|$. For

The American Journal of Engineering and Technology
(ISSN – 2689-0984
Published: April 22, 2021 | **Pages:** 7-16
Doi: https://doi.org/10.37547/tajet/Volume03Issue04-02

IMPACT FACTOR
2021: 5. 705
OCLC - 1121105677

each class this distance is less than or equal to $D(\bar{x}_p)$.

Phase 2: The given initial data are presented in a continuous quantitative form. At this phase the process of converting the feature values of class objects to binary vector is carried out.

The process of converting the binary character features to each of the above-mentioned K_p class objects in vector form is made by typing the following symbols in each class and all character sections:

a). $\bar{x}_p = (\overline{x}_p^1, \overline{x}_p^2, \dots, \overline{x}_p^N)$ vector, K_p mean objects of classes, $p = \overline{1,r}$. Compute its components by the following formula:

$$\overline{x}_p^j = \frac{1}{m_p}\sum_{i=1}^{m_p} x_{pi}^j, p = \overline{1,r};\ j = \overline{1,N};\ i = \overline{1,m_p}. \quad (5)$$

b). Let's define the following vectors $a_p = (a_p^1, a_p^2, \dots, a_p^N)$ and $b_p = (b_p^1, b_p^2, \dots, b_p^N)$, and calculate their components by formulas (6), and (7):

$$a_p^j = \frac{1}{m_p}\sum_{i=1}^{m_p}(\overline{x}_p^j - x_{pi}^j)^2, p = \overline{1,r};\ j = \overline{1,N}. (6)$$

$$b_{pi}^j = (\overline{x}_p^j - x_{pi}^j)^2, p = \overline{1,r};\ j = \overline{1,N}. (7)$$

c). The components of the K_p elements of the training set are converted from the binary number by the procedure in (8).

$$x_{pi}^j == \begin{cases} equal\ 1, & if\ \frac{b_{pi}^j}{a_p^j} \le 1, \\ elseequalto\ 0 \end{cases} (8)$$

After this phase is completed, the characteristic feature values that characterize the three classes objects are converted to binary vector.

Phase 3. This phase considers the classification issue of the objects in the K_p class. The focus is to determine whether each object in the class belongs to own class or a different one.

At the same time, each object belonging to the class $X_p$is compared with objects in its class and other classes, and the function of inter-object proximity in the space of informative features is expressed in (9).

$$\rho_i(x_{p1,}x_{p2}) = \begin{cases} 1 \quad if(x_{p1}^i - x_{p2}^i) = 0, i = \overline{1,N}, \\ 0 \qquad\qquad otherwise. \end{cases} (9)$$

In (9), the first condition denotes the degree of similarity between the two objects, and the second condition indicates that they are different. The total comparative evaluation is based on the formula (10).

$$\Gamma_j(x_{pj,}x_{pk}) = \sum_{\kappa=1}^{m_p}\sum_{i=1}^{N}\rho_i(x_{pj,}x_{pk}), j = \overline{1,m_p};\ k = \overline{1,m_p};\ j \ne k. (10)$$

Using (10) the comparative evaluation is calculated for each class, and the largest of the mean values obtained is the attribution of the object to that class.

At this phase, a classification of class objects is consider to determine if the class objects belong to their class or to another class. This process is carried out step-by-step, excluding objects that do not belong to their class at each step, and the objects in the classes are complete until they reach their full class, which corresponds to 100%. The outcome of the selection procedure has revealed that by the end of the procedure, the following reference training set options are selected: 131 (in Class 1), 115 (in Class 2), and 40 objects (in Class 3). Using the initial data given here, 9 objects from 140 objects in the 1st Class, 120 objects from the 2nd Class were excluded from their class because they passed to another class as a result of the classification of 5 objects. From the 3th class, too, one object did not pass to another class, that is, it remained in its class.

The American Journal of Engineering and Technology
(ISSN – 2689-0984
Published: April 22, 2021 | **Pages:** 7-16
Doi: https://doi.org/10.37547/tajet/Volume03Issue04-02

IMPACT FACTOR
2021: 5. 705
OCLC - 1121105677

The execution of this process is described in all three phases above.

Phase 4. An informative features set is selected using the generated reference table. In this case, the classification results obtained are 100% for the reference table. Now, using the convergence function (9), all of the features will be identified as distinct. According to it, a column with randomly selected features is omitted from the reference table, meaning that in the class we are looking at, there are 62 features in all three classes. As consequence of aforementioned feature omission the classification process is carried out using the remaining 61 characters and the proximity function (9). If at the end of the process all objects in the calculation find 100% of the class, the column removed from the table will not be redirected, otherwise the column will be returned to its original location by a random selection and the process will be restarted. The proposed process lasts up to l. If objects have found a different class in their class (switching to another class), the arbitrarily selected symbol will be returned. This process takes place between 62 features in the issue we are looking at, and at the end of the process, the remaining features are distinguished as informative features. In essence, the work performed at this stage is to select the most useful ℓ element from the set of features that characterize the objects under investigation. This selection of informative symbols, that is, seven sets of 10 informative symbols with identical results from the created software corresponds $\ell = 10$. The program separates the results by analyzing seven groups of 10 feature sets provided from the common features set and determines/evaluates the final result shown in Table 2.

Table 2. Informative Features Set

Informative Features Set
$x_1, x_7, x_{22}, x_{27}, x_{36}, x_{41}, x_{51}, x_{57}, x_{62}$

Table 2 reveals a set of informative variables in the three classes. The results depicted correspond to the essential/important information in the three classes. A selected informative feature set is appropriate with reference table objects.

The method for selecting features which are focused on the use for specific measure of information is developed in this article. Its essence is to use the measure of importance of the initial feature which is properly processed degree of reduction which is called "votes", when this feature is being removed.

The sequential procedure for eliminating signs consists of as follows.

According to the teaching sequence, the table of X is presented for all the features of the

initial system $= (x^1, x^2, \ldots, x^N)$, using a pseudo-random sensor for generating Boolean vectors λ where $\sum_{j=1}^{N} \lambda^j = \ell$. Moreover, the probability of occurrence of each of N features is the same at the beginning and equal to $\frac{1}{N}$. In other words, based on the vector of probabilities is $p = (p^1, p^2 \ldots, p^N)$, where p^j- is the probability of occurrence of feature j. The sensor generates some pseudo-random vector $\lambda = (\lambda^1, \ldots, \lambda^N)$, $\lambda^j \in \{0;1\}$, $j = \overline{1, N}$;

$\sum_{j=1}^{N} \lambda^j = \ell$, and at the beginning is $p^j = \frac{1}{N}$, $j = \overline{1, N}$. At each step, the pseudo-random sensor is $p = (p^1, p^2 \ldots, p^N)$, taking into account the current probability vector, k

The American Journal of Engineering and Technology
(ISSN – 2689-0984
Published: April 22, 2021 | **Pages:** 7-16
Doi: https://doi.org/10.37547/tajet/Volume03Issue04-02

IMPACT FACTOR
2021: 5. 705
OCLC - 1121105677

$\lambda_1,\ldots,\lambda_k$ vectors are generated where (k=I0÷I5)). Among them, a pair of vectors λ_{min}, λ_{max}, is chosen, on which the functional I (λ) respectively takes minimum and maximum values.

Next, the probability vector is changed. j is equal to one for each component, corresponding component of the probability vector decreases by some amount $h << \frac{1}{N}$1/N, which is called the penalty, if the latter does not become negative. In the other cases, the reduction is carried out to zero.

$$\lambda_{min}^{j} = 1 => p^{j} := \max\{0, p^{j} - h\}, j = \overline{1,N}.$$

Then, the probabilities p^j, corresponding to the unit components of the vector λ_{max},are increased to the amount of $d = \frac{H}{\ell}$ (H-total penalty). More precisely

$$\lambda_{max}^{j} = 1 => p^{j} := p^{j} + d, \qquad j = \overline{1,N}.$$

After making changes to the probability vector, the transition to the next step is made.

The proposed method, namely change of the probability vector *p* is carried out step by step until nonzero components of *I* will not exist in it.

CONCLUSION

This paper has developed a method for the easy selection of informative symptoms in the classification of medical data. The method was based on the following steps: (a) Determining the feasibility of the characteristic features of the objects of training set classes. (b) Carrying out a converting process of the characteristic features of the objects in the classes from continuous quantitative to binary. (c) Formation of the reference table as result of exclusion of objects that did not find their own class.

REFERENCES

1. Zhuravlev Yu.I. "Selected scientific works". – M: "Magistr" Publishing house, 1998. - 420 p.
2. Yu, L., H. Liu., "Efficient Feature Selection via Analysis of Relevance and Redundancy". – J. Mach. Learn. Res., Vol. 5, 2004, No Oct, pp. 1205-1224.
3. Yu, L.: "Toward Integrating Feature Selection Algorithm for Classification and Clustering", IEEE Transaction on Knowledge and Data Engineering 17(4), 491-502
4. Yan, K., D. Zhang. "Feature Selection and Analysis on Correlated Gas Sensor Data with Recursive Feature Elimination", – Sensors Actuators, B Chem., Vol. 212, Jun 2015, pp. 353-363.
5. Nishanov A.Kh,, Turakulov Kh.A., Turakhanov Kh.V. "A decision rule for identification of eye pathologies", Biomedical Engineering (1999) 33(4) 178-179.
6. Nishanov A. Kh,,Turakulov Kh.A.,Turakhanov Kh.V. "A decisive rule in classifying diseases of the visual system" Meditsinskaia tekhnika (1999) (4) 16-18.
7. Xiang Fang, Lina Wang. "Feature Selection Based on Fisher Criterion and Sequential Forward Selection for Intrusion Detection", Revista de la Facultad de Ingeniería U.C.V., Vol. 32, N° 1, pp.498-503, 2017.
8. Linhui Sun, Sheng Fu and Fu Wang. "Decision tree SVM model with Fisher feature selection for speech emotion recognition", Sun et al. EURASIP Journal on Audio, Speech, and Music Processing (2019) 2019:2. https://doi.org/10.1186/s13636-018-0145-5
9. Yan, X.,Tan, M., Yan,Y., Lü, M. (2012). "Research on Hidden Markov Model-Based And Neural Network-Based Intrusion Detections", Computer Applications and Software. 29(2), 294–297.
10. Jing, X.,Wang, H., Nie, K., Luo, Z. (2012). "Feature Selection Algorithm Based on

IMGA and MKSVM to Intrusion Detection", Computer Science, 39(7), 96–100.

11. Koushal Kumar, Jaspreet Singh Batth. "Network Intrusion Detection with Feature Selection Techniques using Machine-Learning Algorithms", International Journal of Computer Applications (0975 – 8887) Volume 150 – No.12, September 2016, 1-13.
12. Nishanov A.Kh., Djurayev G.P., Kasanova M.Kh. "Improved algorithms for calculating evaluations in processing medical data", National Institute of Science Communication and Information Resources (NISCAIR)-India, 2019,-3158-3165.
13. Kamilov M., Nishanov A., Beglerbekov R. "Modified stages of algorithms for computing estimates in the space of informative features", International Journal of Innovative Technology and Exploring Engineering (2019) 8(6).
14. Sulaiman M Labadin J. "Feature selection based on mutual information", Publisher: Institute of Electrical and Electronics Engineers (IEEE), 2015 pp: 1-6.
15. Xuelian Deng, Yuqing Li, Jian Weng, Jilian Zhang. "Feature selection for text classification: A review", Multimedia Tools and Applications, 2019, 1007/s11042-018-6083-5.
16. Ma, S. and Huang, J. (2008) "Penalized feature selection and classification in bioinformatics", Briefings in Bioinformatics. doi: 10.1093/bib/bbn027.
17. Sun, L., Fu, S. and Wang, F. (2019) "Decision tree SVM model with Fisher feature selection for speech emotion recognition", Eurasip Journal on Audio, Speech, and Music Processing, 2019(1). doi: 10.1186/s13636-018-0145-5.
18. Krishna, R. S. B. and Aramudhan, M. (2014) "Feature selection based on information theory for pattern classification", in 2014 International Conference on Control, Instrumentation, Communication and Computational Technologies, ICCICCT 2014. doi: 10.1109 / ICCICCT. 2014. 6993149.
19. Mokshin, V.V. *et al.* (2018) "Parallel genetic algorithm of feature selection for complex system analysis", in *Journal of Physics: Conference Series*. doi: 10.1088/1742-6596/1096/1/ 012089.
20. Solorio-Fernández, S., Carrasco-Ochoa, J.A. and Martínez-Trinidad, J.F. (2019) "A review of unsupervised feature selection methods", *Artificial Intelligence Review*.doi: 10.1007/s10462-019-09682-y.
21. Urbanowicz, R. J. *et al.* (2018) "Benchmarking relief-based feature selection methods for bioinformatics data mining", *Journal of Biomedical Informatics*, 85. doi: 10.1016/j.jbi.2018.07.015.
22. Cai, J. *et al.* (2018) "Feature selection in machine learning: A new perspective", *Neurocomputing*, 300. doi: 10.1016/j.neucom.2017.11.077.
23. Ghaemi, M. and Feizi-Derakhshi, M. R. (2016) "Feature selection using Forest Optimization Algorithm", *Pattern Recognition*. Elsevier Ltd, 60, pp. 121–129. doi: 10.1016/j.patcog.2016.05.012.