International Journal of Advance Scientific Research
(ISSN – 2750-1396)
VOLUME 04 ISSUE 03 Pages: 149-155
SJIF IMPACT FACTOR (2022: 5.636) (2023: 6.741) (2024: 7.874)
OCLC – 1368736135

Journal Website: http://sciencebring.com/index.php/ijasr

Research Article

AN ALGORITHM FOR SELECTING AN INFORMATIVE SYMBOLIC COMPLEX BASED ON CLASSIFICATION ERROR COEFFICIENTS AND PROBABILISTIC INDICATORS IN THE REPRESENTATION OF SYMBOLS

Submission Date: March 20, 2024, **Accepted Date:** March 25, 2024,
Published Date: March 30, 2024
Crossref doi: https://doi.org/10.37547/ijasr-04-03-27

Juraev Gulomjon Primovich
Associate Professor of the Department of Methods of Exact and Natural Sciences, National Center for Teacher Training in New Methods of the Kashkadarya Region, Uzbekistan

Saparov Saidkul Khojamurodovich
Doctoral student, Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, Uzbekistan

ABSTRACT

In the primary processing of information, in particular in character recognition, an important issue is the selection and classification of an informative feature or a set of features classifying objects. Despite the fact that a number of methods and algorithms have been proposed to solve these problems, there are many problems in this direction that are waiting to be solved. This is due to the fact that many of the proposed approaches strongly depend on the nature of the object of study, the number of its features, the type of perceived values of features, the size of the study sample, etc., and impose certain requirements on the above. In addition, each method or algorithm will strongly depend on whether the criterion of informative selection of features and the defining rule determining the quality of the choice made are correctly chosen.

This article presents a description of the algorithm developed taking into account the above approaches to the selection of information complexes of signs, as well as recommendations on the application of this algorithm in practical matters of the medical field, i.e. in ischemic heart disease obtained as an object of study (5 classes, 507 objects, 89 signs, including X_1 class “strenuous angina”, X_2 class “Acute myocardial infarction”, class X_3 “Arrhythmic form”, class X_4 “Postinfarction cardiosclerosis”, for class X_5 “Persistent

International Journal of Advance Scientific Research
(ISSN – 2750-1396)
VOLUME 04 ISSUE 03 Pages: 149-155
SJIF IMPACT FACTOR (2022: 5.636) (2023: 6.741) (2024: 7.874)
OCLC – 1368736135

form of atrial fibrillation") formulated training was applied to the selection and positive results were achieved.

KEYWORDS

Primary data processing, training sample, classification, selection of an informative symbolic complex, classification error.

INTRODUCTION

The experience conducted during the pandemic in the world has shown the need to accelerate the work on digitalization of medicine while eliminating problems in the field of medicine. In particular, in the field of medicine, special attention is also paid to the development and development of data mining methods that correspond to human decisions.

Important scientific projects carried out by scientific societies and research institutes engaged in medical research around the world are genomics and the Human Genome project, vaccines and vaccination, medical robots, blood detection and analysis technologies, medical digital technologies, geonomics (bioanalysis), medical rehabilitation technologies, the latest scientific research in the field of medicine and the results count. These researches, aimed at the scientific development of medical science, serve as the basis for the intellectual analysis of medical information in solving classification issues, choosing a set of informative signs, forming symptom complexes and diagnosing diseases based on them. Usually, the size of the medical information that the disease evaluator collects and stores is tens or hundreds of characters. Based on this information, the decision-making process by industry professionals becomes an extremely complex process.

Therefore, issues related to the transition from a large size to a smaller one, which is more important in the primary processing of medical information, that is, with the selection and evaluation of information sign complexes, the formation of symptom complexes in the context of information sign complexes and the development of modern recognition systems that help industry specialists in solving issues such as diagnosis on based on them. In particular, the decision on the transition from a large size to a smaller significant size during primary data processing, that is, on the set of informative symbolic complexes, was of interest to scientists around the world, and in this regard, a number of studies are underway [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15].

This article also develops an algorithm for selecting an informative set of symbols based on classification error coefficients and probabilistic indicators in the representation of symbols.

THE MAIN PART

The standard was chosen for the preparation of $x_{p1}, x_{p2}, \dots, x_{pm_p} \in X_p$, $p = \overline{1,r}$; cognitively, the

International Journal of Advance Scientific Research
(ISSN – 2750-1396)
VOLUME 04 ISSUE 03 Pages: 149-155
SJIF IMPACT FACTOR (2022: 5.636) (2023: 6.741) (2024: 7.874)
OCLC – 1368736135

question of choosing a set of characters is the symbol value with a space in the part $\lambda = (\lambda^1, \lambda^2, \dots, \lambda^N), \lambda^j \in \{0,1\}, j = \overline{1,N};$ is included in the vector.

Here, the main problem of the λ vector is to ensure the transition from an N-dimensional character space to an ℓ -dimensional character space small enough for the N number.

Here λ means that the signs corresponding to the components of the vector that are equivalent together mean that the selected part has a set of characters in space, and signs equal to zero mean that the corresponding characters do not participate in the extracted information character set.

Definition 1. The studied vector space λ is called ℓ -dimensional if, for an arbitrary vector λ under consideration, the sum of its components is equal to

$$\sum_{j=1}^{N} \lambda^j = \ell$$

Similarly, in an ℓ -dimensional vector space, let the set in which the vectors λ are located be denoted by Λ^ℓ.

By definition, the mathematical expression of this set will be:

$$\Lambda^\ell = \left\{ \lambda: \sum_{j=1}^{N} \lambda^j = \ell\,,\ \ \lambda^j \in \{0,1\}, j = \overline{1,N}; \right\}$$

here, set of Λ^ℓ number of vectors λ is equal to $C_N^\ell = \frac{N!}{\ell!(N-\ell)!}$.

if $\ell = \overline{1,N}$, in this set $2^N - 1$ is λ vectors:

$$\sum_{\ell=1}^{N} C_N^\ell = 2^N - 1.$$

Definition 2. Λ^ℓ vectors λ which are elements of the set ℓ are called informative vectors. The number of non-zero components of these vectors is ℓ. Suppose that the objects of the training sample based on the reference table $\theta(N)$ have a classification error, and the number of incorrectly recognized objects is $\varkappa(N)$.

Learning here on the other hand, the total number of objects in the sample is M, then the classification error in N -dimensional space is calculated using the formula $\theta(N) = \frac{\varkappa(N)}{M}$.

Let the mechanism of operation of the proposed algorithm be defined as follows. First, the probability vector for selecting symbols of the objects of the training sample $p_\nu = (p_\nu^1, p_\nu^2, \dots, p_\nu^N)$ is introduced.

Here p_ν means that the ν in the index of the probability vector is the probability vector. Usually, this index will be associated with the choice of a probability vector. The first choice of $\nu = 1$ bo'lib, $p_1^j = \frac{1}{N}; j = \overline{1,N};$

Similarly, $p_\nu = (p_\nu^1, p_\nu^2, \dots, p_\nu^N)$ probability vector ℓ is informative $\lambda \in \Lambda^\ell$ in vector space we want $p_\nu|_\lambda = (\lambda^1 p_\nu^1, \lambda^2 p_\nu^2, \dots, \lambda^N p_\nu^N)$ the form is characterized.

Based on the initial given probability, $\Lambda^\ell$from the set $p_\nu = (p_\nu^1, p_\nu^2, \dots, p_\nu^N)$, to the probability vector λ, the vector is randomly selected, and this is denoted as $\lambda_{\nu 1}, \lambda_{\nu 2}, \dots, \lambda_{\nu k}$. .

In the context of all the features of this training sample, the classification process is

International Journal of Advance Scientific Research
(ISSN – 2750-1396)
VOLUME 04 ISSUE 03 Pages: 149-155
SJIF IMPACT FACTOR (2022: 5.636) (2023: 6.741) (2024: 7.874)
OCLC – 1368736135

carried out on the basis of sequential calculation of the following Formulas (1), (2) and (3:

1. Initially, each object belonging to the X_p class is compared with objects in its class, as well as with objects in other classes x_{pi} the proximity function between the object and objects x_{pq} of the X_p class $\rho^j(x_{pi,}x_{pq})$ is calculated as:

$$\rho^j(x_{pi,}x_{pq}) = \begin{cases} 1, \text{if } |x_{pi}^j - x_{pq}^j| = 0 \\ 0, \text{otherwise} \end{cases} \quad (1)$$

Here $j = \overline{1,N}$; *is equal* $i, q = \overline{1,m_p}; i \neq q;$.

2. Comparative evaluation for each class, that is, each object of this class X_p is $\Gamma_p(x_{pi,}x_{pq})$ for x_{pi}

$$\Gamma_p(x_{pi,}x_{pq}) = \frac{1}{m_p}\sum_{q=1}^{m_p}\sum_{j=1}^{N}\rho^j(x_{pi,}x_{pq}), i = \overline{1,m_p}; i \neq q; \quad (2)$$

3. Based on the results obtained in a comparative assessment, the maximization problem is solved using the formula

$$\Gamma_p^*(x_{pi,}x_{pq}) = \max_{p=\overline{1,r;}} \Gamma_p(x_{pi,}x_{pq}), i, q = \overline{1,m_p}; \; i \neq q; \quad (3)$$

And in result $x_{pi} \in X_p$.

It determines the levels of individual significance of all features in the reference sample of training. This is a reference training without selection $p_v = (p_v^1, p_v^2, \dots, p_v^N)$ using the probability vector, the column containing one character is excluded from the training sample and the classification error coefficient is calculated on the segment of the remaining characters $\theta(N-1)$. These processes include the following two important cases.

The first case: the difference in classification errors if $\theta(N) = \theta(N-1)$, in this case, this character is completely removed from the training sample, and the number of characters is reduced by one character.

The second case: otherwise, $\theta(N) \neq \theta(N-1)$, in this case, this symbol will be left in the training sample, and the process will start againIn the second case, the number of signs of educational selection does not change.

Description of the proposed algorithm

1-step. Initial classification error is $\theta(N)$;

2-step. the value ℓ is entered, $\ell << N$;

3-step. In $v = 1$ is $p_v = (p_v^1, p_v^2, \dots, p_v^N)$ for $p_v^j = \frac{1}{N}; j = \overline{1,N}$; initial values are assigned;

4-step. One character is randomly selected from N characters and subtracted from the array. Than, In the cross section of $N-1$characters, the classification process is performed. In a result for $N-1$ characters, the error coefficient is equal to $\theta(N-1)$;

5-step. If $\theta(N) = \theta(N-1)$, here, in this case, this symbol is completely removed from the training sample, otherwise this symbol remains in the training sample, and the process begins with step 4. This process is reversible as long as $N - h = \ell$.

If $N - h = \ell$, in this case is formed number of the signs ℓ and a signal solution with multiple

International Journal of Advance Scientific Research
(ISSN – 2750-1396)
VOLUME 04 ISSUE 03 Pages: 149-155
SJIF IMPACT FACTOR (2022: 5.636) (2023: 6.741) (2024: 7.874)
OCLC – 1368736135

characters is formulated and proceeds to step 6. Similarly, if equality is not satisfied, the process continues until all characters are processed one by one. However, after all the symbols are considered, we proceed to the next step;

6-step. Suppose that $N - h = \ell$, then given the following $p^j = \frac{1}{2N-h}$, $j = \overline{1,h;}$ $p^j = \frac{2}{2N-h}, j = \overline{h+1,N;}$ with probability 4 –process, in the reverse step, only those symbols that were selected earlier, which should be chosen randomly, and those that remained in the system differ from each other in probabilities.

Then go to **step 5.**

With the help of a software package created based on this algorithm, informative symbolic medicine complexes for coronary heart disease were selected (see Table 1).

1-table

information symbolic complexes selected in 103 variants when $\ell = 9$

$\ell = 9$	$\ell = 9$	$\ell = 9$	$\ell = 9$
$x^3,x^5,x^{10},x^{11},x^{41},x^{52},x^{53},x$	$x^3,x^5,x^{11},x^{52},x^{53},x^{56},x^{62},x$	$x^3,x^5,x^{10},x^{11},x^{40},x^{41},x^{52},x$	$x^3,x^{11},x^{40},x^{41},x^{48},x^{52},x^{53},x$
$x^3,x^5,x^{11},x^{41},x^{52},x^{53},x^{64},x$	$x^3,x^5,x^{11},x^{52},x^{53},x^{62},x^{63},x$	$x^{10},x^{11},x^{40},x^{41},x^{52},x^{53},x^{59}$	$x^3,x^5,x^{11},x^{40},x^{49},x^{52},x^{53},x$
$x^3,x^5,x^{10},x^{11},x^{52},x^{53},x^{62},x$	$x^3,x^5,x^{11},x^{41},x^{52},x^{53},x^{68},x$	$x^3,x^5,x^{11},x^{14},x^{52},x^{53},x^{62},x$	$x^3,x^{11},x^{40},x^{52},x^{53},x^{59},x^{62},x$
$x^3,x^5,x^{11},x^{52},x^{53},x^{62},x^{80},x$	$x^3,x^5,x^{11},x^{34},x^{52},x^{53},x^{62},x$	$x^3,x^5,x^{11},x^{41},x^{48},x^{52},x^{53},x$	$x^3,x^5,x^{11},x^{25},x^{52},x^{53},x^{62},x$
$x^3,x^5,x^{11},x^{52},x^{53},x^{62},x^{64},x$	$x^{10},x^{11},x^{40},x^{52},x^{53},x^{59},x^{62}$	$x^3,x^5,x^{11},x^{41},x^{52},x^{53},x^{80},x$	$x^3,x^5,x^{11},x^{32},x^{41},x^{52},x^{53},x$
$x^3,x^5,x^{11},x^{25},x^{41},x^{52},x^{53},x$	$x^3,x^5,x^{10},x^{11},x^{41},x^{52},x^{53},x$	$x^3,x^5,x^{11},x^{24},x^{52},x^{53},x^{62},x$	$x^3,x^5,x^{11},x^{41},x^{42},x^{52},x^{53},x$
$x^{10},x^{11},x^{40},x^{52},x^{53},x^{59},x^{60}$	$x^3,x^5,x^{11},x^{14},x^{41},x^{52},x^{53},x$	$x^3,x^{11},x^{40},x^{41},x^{52},x^{53},x^{59},$	$x^3,x^5,x^{11},x^{38},x^{41},x^{52},x^{53},x$
$x^3,x^5,x^{11},x^{42},x^{52},x^{53},x^{62},x$	$x^3,x^5,x^{11},x^{21},x^{41},x^{52},x^{53},x$	$x^3,x^5,x^{11},x^{19},x^{41},x^{52},x^{53},x$	$x^3,x^5,x^{11},x^{40},x^{41},x^{52},x^{53},x$
$x^3,x^5,x^{11},x^{41},x^{52},x^{53},x^{77},x$	$x^3,x^5,x^{11},x^{24},x^{52},x^{53},x^{62},x$	$x^3,x^5,x^{11},x^{24},x^{41},x^{52},x^{53},x$	$x^3,x^{11},x^{24},x^{41},x^{52},x^{53},x^{59},x$
$x^3,x^5,x^{11},x^{41},x^{52},x^{53},x^{80},x$	$x^3,x^5,x^{11},x^{40},x^{41},x^{52},x^{53},x$	$x^3,x^{11},x^{40},x^{41},x^{52},x^{53},x^{59},$	$x^3,x^4,x^5,x^{11},x^{41},x^{52},x^{53},x^8$
$x^3,x^5,x^{11},x^{12},x^{24},x^{52},x^{53},x$	$x^3,x^5,x^{11},x^{40},x^{42},x^{52},x^{53},x$	$x^3,x^5,x^{11},x^{41},x^{52},x^{53},x^{63},x$	$x^3,x^5,x^{11},x^{26},x^{41},x^{52},x^{53},x$
$x^3,x^5,x^{11},x^{48},x^{52},x^{53},x^{62},x$	$x^3,x^5,x^{11},x^{49},x^{52},x^{53},x^{62},x$	$x^3,x^5,x^{11},x^{40},x^{52},x^{53},x^{62},x$	$x^3,x^5,x^{11},x^{38},x^{52},x^{53},x^{62},x$
$x^3,x^5,x^{11},x^{38},x^{41},x^{52},x^{53},x$	$x^3,x^{11},x^{12},x^{24},x^{52},x^{53},x^{59},$	$x^3,x^5,x^{11},x^{26},x^{40},x^{41},x^{52},x$	$x^3,x^5,x^{11},x^{20},x^{52},x^{53},x^{62},x$
$x^{10},x^{11},x^{40},x^{52},x^{53},x^{59},x^{62}$	$x^3,x^{11},x^{38},x^{40},x^{41},x^{52},x^{53},$	$x^3,x^5,x^{11},x^{38},x^{40},x^{52},x^{53},x$	$x^3,x^5,x^{11},x^{19},x^{52},x^{53},x^{62},x$
$x^3,x^5,x^{11},x^{52},x^{53},x^{62},x^{77},x$	$x^3,x^5,x^{11},x^{52},x^{53},x^{54},x^{62},x$	$x^3,x^5,x^{11},x^{38},x^{40},x^{41},x^{52},x$	$x^3,x^5,x^{11},x^{52},x^{53},x^{54},x^{62},x$
$x^3,x^{11},x^{52},x^{53},x^{59},x^{62},x^{64},$	$x^{10},x^{11},x^{40},x^{52},x^{53},x^{59},x^{62}$	$x^{10},x^{11},x^{40},x^{42},x^{52},x^{53},x^{59}$	$x^3,x^5,x^{11},x^{18},x^{52},x^{53},x^{62},x$

International Journal of Advance Scientific Research
(ISSN – 2750-1396)
VOLUME 04 ISSUE 03 Pages: 149-155
SJIF IMPACT FACTOR (2022: 5.636) (2023: 6.741) (2024: 7.874)
OCLC – 1368736135

$x^3, x^5, x^{11}, x^{40}, x^{52}, x^{53}, x^{62}, x$	$x^3, x^5, x^{11}, x^{32}, x^{52}, x^{53}, x^{62}, x$	$x^3, x^5, x^{11}, x^{40}, x^{41}, x^{52}, x^{53}, x$	$x^3, x^5, x^{11}, x^{26}, x^{52}, x^{53}, x^{62}, x$
$x^3, x^5, x^{11}, x^{40}, x^{41}, x^{52}, x^{53}, x$	$x^3, x^{11}, x^{40}, x^{41}, x^{52}, x^{53}, x^{59},$	$x^3, x^5, x^{11}, x^{36}, x^{40}, x^{41}, x^{52}, x$	$x^3, x^{11}, x^{40}, x^{52}, x^{53}, x^{59}, x^{62},$
$x^3, x^5, x^{11}, x^{22}, x^{52}, x^{53}, x^{62}, x$	$x^3, x^{11}, x^{40}, x^{52}, x^{53}, x^{59}, x^{60},$	$x^3, x^5, x^{11}, x^{41}, x^{45}, x^{52}, x^{53}, x$	$x^{10}, x^{11}, x^{52}, x^{53}, x^{59}, x^{62}, x^{64},$
$x^3, x^5, x^{11}, x^{22}, x^{41}, x^{52}, x^{53}, x$	$x^3, x^5, x^{11}, x^{35}, x^{52}, x^{53}, x^{62}, x$	$x^3, x^5, x^{11}, x^{52}, x^{53}, x^{62}, x^{68}, x$	$x^{10}, x^{11}, x^{40}, x^{41}, x^{52}, x^{53}, x^{59},$
$x^3, x^{11}, x^{40}, x^{52}, x^{53}, x^{59}, x^{62},$	$x^3, x^{11}, x^{40}, x^{52}, x^{53}, x^{59}, x^{62},$	$x^3, x^4, x^5, x^{11}, x^{52}, x^{53}, x^{62}, x^8$	$x^3, x^5, x^{11}, x^{40}, x^{52}, x^{53}, x^{62}, x$
$x^3, x^5, x^{11}, x^{52}, x^{53}, x^{61}, x^{62}, x$	$x^3, x^{11}, x^{40}, x^{49}, x^{52}, x^{53}, x^{59},$	$x^3, x^{11}, x^{40}, x^{41}, x^{52}, x^{53}, x^{59},$	$x^3, x^{11}, x^{19}, x^{41}, x^{52}, x^{53}, x^{59}, x$
$x^3, x^5, x^{11}, x^{36}, x^{52}, x^{53}, x^{62}, x$	$x^3, x^5, x^{11}, x^{41}, x^{49}, x^{52}, x^{53}, x$	$x^3, x^5, x^{11}, x^{29}, x^{41}, x^{52}, x^{53}, x$	$x^3, x^5, x^{11}, x^{25}, x^{40}, x^{41}, x^{52}, x$
$x^3, x^5, x^{11}, x^{29}, x^{52}, x^{53}, x^{62}, x$	$x^3, x^5, x^{11}, x^{40}, x^{52}, x^{53}, x^{62}, x$	$x^3, x^5, x^{11}, x^{40}, x^{41}, x^{48}, x^{52}, x$	$x^3, x^5, x^{11}, x^{41}, x^{52}, x^{53}, x^{63}, x$
$x^3, x^{10}, x^{11}, x^{41}, x^{52}, x^{53}, x^{62},$	$x^3, x^5, x^{11}, x^{40}, x^{52}, x^{53}, x^{60}, x$	$x^3, x^{11}, x^{40}, x^{42}, x^{52}, x^{53}, x^{59},$	$x^3, x^5, x^{11}, x^{22}, x^{41}, x^{52}, x^{53}, x$
$x^3, x^5, x^{11}, x^{40}, x^{48}, x^{52}, x^{53}, x$	$x^3, x^5, x^{11}, x^{45}, x^{52}, x^{53}, x^{62}, x$	$x^3, x^5, x^{11}, x^{20}, x^{41}, x^{52}, x^{53}, x$	

Conclusion

As a criterion of informativeness in solving the problem of choosing a set of informative features, which is one of the main tasks of primary data processing, the classification error coefficient and the probability vector of the features of the sample objects were obtained, aimed at reducing the classification error. In addition, the probability vector used when selecting symbols prevents the inappropriate exclusion of significant object symbols from the selection.

With the help of a software package developed on the basis of the proposed algorithm, it was possible to select and diagnose (classify) based on a set of informative signs that are part of coronary heart disease, which can serve as a basis for the diagnosis of the disease, that is, the class is considered as class X_1- "Strenuous angina", class X_2 "Acute myocardial infarction", class X_3 "Arrhythmic form", class X_4 "Postinfarction cardiosclerosis", class X_5 "Persistent form of atrial fibrillation, in particular, based on clinical signs in patients.

References

1. Bykova V.V., Kataeva A.V. Methods and means of analyzing the informative value of signs in the processing of medical data//Software products and systems /Software & Systems № 2 (114), 2016.-c. 172-178.
2. Fazylov Sh.Kh., Nishanov A.Kh., Mamatov N.S. Methods and algorithms for selecting informative features based on heuristic criteria of informativeness//Monograph.-T.: Fan wa technology.-Tashkent, 2017.-132 p.

International Journal of Advance Scientific Research
(ISSN – 2750-1396)
VOLUME 04 ISSUE 03 Pages: 149-155
SJIF IMPACT FACTOR (2022: 5.636) (2023: 6.741) (2024: 7.874)
OCLC – 1368736135

3. Ashok B., Aruna P. Comparison of Feature selection methods for diagnosis of cervical cancer using SVM classifier//Journal of Engineering Research and Applications. ISSN: 2248-9622, Vol. 6, Issue 1, (Part-1) January 2016.-pp. 94-99.
4. Bolón-Canedo, V. & Alonso-betanzos, A. Ensembles for feature selection: A review and future trends//Information Fusion 52(2019).-pp. 1-12.
5. Emary, E., Zawbaa, H. M. & Hassanien, A. E. Binary grey wolf optimization approaches for feature selection//Neurocomputing 172, 2016.-pp.371-381.
6. Faris, H. et al. An efficient binary Salp Swarm Algorithm with crossover scheme for feature selection problems//Knowledge-based Systems 154, 2018.-pp. 43-67.
7. Gao, W., Hu, L. & Zhang, P. Class-specific mutual information variation for feature selection//Pattern Recognition 79, 2018.-pp. 328-339.
8. Gao, W., Hu, L., Zhang, P. & He, J. Feature selection considering the composition of feature relevancy//Pattern Recognition Letters 112, 2018.-pp. 70-74.
9. Hussien A., Hassanien A., Houssein E., et al.See more. S-shaped binary whale optimization algorithm for feature selection//Advances in Intelligent Systems and Computing, Vol. 727, 2019.-pp. 79-87.
10. Li, J. & Liu, H. Challenges of Feature Selection for Big Data Analytics//IEEE Intelligent Systems 32, (2017).-pp. 9-15.
11. Liu, C., Wang, W., Zhao, Q., Shen, X. & Konan, M. A new feature selection method based on a validity index of feature subset. Pattern Recognition Letters 92, (2017).-pp. 1-8.
12. Nishanov A.Kh., Djurayev G.P., Kasanova M.Kh. Improved algorithms for calculating evaluations in processing medical data//Compusoft: An International Journal of Advanced Computer Technology, 8(6), June-2019.-pp. 3158-3165.
13. Nishanov A.Kh., Akbaraliev B.B., Juraev G.P., Khasanova M.A., Maksudova M.Kh., Umarova Z.F. The algorithm for selection of symptom complex of ischemic heart diseases based on flexible search//Journal of Cardiovascular Disease Research, Vol. 11(2), 2020.-pp. 218-223.
14. Nishanov A.K., Akbaraliev B.B. and Djurayev G.P., A Symptom Selection Algorithm Based on Classification Errors//International Conference on Information Science and Communications Technologies (ICISCT), 2020.-pp. 1-4.
15. Nishanov A.Kh., Djurayev, G.P., Khasanova, M.A. Classification and feature selection in medical data preprocessing//Compusoft: An International Journal of Advanced Computer Technology, 9(6), June-2020.-pp. 3725-3732.