

# Analytical method for selection an informative set of features with limited resources in the pattern recognition problem

*Akhram Nishanov*<sup>1</sup>, *Bakhtiyorjon Akbaraliev*<sup>1</sup>, *Rasul Beglerbekov*<sup>2</sup>, *Oybek Akhmedov*<sup>1</sup>, *Shukhrat Tajibaev*<sup>3</sup>, and *Rashidjon Kholiknazarov*<sup>1\*</sup>

<sup>1</sup>Tashkent University of Information Technologies named after Muhammad Al-Khwarizimi, 100200, 108 Amir Temur main street, Tashkent, Uzbekistan

<sup>2</sup>Nukus branch of Tashkent state agrarian university, 230100, Nukus Abdambetov st., Nukus, Uzbekistan

<sup>3</sup>Karakalpak State University named after Berdakh, 742000, Acad. Abdirov st. 1., Nukus, Uzbekistan

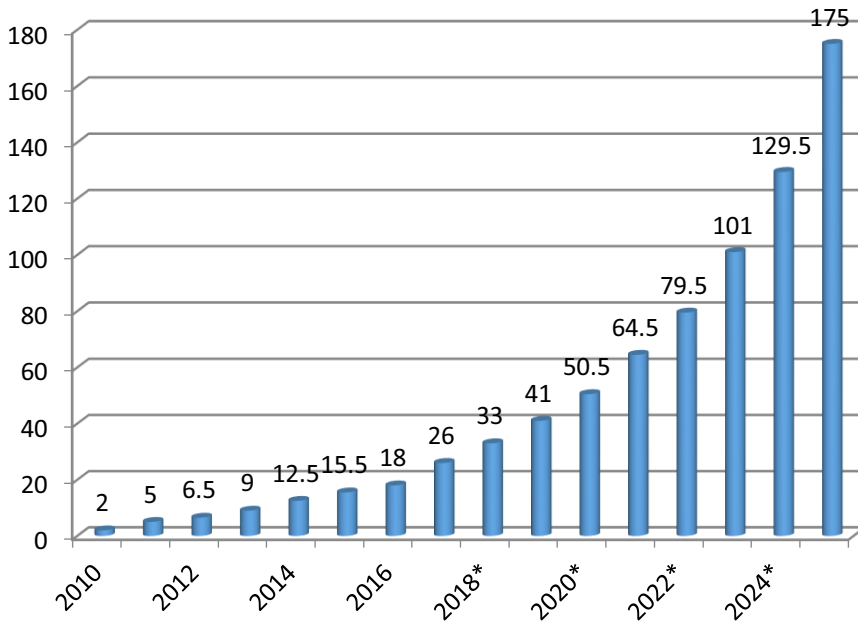
**Abstract.** Feature selection is one of the most important issues in Data Mining and Pattern Recognition. Correctly selected features or a set of features in the final report determines the success of further work, in particular, the solution of the classification and forecasting problem. This work is devoted to the development and study of an analytical method for determining informative attribute sets (IAS) taking into account the resource for criteria based on the use of the scattering measure of classified objects. The areas of existence of the solution are determined. Statements and properties are proved for the Fisher type informativeness criterion, using which the proposed analytical method for determining IAS guarantees the optimality of results in the sense of maximizing the selected functional. The relevance of choosing this type of informativeness criterion is substantiated. The universality of the method with respect to the type of features is shown. An algorithm for implementing this method is presented. In addition, the paper discussed the dynamics of the growth of information in the world, problems associated with big data, as well as problems and tasks of data preprocessing. The relevance of reducing the dimension of the attribute space for the implementation of data processing and visualization without unnecessary difficulties is substantiated. The disadvantages of existing methods and algorithms for choosing an informative set of attributes are shown.

## 1 Introduction

The total amount of data created in the world is forecast to increase dramatically in the coming years, reaching 175 zettabytes in 2025. The rapid development of digitalization contributes to the ever-growing global datasphere [1].

---

\* Corresponding author: [sb.i.l@mail.ru](mailto:sb.i.l@mail.ru)



**Fig. 1.** Data growth dynamics worldwide from 2010 to 2025 (in zebabytes). Source by: [www.statista.com](http://www.statista.com).

Such data (Big Data) is often created by integrating various data sources that meet different standards. In such cases, unfortunately, we have practically no opportunity to analyze, systematize and track all these data. In the absence of such opportunities, often, the proportion of "dirty" data, such as inaccurate, incomplete, duplicated, contradictory, noisy, useless, increases in proportion to the amount of data.

Typically, such problems are solved using methods and algorithms for data mining (Data Mining). The practical application of Data Mining methods involves a multi-step procedure, where one of the key steps is data preprocessing. At this stage are carried out [2-4]:

- Data cleaning, i.e. elimination of contradictions, omissions, accidental emissions and interference;

- Data integration, i.e. combining data from several possible sources in one repository;

- Data transformation, i.e. data aggregation and compression, attribute discretization and dimensionality reduction, etc.

Modern data arrays, to which one or another Data Mining methods can be applied, can be characterized by a large number of features that form a feature space of large dimension. Therefore, the urgent task is dimensionality reduction of such a space to a dimension that allows data processing and / or visualization without unnecessary difficulties.

To date, a variety of approaches, methods and algorithms have been developed and investigated dimensionality reduction of features [5-36].

All approaches to dimensionality reduction of the original attribute space can be divided into two large classes.

The first class provides for the transformation of attribute space. One of the best known and practiced approaches of this class is the principal component method.

Another class of methods consists in choosing the most informative, useful features and exceptions from the consideration of non-informative features without transforming the original space. Various methods and approaches are used here:

full or truncated search;  
 branches and borders;  
 evolutionary;  
 with random selection.

The analysis shows that many existing methods and algorithms for determining an informative attribute sets (IAS) represent one or another way of partial enumeration and therefore do not guarantee the achievement of an optimal result in the sense of the informativeness criterion used. Moreover, these methods do not take into account many factors associated with the costs of identifying attributes, except for computational costs. In this connection, it is advisable to develop new methods that take into account other costs and would be based on the use of a developed mathematical programming apparatus and give an optimal (or close to it) solution.

The aim of this work is to develop an analytical method of determining the IAS taking into account the resource for criteria based on the use of the scattering measure of classified objects.

## 2 Research methods

Let training sample objects  $X$  be given:

$$X = \bigcup_{p=1}^r X_p, X_p \cap X_q = \emptyset, (p \neq q, p, q = \overline{1, r}), X_p = \{x_{pi} : i = \overline{1, m_p}\}, \quad (1)$$

where  $X_p$  ( $p = \overline{1, r}$ ) –classes,  $r$  –number of classes,  $x_{pi}$  – object  $pth$ – class,  $m_p$ – number of objects of the  $pth$  –class.

Then the total number of training sample objects is:

$$M = \text{card}(X) = \sum_p \text{card}(X_p) = \sum_p m_p.$$

Suppose that each object in the training sample is a dimensional vector of features, then  $\forall x_{pi} = (x_{pi}^1, x_{pi}^2, \dots, x_{pi}^N) \in X_p$  and  $\dim(X) = N$ .

For dimensionality reduction of the initial feature space and selection of the IAS, we use the  $\ell$ -dimensional vector  $\lambda = (\lambda^1, \lambda^2, \dots, \lambda^N)$ , i.e.

$$\lambda: X \rightarrow X|_{\lambda} = \{x | x = (\lambda^1 x^1, \lambda^2 x^2, \dots, \lambda^N x^N)\} \quad (2)$$

where  $\lambda^j \in \{0, 1\}$  ( $j = \overline{1, N}$ ).

## 3 Results

Here  $\lambda^j = 0$  indicates the absence or  $\lambda^j = 1$  the presence  $j$ -th feature in this set.

Definition 1[9]. A vector  $\lambda = (\lambda^1, \lambda^2, \dots, \lambda^N)$  is called  $\ell$ informative, if  $\sum_{j=1}^N \lambda^j = \ell$ .

The set of all  $\ell$  informative vectors is denoted by  $\Lambda^{\ell}$ :

$$\Lambda^{\ell} = \left\{ \lambda \mid \sum_{j=1}^N \lambda^j = \ell, \lambda^j \in \{0, 1\}, j = \overline{1, N} \right\}. \quad (3)$$

We see that the cardinality of the set  $\Lambda^{\ell}$  equals to

$$|\Lambda^{\ell}| = C_N^{\ell} = \frac{N!}{\ell!(N-\ell)!} \quad (4)$$

From (2) and (3) follows

$$\Lambda^\ell: X \rightarrow X|_{\Lambda^\ell} = \{X|_\lambda \mid \lambda \in \Lambda^\ell\}. \quad (5)$$

Taking into account that the task of determining the optimal IAS is usually associated with assessing the quality of classification, we will use the functional  $I(\lambda)$  as an informativeness (effectiveness) criterion to select the optimal IAS from (5).

Definition 2. We say that a system (or IAS)  $X|_\lambda \in X_{\Lambda^\ell}$  is optimal, if  $\exists \lambda \in \Lambda^\ell$  for which  $I(\lambda) = \text{extr}_{\mu \in \Lambda^\ell} I(\mu)$  is true.

Then the problem of determining an informative attribute set can be reduced to an optimization problem

$$\begin{cases} I(\lambda) \rightarrow \text{extr} \\ \lambda \in \Lambda^\ell \end{cases} \quad (6)$$

Suppose the definition of an object, i.e. each feature of an object requires certain costs (technical, computational, time, etc.) and a fixed resource is allocated to determine the object.

To denote the above, we introduce the vector  $c = (c^1, c^2, \dots, c^N)$ , where  $c^j$  – indicates the total costs for determining the attribute of  $j$ th attribute ( $j = \overline{1, N}$ ) and let  $c_0$  – a fixed resource allocated to determine the object.

If we take into account that each  $\lambda \in \Lambda^\ell$  vector uniquely determines a specific IAS, then  $(c, \lambda)$  will show how much a resource is required to get this IAS.

Can we get this informative attribute set? If  $(c, \lambda) \leq c_0$ , then yes.

Now we can formulate a general mathematical statement of the problem of choosing a set of the most informative features with limited resources as follows:

$$\begin{cases} I(\lambda) \rightarrow \text{extr} \\ \lambda \in \Lambda^\ell \\ (c, \lambda) \leq c_0 \end{cases} \quad (7)$$

where  $(*, *)$  – scalar product of vectors.

If it is not specifically known about objects and classes, except for the value of the features of the object, then we can assume the following:

Hypothesis 1. If the training sample has the form (1), then the value of features between objects of the same class is more similar (close) than the value of different classes.

Hypothesis 2. If hypothesis 1 is true, then the optimal IAS brings together objects of one class and separates objects of different classes better than others.

Based on the above hypotheses of the informativeness (effectiveness) criterion for the selection of IAS, we define in the form of a functional as follows [9]:

$$I(\lambda) = \frac{(a, \lambda)}{(b, \lambda)} \quad (8)$$

where  $I(\lambda)$  – is a homogeneous functional of the Fisher type,  $a = (a^1, a^2, \dots, a^N)$  – interclass distance,  $b = (b^1, b^2, \dots, b^N)$  – intraclass distance.

The coefficients  $a^j, b^j$  are independent of  $\lambda$  and are calculated in advance.

For the given functional, it is considered that the IAS is optimal if the functional values are greater. Among the main advantages of the functional, its relative simplicity should be highlighted.

On the other hand, the simplicity of the Fisher Functional acts as a disadvantage, since here one can "overlook" the complex nonlinear properties of the classes being analyzed.

However, in favor of this functionality is the fact that simple quality criteria, as a rule, turn out to be more reliable, i.e. if not the most informative, then at least a fairly informative subsystem of attributes is distinguished. Conversely, complex criteria, which in most cases allow more informative subsystems, still allow you to select subsystems for which it is difficult to construct a decisive rule.

Then (7) will have the following form

$$\begin{cases} I(\lambda) = \frac{(a,\lambda)}{(b,\lambda)} \rightarrow \max \\ \lambda \in \Lambda^\ell \\ (c,\lambda) \leq c_0 \end{cases} \quad (9)$$

Before proceeding to solution (9), we first determine the field of existence of the solution.

We will analyze the resource vector  $c = (c^1, c^2, \dots, c^N)$ . There is a sequence of pairwise different indices  $j_1, j_2, \dots, j_N$  such that

$$c^{j_1} \leq c^{j_2} \leq \dots \leq c^{j_N} \quad (10)$$

We construct numerical sequences, taking into account (10), as follows

$$f_1 = \sum_{i=1}^{\ell} c^{j_i}, f_2 = \sum_{i=2}^{\ell+1} c^{j_i}, \dots, f_k = \sum_{i=k}^{\ell+k-1} c^{j_i}, \dots, f_{N-\ell+1} = \sum_{i=N-\ell+1}^N c^{j_i} \quad (11)$$

## 4 Discussion

Proposition 1. From the numerical sequence (11) the following are obtained:

if  $f_1 > c_0$ , then (9) has no solutions;

if  $f_1 \leq c_0$ , then (9) has at least one solution;

if  $\exists t > \ell$  such that  $f_1 - c^{j_t} + c^{j_t} \leq c_0 < f_1 - c^{j_t} + c^{j_{t+1}}$ , then the features of the objects corresponding to the indices  $j_{t+1}, j_{t+2}, \dots, j_N$  will not participate in the IAS and should be excluded from further consideration, i.e. these features should be an exception from the attribute space.

Proof.

1. Let  $f_1 > c_0$ . Then  $\forall \lambda \in \Lambda^\ell$  vector  $(c, \lambda) \geq c_0$ . Therefore (9) has no solutions ■

2. Let  $f_1 \leq c_0$ . Consider vector  $\lambda = (\lambda^1, \lambda^2, \dots, \lambda^N)$ , where  $\lambda^j = \begin{cases} 1, & \text{if } j \in \{j_1, j_2, \dots, j_N\} \\ 0, & \text{otherwise} \end{cases}$ .

Then  $\lambda \in \Lambda^\ell$  and  $(c, \lambda) \leq c_0$ . Therefore (9) has no solutions ■

3. Suppose that conditions 3) of Proposition 1 are satisfied. If  $\lambda \in \Lambda^\ell$  and  $\sum_{i=t+1}^N \lambda^{j_i} > 0$ , then  $(c, \lambda) \geq c_0$  ■

We consider the following optimization problem

$$\begin{cases} I(\lambda) = \frac{(a,\lambda)}{(b,\lambda)} * s(\lambda) \rightarrow \max \\ \lambda \in \Lambda^\ell \end{cases}, \quad (12)$$

$$\text{where } s(\lambda) = \begin{cases} -1, & c_0 - (c, \lambda) < 0 \\ 1, & c_0 - (c, \lambda) \geq 0 \end{cases}$$

By analogy [\*\*\*] to solve problem (12), we introduce a vector function

$$\varphi(\lambda) = a(b, \lambda) - b(a, \lambda), \quad (13)$$

which indicates the direction of the fastest growth of the functional  $I(\lambda)$  at the point  $\lambda$ .

Let  $b^j > 0$  ( $j = \overline{1, N}$ ) and  $\lambda, \mu \in \Lambda^\ell$ .

We define the connections between (12) and (13).

Let  $(\varphi(\lambda), \mu) \geq 0$ . Then

$$\begin{aligned} (a(b, \lambda) - b(a, \lambda), \mu) \geq 0 &\Rightarrow (a, \mu)(b, \lambda) - (b, \mu)(a, \lambda) \geq 0 \Rightarrow \\ &\Rightarrow (a, \mu)(b, \lambda) \geq (b, \mu)(a, \lambda) \Rightarrow \frac{(a, \mu)}{(b, \mu)} \geq \frac{(a, \lambda)}{(b, \lambda)}. \end{aligned} \quad (14)$$

This implies the following

Property 1. If  $(\varphi(\lambda), \mu) \geq 0$ , then

$I(\mu) \leq I(\lambda) \leq 0$  for  $\mathbf{s}(\lambda) = -\mathbf{1}, \mathbf{s}(\mu) = -\mathbf{1}$ ;

$I(\mu) \geq -I(\lambda) \geq 0$  for  $\mathbf{s}(\lambda) = -\mathbf{1}, \mathbf{s}(\mu) = \mathbf{1}$ ;

$I(\mu) \leq -I(\lambda) \leq 0$  for  $\mathbf{s}(\lambda) = \mathbf{1}, \mathbf{s}(\mu) = -\mathbf{1}$ ;

$I(\mu) \geq I(\lambda) \geq 0$  for  $\mathbf{s}(\lambda) = \mathbf{1}, \mathbf{s}(\mu) = \mathbf{1}$ .

Let  $I(\mu) \geq I(\lambda)$ . Then

$$\begin{aligned} \frac{(a, \mu)}{(b, \mu)} s(\mu) \geq \frac{(a, \lambda)}{(b, \lambda)} s(\lambda) &\Rightarrow (a, \mu)(b, \lambda)s(\mu) - (a, \lambda)(b, \mu)s(\lambda) \geq 0 \Rightarrow \\ &\Rightarrow (a(b, \lambda)s(\mu) - b(a, \lambda)s(\lambda), \mu) \geq 0. \end{aligned} \quad (15)$$

This implies the following

Property 2. If  $I(\mu) \geq I(\lambda)$ , then

$(\varphi(\lambda), \mu) \leq 0$  for  $\mathbf{s}(\lambda) = -\mathbf{1}, \mathbf{s}(\mu) = -\mathbf{1}$ ;

$(a(b, \lambda) + b(a, \lambda), \mu) \geq 0$  for  $\mathbf{s}(\lambda) = -\mathbf{1}, \mathbf{s}(\mu) = \mathbf{1}$ ;

$(a(b, \lambda) + b(a, \lambda), \mu) \leq 0$  for  $\mathbf{s}(\lambda) = \mathbf{1}, \mathbf{s}(\mu) = -\mathbf{1}$ ;

$(\varphi(\lambda), \mu) \geq 0$  for  $\mathbf{s}(\lambda) = \mathbf{1}, \mathbf{s}(\mu) = \mathbf{1}$ .

Proposition 2. If  $s(\lambda) = s(\mu) = 1$ , then  $I(\mu) \geq I(\lambda)$  if and only if  $(\varphi(\lambda), \mu) \geq 0$ .

Proof.

$\Rightarrow$ ) follows from d) property 2.

$\Leftarrow$ ) follows from d) property 1 ■

For the components of the vector function  $\varphi(\lambda)$  there exists a sequence of pairwise distinct indices  $\exists t_1, t_2, \dots, t_N$  such that  $\varphi(\lambda)^{t_1} \geq \varphi(\lambda)^{t_2} \geq \dots \geq \varphi(\lambda)^{t_N}$ .

Proposition 3. If for the vector  $\mu(\lambda): \mu^{t_1} = \mu^{t_2} = \dots = \mu^{t_\ell} = 1$  и  $\mu^{t_{\ell+1}} = \dots = \mu^{t_N} = 0$ , then  $\mu(\lambda) \in \Lambda^\ell$  and  $(\varphi(\lambda), \mu(\lambda)) = \max\{(\varphi(\lambda), \eta) \mid \eta \in \Lambda^\ell\}$ .

Proposition 4.  $\forall \lambda \in \Lambda^\ell$  is true  $(\varphi(\lambda), \mu(\lambda)) \geq 0$ .

Proof.

$$(\varphi(\lambda), \mu(\lambda)) = (a(b, \lambda) - b(a, \lambda), \mu(\lambda)) \geq (a(b, \lambda) - b(a, \lambda), \lambda) = (a, \lambda)(b, \lambda) - (b, \lambda)(a, \lambda) = 0. \quad \blacksquare$$

Let there exist solutions for problem (9), then we have the following:

Proposition 5.  $\forall \lambda \in \Lambda^\ell$  there exists  $\exists s(\mu(\lambda)) = 1$  and  $I(\mu(\lambda)) \geq I(\lambda)$ .

Theorem. If  $s(\mu(\lambda)) = 1$  and  $I(\lambda) = I(\mu(\lambda))$ , then  $I(\lambda) = \max\{I(\eta) \mid \eta \in \Lambda^\ell\}$ .

Proof.

Let  $s(\mu(\lambda)) = 1$  and  $I(\lambda) = I(\mu(\lambda))$ . Then property 2 implies that  $s(\lambda) = 1$ .

Proposition 2 implies  $(\varphi(\lambda), \mu(\lambda)) = 0$ . Hence, according to statement 3

$$0 = (\varphi(\lambda), \mu(\lambda)) = \max\{(\varphi(\lambda), \eta) \mid \eta \in \Lambda^\ell\} \Rightarrow (\varphi(\lambda), \eta) \leq 0 \text{ for } \forall \eta \in \Lambda^\ell \Rightarrow I(\lambda) \leq I(\eta). \quad \blacksquare$$

Note that the theorem guarantees the optimality of the obtained solution, i.e. the value of the functional  $I(\lambda)$  for the found solution  $\lambda$  reaches its maximum on the set  $\Lambda^\ell$ .

So, the proposed method is as follows. At the first step, an arbitrary  $\lambda \in \Lambda^\ell$  vector is chosen, for example,

$$\lambda = (\underbrace{1, 1, \dots, 1}_\ell, 0, \dots, 0)$$

As  $\mu(\lambda)$  we take

$$\max\{\eta(\lambda) \mid (\varphi(\lambda), \eta(\lambda)) \geq 0, s(\eta(\lambda)) = 1, \eta(\lambda) \in \Lambda^\ell\}. \quad (16)$$

Further, at each step, the vector  $\lambda$  is replaced by  $\mu(\lambda)$ , satisfying conditions (16), as long as the functional  $I(\lambda)$  increases, i.e. until conditions  $I(\lambda) = I(\mu(\lambda))$  are met. As soon as the functional stops increasing, the calculation ends.

We formulate the considered method in the form of an algorithm:

Step 1. Input parameters are the values:  $\ell$ - the required number of features;  $N$  is the number of features;  $c, b, c$  are  $N$ -dimensional vectors;  $c_0$  is the allocated resource.

Step 2. Setting the vector:  $\lambda = (\underbrace{1, 1, \dots, 1}_\ell, 0, \dots, 0)$ ;

Step 3. Calculation:  $I(\lambda)$ ;

Step 4. Definition:  $\mu(\lambda) = \max\{\eta(\lambda) \mid (\varphi(\lambda), \eta(\lambda)) \geq 0, s(\eta(\lambda)) = 1, \eta(\lambda) \in \Lambda^\ell\}$ ;

Step 5. Calculation:  $I(\mu(\lambda))$ .

Step 6. Comparison:  $I(\mu(\lambda)) > I(\lambda)$ . If the inequality holds, then we set  $\lambda = \mu(\lambda)$ ,  $I(\lambda) = I(\mu(\lambda))$  and we go to Step 4. Otherwise, the procedure ends and  $\lambda$  is the optimal solution.

Step 7. Output parameters:  $\lambda, I(\lambda)$ .

The following is a method for determining the measure of similarity (proximity) or the scattering measure for objects of the training sample.

Let the features of objects take binary (i.e.  $\{0,1\}$ ) and/or or continuous (on the interval  $[\alpha, \beta]$ ) values. We consider both cases in more detail.

As noted above,  $a = (a^1, a^2, \dots, a^N)$  is the interclass distance and  $b = (b^1, b^2, \dots, b^N)$  is the intraclass distance.

Let  $j_k$ th ( $j_k \in \{1, 2, \dots, N\}$ ) the attribute takes continuous values on the interval, i.e.  $x^{j_k} \in [\alpha_{j_k}, \beta_{j_k}] \subset R$  (here  $R$  is the set of real numbers). For such features, the Euclidean metric can be used as the scattering measure. Then  $a^{j_k}$  and  $b^{j_k}$  is defined as follows:

$$a^{j_k} = \sum_{p,q=1}^r (\bar{x}_p^j - \bar{x}_q^j)^2,$$

$$b^{j_k} = \sum_{p=1}^r \left[ \frac{1}{m_p} \sum_{i=1}^{m_p} (\bar{x}_p^j - x_{pi}^j)^2 \right],$$

$$\bar{x}_p^j = \frac{1}{m_p} \sum_{i=1}^{m_p} x_{pi}^j,$$

where  $\bar{x}_p = (\bar{x}_p^1, \bar{x}_p^2, \dots, \bar{x}_p^N)$  is averaged object of class  $X_p$  ( $p = \overline{1, r}$ ).

Let  $t_k \in \{1, 2, \dots, N\}$  the attribute takes binary values, i.e.  $x^{t_k} \in \{0,1\}$ . Then  $a^{t_k}$  and  $b^{t_k}$  can be defined as follows:

$$a^{t_k} = \sum_{p,q=1}^r \left[ \left| \frac{m_p^{t_k(0)}}{m_p} - \frac{m_q^{t_k(1)}}{m_q} \right| * \left| \frac{m_p^{t_k(1)}}{m_p} - \frac{m_q^{t_k(0)}}{m_q} \right| \right],$$

$$b^{t_k} = \sum_{p=1}^r \frac{m_p^{t_k(0)} * m_p^{t_k(1)}}{m_p},$$

where  $m_p^{t_k}(i)$  is the number of objects of class  $X_p$  ( $p = \overline{1, r}$ ), that takes values  $i$ .

## 5 Conclusion

The main result of this work is to solve the scientific problem of choosing of informative attribute set, taking into account the resource in the problem of pattern recognition and data preprocessing. In the course of solving this problem, the following work was carried out:

- studied the main causes and types of big data problems;
- the main tasks of data preprocessing are given;
- shown the need of dimensionality reduction of the attribute space and the shortcomings of existing methods for the selection of IAS;
- an effective analytical method for determining IAS was developed and investigated taking into account the resource;
- an approach is proposed for determining the area of existence of a solution for the problem of choosing an IAS with limited resource;
- it is proved that the proposed method for determining the IAS guarantees the optimality of the results in the sense of maximizing the selected functional;
- the relevance of the choice of this type of informativeness criterion is substantiated;
- shown the universality of the method with respect to the type of features;
- given an algorithm for implementing this method.

## References

1. <https://www.statista.com/statistics/871513/worldwide-data-created/>
2. J. Han, M. Kamber, J. Pei, Data mining : concepts and techniques (3rd ed. by Elsevier Inc., USA, 2012)
3. A.V. Zamyatin, Introduction to Data Mining (Tomsk : Pub.house: Tomsk State University, 2016)
4. J.L. Zhou, F.Chen, Springer, Human-Computer interaction Series, Switzerland, 482 (2018)
5. N.G. Zagoruyko, Applied methods of data and knowledge analysis (Novosibirsk: IM SO RAN, 1999)
6. M. Nasma, Moscow Economic Journal, **6**, (2019)
7. A.X. Nishanov, B.B. Akbaraliev, O.B. Ruzibaev, O.K. Xujaev, Chemical Technology, Control and Management, International Scientific and Technical Journal, **6**, 78-82 (2014)
8. M.M. Kamilov, A.H. Nishanov, B.B. Akbaraliev, Proceedings of ICEIC2008, June 24-27, 2008, Tashkent, pp. 476-478 (2017)
9. M.M. Kamilov, A.H. Nishanov, B.B. Akbaraliev, *Methods of forming of optimal sign space for object recognition in the class of logic-heuristic algorithms*, Fourth World Conference on Intelligent Systems for Industrial Automation - WCIS 2006, Tashkent (2006)
10. B.B. Akbaraliev, Bulletin of TUIT, **2**, 47-50 (2007)
11. A.T. Raxmanov, B.B. Akbaraliev, A.K. Ergashev, Problems of Informatics and Energy Uzbek journal, **1**, 76-79 (2011)
12. S. Beniwal, J. Arora, International Journal of Engineering Research & Technology (IJERT), **1**, 6, (2012)
13. H. Liu, J. Li, L. Wong, Genome Informatics, **13**, 51-60 (2002)



14. G. S. Lbov, Methods for processing heterogeneous experimental data (Novosibirsk: Nauka, Sib.otd., 1981)
15. Yu.I. Zhuravlev, Selected scientific works (Publishing house: Master, Moscow, 1998)
16. L. Zhang, M. Luo, J. Liu, Z. Li, Q. Zheng, Pattern Recognition Letters **130**, 275-283 (2020)
17. D. Santra, S.K. Basu, J.K. Mandal, S. Goswami, Expert Systems with Applications **145**, 113084 (2020)
18. Y. Xiong, R. Zuo, Computers and Geosciences **140**, 104484 (2020)
19. J. Gai, J. Shen, H. Wang, Y. Hu, Shock and Vibration, **2020**, 4294095 (2020)
20. P.S. Raja, K. Thangavel, Soft Computing **24(6)**, 4361-4392 (2020)
21. D. Wang, F. Tian., S.X. Yang, D. Jiang, B. Cai, Sensors (Switzerland) **20(3)**, 874 (2020)
22. P. Lou, A. Jimeno Yepes, Z. Zhang, C. Li, J. Wren, Bioinformatics **36**, 2, 611-620 (2020)
23. S. Fu, X. Liu, Eurasip Journal on Wireless Communications and Networking **2020**, 1, 8 (2020)
24. D. Wei, T. Chen, S. Li, Y. Zhao, T. Li, Eurasip Journal on Advances in Signal Processing **2020**, 1, 20 (2020)
25. M. Ala'raj, M. Majdalawieh, M.F. Abbod, Journal of Big Data, **7**, 1, 15 (2020)
26. G. Mishra, V.P. Vishwakarma, A. Aggarwal, Optics and Laser Technology **129**, 106232 (2020)
27. T.C.G. Kibbey, R. Jabrzemski, D.M. O'Carroll, Chemosphere **252**, 126593 (2020)
28. Z. Shen, Z. Man, Z. Cao, J. Zheng, Computers and Electrical Engineering **84**, 106641 (2020)
29. A. Nishanov, G. Djurayev, M. Khasanova, Compusoft: An International Journal of Advanced Computer Technology, **8(6)**, 3158-3165 (2019)
30. M. Kamilov, A. Nishanov, R. Beglerbekov, International Journal of Innovative Technology and Exploring Engineering **8(6)** (2019)
31. A. Nishanov, E. Avazov, B. Akbaraliyev, International Journal of Innovative Technology and Exploring Engineering **8(6)**, 696-698(2019)
32. E. Emary, H. Zawbaa, A. Hassanien, Neurocomputing, **213** (2016) DOI 10.1016/j.neucom.2016.03.101
33. Z. Yong, G. Dun-wei, Z. Wan-qiu, Neurocomputing, **171** (2016) DOI 10.1016/j.neucom.2015.07.057
34. Y. Zhang, D. Gong, X. Sun, Y. Guo, Scientific Reports, **7 (1)** (2017) DOI 10.1038/s41598-017-00416-0
35. A.Kh. Nishanov, B.B.Akbaraliyev, G.P.Juraev, M.A. Khasanova, M.Kh. Maksudova, Z.F. Umarova, J.Cardiovasc Disease Res., **11(2)**, 218-223 (2020)
36. A.Kh. Nishanov, B.B. Akbaraliyev, B.S. Samandarov, O.K. Akhmedov, S.K. Tajibaev, Webology, **17**, 1, 341-364 (2020) DOI: 10.14704/WEB/V17I1/WEB17009