Research Article

Criteria Study in Solving Data Science Classification Problems

Akhram Khasanovich Nishanov[1], Bakhtiyorjon Bakirovich Akbaraliev[2], Utkir Sheralievich Achilov[3], Saidrasulov Sherzod Norboy o'g'li[4], Tulkinjon Gaybullaevich Rayimov[5], Botir Tokhirjonovich Sobirov[6]

[1]DSc, Professor, Department of Software of Information Technologies, Tashkent University of Information Technologies named after Muhammad Al-Khwarizimi, Tashkent, Uzbekistan.
[2]PhD, Docent, Dean of Software Engineering Faculty, Tashkent University of Information Technologies named after Muhammad Al-Khwarizimi, Tashkent, Uzbekistan.
[3]Docent, Head of a department, The Academy of the Armed Forces of the Republic of Uzbekistan.
[4]Chief specialist, Center for research development of higher education and implementation of advanced technologies under the Ministry of higher and secondary specialized education of the Republic of Uzbekistan.
[5]Senior teacher, The Academy of the Armed Forces of the Republic of Uzbekistan.
[6]Senior teacher, The Academy of the Armed Forces of the Republic of Uzbekistan.

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 20 April 2021

Abstract: The given article examines and analyzes DATA SCIENCE, in particular, the methods and algorithms used in solving the issues of forming information systems describing objects, classification and clustering, as well as information content criteria. As a result, it was discovered that the issues of classification, clustering, and character space reduction were poorly studied in a comprehensive manner and in resource-constrained conditions. A number of criteria evaluating the efficiency, information value and reliability indicators of algorithms and methods which are widely used in solving the issues of classification, clustering and character space reduction have been thoroughly studied in this article.

Keywords: attribute (feature), pattern recognition, informative attribute set, information value criterion, dimensionality reduction, the scattering measure, informative vector, attribute (feature) space.

1. Introduction

The main objective of reducing the size of the character space in intellectual data analysis, that is, developing information systems that describe objects, is to increase efficiency, reduce workload, in particular, computation and labour costs.

It is advisable to develop or improve existing methods and algorithms that are easy to interpret and comfortable to use to achieve the set goal.

Let $X = \bigcup_{p=1}^{r} X_p, X_p \cap X_q = \emptyset, p \neq q, \quad p,q = \overline{1,r}$ is given to us,
Where

$$X_1 = \{x_{11}, x_{12}, \dots, x_{1m_1}\},$$
$$X_2 = \{x_{21}, x_{22}, \dots, x_{2m_2}\},$$
$$\dots\dots\dots\dots\dots\dots\dots\dots$$
$$X_r = \{x_{r1}, x_{r2}, \dots, x_{rm_r}\}$$

Here $x_{pi} = (x_{pi}^1, x_{pi}^2, \dots, x_{pi}^N), i = \overline{1, m_p}$.

Hypothesis 1. (Nishanov et al., 2019; Nishanov et al., 2019;Kamilov et al., 2019; Nishanov et al., 1999; Nishanov et al., 1999; Nishanov et al., 2016; Nishanov et al., 2020)If the study sample is defined in the above type, then objects belonging to the same class are close to each other (similar) to the ones which belong to different classes.

Hypothesis 2(Nishanov et al., 2002; Nishanov et al., 2020; Nishanov et al., 2020; Nishanov et al., 2020; Nishanov et al., 2020; Nishanov et al., 2020; Nishanov et al., 2020). Characters that optimally describe the objects of study sample bring objects of the same class closer than other characters, and removes objects belonging to different classes.

On the assumption of these hypotheses, the study of information criteria plays an significant role in the formation ofℓ information systems describing objects and assessing their quality on the basis of a given sample of training.

Some concepts and definitions

Let's assume we are given $\Omega = \{\omega\}$ set of objects, and each $\omega \in \Omega$object is defined with Nsign.
These characters represent the properties, features, and other descriptions of a given object.

So, for $\forall\omega \in \Omega$ there is a set of characters $\exists(\alpha_1, \alpha_2, \ldots, \alpha_N)$ and they perfectly define each other, that is, $\omega \leftrightarrow (\alpha_1, \alpha_2, \ldots, \alpha_N)$.

Generally, each α_i $(i = \overline{1,N})$character may be admitted in various values. For example (Nishanov et al., 2019; Nishanov et al., 2019;Kamilov et al., 2019; Nishanov et al., 1999; Nishanov et al., 1999; Nishanov et al., 2016; Nishanov et al., 2020; Nishanov et al., 2002; Nishanov et al., 2020; Nishanov et al., 2020; Nishanov et al., 2020; Nishanov et al., 2020; Nishanov et al., 2020; Nishanov et al., 2020),

a) $\alpha_i \in \{0,1\}$, means there is a suitable property for the i character of$\alpha_i = 1$object, but$\alpha_i = 0$ vice verse;

b) $\alpha_i \in \{0,1,-\}$, here in addition to the abovementioned means there is no information about $i$character within the object $\alpha_i = " - "$;

c) $\alpha_i \in \{1,2, \ldots, K\}$. Here the value that a character receives describes the level of expression of the i mark of an object;

d) $\alpha_i \in (a,b) \subset Ror\alpha_i \in [a,b] \subset R$;

e) $\alpha_i \in \{\mu\}$ – set of probability measurements and etc.

f) Let's mark the set of values with $D_i(i = \overline{1,N})$ that i character of an object may receive. Then $D = D_1 \times D_2 \times \ldots \times D_N$ comprises the character space that define the objects, here $\dim(D) = N$

So, it is possible to imagine each object as multidimensional (in particular N dimensional) vector, i.e.$\omega = (\alpha_1, \alpha_2, \ldots, \alpha_N)$.

Definition 1. $D_i(i = \overline{1,N})$is called the character alphabet.

Definition 2. (Nishanov et al., 1999; Nishanov et al., 2016; Nishanov et al., 2020; Nishanov et al., 2002; Nishanov et al., 2020). $\omega \in \Omega$ **Is called permitted object, if it is** $\alpha_i \in D_i(i = \overline{1,N})$

Suppose that a set of Ω objects is divided into sets that do not intersect according to a certain rule

$$\Omega = \bigcup_{p=1}^{r} \Omega_p\, , \Omega_p \cap \Omega_q = \emptyset, p \neq q, \quad p,q = \overline{1,r}.$$

For the study of this set, its sub-sets and their constituent objects, experts have given $X \subset \Omega$ selection, for this selection the following may be assumed appropriate

$$X = \bigcup_{p=1}^{r} X_p\, , X_p \cap X_q = \emptyset, p \neq q, \quad p,q = \overline{1,r},$$

$$X_p = \{x_{pi} = (x_{pi}^1, x_{pi}^2, \ldots, x_{pi}^N) : i = \overline{1,m_p}\} \subset D,$$

here$X_p \subset \Omega_p(p = \overline{1,r})$, m_p and the number of objects in X_p.

The X set determined in this method is called study sample, $X_p(p = \overline{1,r})$and is called class.

On the assumption of the research purposes, an object or set of objects that represents the class $X_p(p = \overline{1,r})$ with sufficient accuracy is called a reference object or standard of that class. The set of all class standards forms the benchmark table for the Xstudy sample.

Below we define the functions of proximity (similarity) between objects and classes, which is one of the most important concepts in DATA SCIENCE. In general, proximity functions are usually defined in relation to the values that object symbols can accept, i.e., the alphabet of characters $D_i(i = \overline{1,N})$.

Explanation 2. While X does not require additional precision for study sample, in order to simplify the designations, X is sometimes used on its own to define character space.

2. Criteria study

As stated above, classification in DATA SCIENCE is important, and a number of approaches, methods, and algorithms have been developed to address this issue, at the same time this process is still ongoing. This is due, firstly, to the fact that there is not any single and mathematically proved method for solution of this issue yet, and secondly, since all available methods are based on heuristic or statistical approaches, their quality, efficiency, reliability, etc. strongly dependent on the objects of the research filed, in particular, on the characteristics of the practical case, which were resolved on the basis of certain additional conditions.

Below are a number of criteria that assess the effectiveness, information value, and reliability of the decisive rules for the classification issue.

Below are a number of criteria that assess the effectiveness, information value, and reliability of the decisive rules for the classification issue.

It is known that the most important criterion in the matter of classification is the quality and information value of the decisive rule.

Suppose X study sample and any $\psi: X \to \{0,1\}$ preidcate function.

Definition 3. If $x \in X$ for it is $\psi(x) = 1$, then ψ predicate separates *x* object or covers it.

Below we will include marking:

$$X|_{\psi}^{1} = \{x \in X: \psi(x) = 1\}, X|_{\psi}^{0} = \{x \in X: \psi(x) = 0\}$$

$$X_p\Big|_{\psi}^{1} = \{x \in X_p: \psi(x) = 1\}, X_p\Big|_{\psi}^{0} = \{x \in X_p: \psi(x) = 0\}, p = \overline{1,r}.$$

Definition 4. If for $\exists(!)q \in \{1,2,\dots,r\}$ the $card\left(X_q\Big|_{\psi}^{1}\right) \gg card\left(X_q\Big|_{\psi}^{0}\right)$ and $card\left(X_q\Big|_{\psi}^{1}\right) \gg card\left((X\backslash X_q)\Big|_{\psi}^{1}\right)$ is proper, then ψ predicate is called standard for X_q class.

Definiton 5. if it is$card\left(X_q\Big|_{\psi}^{1}\right) \gg card\left(X_q\Big|_{\psi}^{0}\right)$, then ψregularity is considered optimal forX_q.

Laws that can be classified by simple logical formulas are important, they are called rules. The process of searching for such rules through the study of a given study sample is called knowledge discovery. The main requirement here is that the extracted knowledge should be clear to the user.

It should be noted that, as a rule, any law classifies only a certain part of the objects presented. Therefore, by combining or generalizing several laws, it will be possible to classify all the objects presented for research.

Rules frequently used to resolve the issues of classification

Boundary terms (decision stump):

$$\psi(x) = \begin{cases} 1, & x^j \le a_j \\ 0, & x^j > a_j \end{cases} \text{ or } \psi(x) = \begin{cases} 1, & a_j \le x^j \le b_j \\ 0, & otherwise \end{cases}.$$

1. Decision stump conjunction:

$$\psi(x) = \begin{cases} 1, & \bigwedge_{j \in J}[a_j \le x^j \le b_j] \\ 0, & otherwise \end{cases}.$$

2. In $(d = |J|$syndromeit is conjunction, in $d = 1$it becomes disjunction):

$$\psi(x) = \begin{cases} 1, & \sum_{j \in J}[a_j \le x^j \le b_j] \ge d \\ 0, & otherwise \end{cases},$$

thesea_j, b_j, J, dparameters are customized through optimization of information value criterion on study sample.

3. Half plane - linear boundary function:

$$\psi(x) = \begin{cases} 1, & \sum_{j \in J} \omega_j x^j \ge \omega_0 \\ 0, & otherwise \end{cases}$$

4. Sphere – function of boundary proximity:

$$\psi(x) = \begin{cases} 1, & r(x, x_0) \le \omega_0 \\ 0, & otherwise \end{cases}.$$

So, for class$X_p (p = \overline{1,r})$any ψ law is determined through the following

$$\begin{cases} t(\psi) = card\left(X_p\Big|_{\psi}^{1}\right) \to max \\ n(\psi) = card\left((X\backslash X_p)\Big|_{\psi}^{1}\right) \to min \end{cases} \quad (1)$$

Here $t(\psi)$ is the number of objects which are properly classified through $-\psi$, and $n(\psi)$ – is vice verse, that is$x \overline{\in} X_p$, but $\psi(x) = 1$.

$t(\psi)$ and $n(\psi)$ determined as (1) is considered information value criterion for ψlaw.

Let's introduce the following marking

$$E(\psi) = \frac{n(\psi)}{t(\psi) + n(\psi)}, \qquad D(\psi) = \frac{t(\psi)}{m_p}.$$

Definition 6. $\psi(x)$ predicate is regarded ε, δ logical law for $X_p$class, if it is$E(\psi) \le \varepsilon$and $D(\psi) \ge \delta$, here it becomes$\varepsilon, \delta \in [0,1]$.

If it is$n(\psi) = 0$, thenψlaw is pure or non-contradictory, or else it is called partly pure or non-contradictory.

In general, using $t(\psi)$ and $n(\psi)$, it is possible to form qualitative criteria for crucial as follows:

1) $I(t(\psi),n(\psi))=\frac{t(\psi)}{n(\psi)+1}\to max$

2) $I(t(\psi),n(\psi))=\frac{t(\psi)}{t(\psi)+n(\psi)}\to max$

3) $I(t(\psi),n(\psi))=t(\psi)-n(\psi)\to max$

4) $I(t(\psi),n(\psi))=t(\psi)-Cn(\psi)\to max,$ бунда $C-const.$

5) $I(t(\psi),n(\psi))=\frac{t(\psi)}{m_p}-\frac{n(\psi)}{m-m_p}\to max$

6) $IGain\big(t(\psi),n(\psi)\big)=h\left(\frac{m_p}{m}\right)-\frac{t(\psi)+n(\psi)}{m}h\left(\frac{t(\psi)}{t(\psi)+n(\psi)}\right)-$

$$-\frac{m-t(\psi)-n(\psi)}{m}h\left(\frac{m_p-t(\psi)}{m-t(\psi)-n(\psi)}\right)\to max,$$

бу ерда $h(z)=-z\log_2 z-(1-z)\log_2(1-z)$.

7) While there is Gini (Gini) criterion: $Gini\big(t(\psi),n(\psi)\big)=IGain\big(t(\psi),n(\psi)\big)$, here $h(z)=4z(1-z)$.

8) (Ficher's Exact Test):

$IStat\big(t(\psi),n(\psi)\big)=-\frac{1}{m}\log_2\frac{C_{m_p}^{t(\psi)}C_{m-m_p}^{n(\psi)}}{C_m^{t(\psi)+n(\psi)}}\to max$, бунда $C_n^k=\frac{n!}{k!(n-k)!}$.

9) Boosting criterion:

$I\big(t(\psi),n(\psi)\big)=\sqrt{t(\psi)}-\sqrt{n(\psi)}\to max,$

10) normal boosting criterion: $I\big(t(\psi),n(\psi)\big)=\sqrt{\frac{t(\psi)}{m_p}}-\sqrt{\frac{n(\psi)}{m-m_p}}\to max.$

It can be seen that the first five of the quality criteria cited are based on a heuristic approach, which is simple and logical. Therefore, we will not dwell on them. The rest is based on a statistical and entropic approach. Below is an analysis of them

Suppose, X – to be probability space.

H_0 hypothesis: $\Re(x)$ and $\psi(x)$ are unrelated sudden criteria, here $\Re(x)$ – is decisive rule.

Thenthe occurrence probability of (t,n) pair is defined with hypergeometric distribution, and it equals the following

$$P(p,n)=\frac{C_{m_p}^{t}\cdot C_{m-m_p}^{n}}{C_m^{t+n}} \qquad (2)$$

here $C_{m_p}^t$, $C_{m-m_p}^n$, C_m^{t+n} – binomial coefficient.

Then the information value of $\psi(x)$ predicate in relation to X_p class can be determined as follows

$$IStat\big(t(\psi),n(\psi)\big)=-\frac{1}{m}\log_2\frac{C_{m_p}^{t(\psi)}C_{m-m_p}^{n(\psi)}}{C_m^{t(\psi)+n(\psi)}}.$$

Definition 7. $\psi(x)$ predicate is called statistic law for X_p class, if the $IStat\big(t(\psi),n(\psi)\big)\geq I_0$ is proper for the sufficient big I_0

I_0 valueis selected in relation to the significance level (2). For instance, if its significance level equals 0.05, then it is taken as $I_0=-\log_2 0.05\approx 4$.

The following is the definition of information value through information theory.

Suppose the two results ω_0 and ω_1 with probability q and $1-q$. Then the suitable information amount is:

$I(\omega_0)=-\log_2 q,$

$I(\omega_1)=-\log_2(1-q).$

Mathematic expectation of information amount,that is entropy:

$h(q)=-q\log_2 q-(1-q)\log_2(1-q).$

If we assume that the occurrence of X_p class objects is ω_0 and that of other class objects is ω_1, then the entropy of X study sample is as follows:

$$H\big(X_p\big)=h\left(\frac{m_p}{m}\right).$$

Suppose the ψ predicate separates $t(\psi)$ from m_p object belonging to class X_p and X_p from $m-m_p$ object not belonging to class X_p. Then the selection entropy $\{x\in X:\psi(x)=1\}$:

$$H\big(X_p|\psi=1\big)=\frac{t(\psi)+n(\psi)}{m}h\left(\frac{t(\psi)}{t(\psi)+n(\psi)}\right)$$

Similarly, $\{x\in X:\psi(x)=0\}$ selection entropy:

$$H\big(X_p|\psi=0\big)=\frac{m-t(\psi)-n(\psi)}{m}h\left(\frac{m_p-t(\psi)}{m-t(\psi)-n(\psi)}\right)$$

Hence, after obtaining the data on ψ, the entropy of X study sample appears as follows:

$H(X_p|\psi) = \frac{t(\psi)+n(\psi)}{m} h\left(\frac{t(\psi)}{t(\psi)+n(\psi)}\right) + \frac{m-t(\psi)-n(\psi)}{m} h\left(\frac{m_p-t(\psi)}{m-t(\psi)-n(\psi)}\right).$

We will have the following result:

$$IGain(t(\psi), n(\psi)) = H(X_p) - H(X_p|\psi) \quad (3)$$

X represents the information gain (3) in the separation of the objects of study sample by the ψ predicate whether it belongs to class X_p or not belongs to this class (3).

Definition 8. The ψ predicate is called the law of the entropy criterion for class X_p if $IGain(t(\psi), n(\psi)) \geq G_0$ is appropriate for any predefinedG_0.

It should be noted that the entropy criterion ($IGain$) is asymptotic equivalent to the statistical criterion ($IStat$), i.e.

$m \to \infty$ да $IStat(t(\psi), n(\psi)) \to IGain(t(\psi), n(\psi))$.

Using the abovementioned, it will be possible to form informative criteria that distinguish not only one but also several classes from others in the X learning selection through the ψ predicate, e.g.

1. $IStat(\psi, X) = -\frac{1}{m}\log_2 \frac{C_{m_1}^{t_1} \cdot C_{m_2}^{t_2} \cdot \ldots \cdot C_{m_k}^{t_k}}{C_m^t}$, here
$m_q = card(X_q), t_q = card\{x \in X_q: \psi(x) = 1\}, t = \sum_{q=1}^{k} t_q, (q = \overline{1,k})$.
2. $IGain(\psi, X) = \sum_{q\in\{1,2,..,k\}} \left(h\left(\frac{m_q}{m}\right) - \frac{t}{m} h\left(\frac{t_q}{t}\right) - \frac{m-t}{m} h\left(\frac{m_q-t_q}{m-t}\right)\right)$
here$h(z) \equiv -z\log_2 z$.
3. Gini criterion: $IGini(\psi, X) = card\{(x, y): \psi(x) = \psi(y)\ and\ x, y \in X_q\}$.
4. D-criterion$I(\psi, X) = card\{(x, y): \psi(x) \neq \psi(y) and\ x \in X_q, y \overline{\in} X_q\}$.

Boosting criterion. The busting algorithm was proposed in 1995 by American scientists Freund and Schapire as a universal method of constructing a convex combination of classifiers.

At the same time, the laws are formed sequentially, and with each new law, the "weights" of the allocated objects are changed, that is, the weights of the correctly allocated objects are reduced, and those of incorrectly allocated ones are increased. The updated vector of weights ω is used in the search for the next law ψ on the maximum criterion of weighted information. As a result, the next law seeks to separate the objects that are "the most difficult" for the previous laws, that is, the "least separated". This, in turn, helps to increase the differences in the laws, to cover the objects relatively evenly, and to increase the possibility of generalizing the convex combinations of the laws.

Suppose that Tregularities for classification are defined, through which the classification algorithm $a_T(x)$ is formed. Let Q_T and $Q_{T+1}(\psi, \alpha)$, respectively, before and after the addition of the next α - ψ law to this algorithm, where α is the weight of the law. In that case, the following would be appropriate:

1) if it is$n_X^\omega(\psi) \neq 0$ ва $\psi_{X_p}^* = arg\max_{\psi} I_{X_p}^\omega(\psi, X)$, $\alpha^* = \frac{1}{2} ln \frac{t_{X_p}^\omega(\psi)}{n_X^\omega(\psi)}$, then$Q_{T+1}(\psi, \alpha) = min$, here$I_{X_p}^\omega(\psi, X) = \sqrt{t_{X_p}^\omega(\psi)} - \sqrt{n_X^\omega(\psi)}$.

2) suppose there is$Q_T \leq m\prod_{i=1}^{T}\left(1 - \frac{I_i^2}{m}\right)$and $\psi_{X_p}^i$ regularities whose information value in each step equals$I_i > I > 0$, then maximumly in $T_0 = \frac{mlnm}{I^2}$step $Q_T$equals zero.

Here $I_i = \sqrt{t_{X_p}^\omega(\psi_{X_p}^i)} - \sqrt{n_X^\omega\left(\psi_{X_p}^i\right)}$.

It should be noted that the arbitrary criterion chosen to assess the quality and / or reliability of a rule built to solve a classification issue may not always provide a logically correct result. We illustrate this in the following model issues.

Suppose that $X = \bigcup_{i=1}^{r} X_i$ is a study sample and ψ predicate, for which the following is appropriate:
$card(X) = m, card(X_p) = m_p$,
$t = card\{x: \psi(x) = 1, x \in X_p\}, n = card\{x: \psi(x) = 1, x \in X \backslash X_p\}$.

Model issues

1- Model problem (disproportion of heuristic criteria). Suppose that $m = 300$ and $m_p = 200$. Using the above informative criteria, we evaluate the quality of the ψ predicate (t, n) in pairs. If pairs (50,0) and (100,50) are given, logically, if the first separated pair is of better quality than the second, but evaluated on the t-n criterion, their quality indicators will be the same. Similarly, $\frac{t}{n+1}$ or $t - 5n$ criteria for pairs (50,9) and (5,0), $\frac{t}{t+n}$ criteria for

(100,0) and (5,0), The criteria $\frac{t}{m_p}-\frac{n}{m-m_p}$for (100.0) and (140.20) give a homogeneous estimate. Even for (100,0) and (140,20) the t-n criterion is for the first pair

The estimation for all pairs on suggested criteria is given in table 1.

Table 1. Inconsistency of simple (heuristic) criteria

Object		Criteria of information value								
t	N	$\frac{t}{n+1}$	$\frac{t}{t+n}$	t-n	t-5n	$\frac{t}{m_p}-\frac{n}{m-m_p}$	m*IGain	m*IGini	m*IStat	$\sqrt{t}-\sqrt{n}$
50	0	50	1,00	50	50	0,25	32,75	26,67	32,68	7,07
100	50	1,96	0,67	50	-150	0,00	0,00	0,00	3,36	2,93
50	9	5	0,85	41	5	0,16	8,66	9,60	11,35	4,07
5	0	5	1,00	5	5	0,03	2,96	2,26	2,95	2,24
100	0	100	1,00	100	100	0,50	75,49	66,67	75,28	10,00
140	20	6,67	0,88	120	40	0,50	50,59	59,52	53,50	7,36

2- Model issue (disproportion of statistical and entropy criteria). Suppose that $m = 700$ and $m_p = 245$. The analysis shows that the IGini criterion in pairs (175,0) and (210,32), and the IGain, IGini, and IStat criteria for pairs (175,25) and (220,80) provide results that are relatively incomprehensible (Table 2.2).

Table 2. Disproportion of statistic and entropic criteria

t	n	$\frac{t}{n+1}$	$\frac{t}{t+n}$	t-n	t-7n	$\frac{t}{m_p}-\frac{n}{m-m_p}$	m*IGain	m*IGini	m*IStat	$\sqrt{t}-\sqrt{n}$
175	0	175,00	1,00	175	175	0,71	356,43	394,33	355,74	13,23
210	32	6,36	0,87	178	-14	0,79	339,11	396,62	341,69	8,83
145	45	3,15	0,76	100	-170	0,49	139,65	178,06	143,03	5,33
100	5	16,67	0,95	95	65	0,40	148,17	179,30	150,38	7,76
175	25	6,73	0,88	150	0	0,66	253,02	308,70	229,64	8,23
220	80	2,72	0,73	140	-340	0,72	267,94	308,58	270,83	5,89

In the cases considered, only the boosting criterion showed that it was stable in relation to the rest, moreover, it was more convenient and simple in terms of understanding and calculation than the statistical and entropy criteria.

3. Conclusion

The article provides the basic concepts, definitions and determinations needed. Ways to determine similarity and proximity measurements between objects have been identified when the characteristics of the objects belong to different types.

Qualitative criteria based on a heuristic, statistical, and entropic approach that assessed the reliability and effectiveness of the decisive rule developed for the classification problem with respect to the $\psi(x)$ predicate function identified in the study sample were cited and analyzed. Heuristic criteria were found to have an advantage over other criteria in terms of logical comprehensibility and ease of implementation. The shortcomings of the criteria cited through the model issues were highlighted.

The criteria for assessing the quality of clustering algorithms developed for DATA SCIENCE issues were divided into three categories, namely, external, internal and relative. Criteria for different categories and the principles of their operation were analyzed, and the results of a comparative analysis of a number of criteria on the level of complexity were presented. As a result of the analysis, it was found that the assessment of the reliability of the clustering method was not theoretically fully resolved.

A number of heuristic criteria used in the formation of informative and optimal information systems describing objects have been studied. Their working principles and some features are described.

References

A. Nishanov, A.Kh., G. Djurayev andM.A. Khasanova.2019. Improved algorithms for calculating evaluations in processing medical data. Compusoft, 8(6): 3158-3165.
B. Nishanov, A.,E. Avazov and B.Akbaraliyev.2019. Partial selection method and algorithm for determining graph-based traffic routes in a real-time environment. International Journal of Innovative Technology and Exploring Engineering, 8(6): 696-698.
C. Kamilov, M., A. Nishanov andR. Beglerbekov. 2019.Modified stages of algorithms for computing estimates in the space of informative features. International Journal of Innovative Technology and Exploring Engineering, 8(6): 714-717.
D. Nishanov, A.Kh., Kh.A. Turakulov andKh.V. Turakhanov.1999. A decision rule for identification of eye pathologies. Biomedical Engineering,33(4): 178-179.
E. Nishanov, A. Kh., Kh.A. Turakulov, and Kh.V.Turakhanov. 1999. A decisive rule in classifying diseases of the visual system. Meditsinskaia tekhnika, 4: 16-18.
F. Nishanov, A.,O. Ruzibaev,and N.Tran.2016. Modification of decision rules 'ball Apolonia' the problem of classification. ICISCT. DOI: 10.1109/ICISCT.2016.7777382.
G. Nishanov, A.Kh., G. Djurayev, and M.A. Khasanova. 2020. Classification and feature selection in medical data preprocessing. Compusoft, 9(6): 3725-3732.
H. Nishanov, A.Kh., B.B. Akbaraliev,G.P. Juraev,M.A. Khasanova,M.Kh. Maksudova,andZ.F. Umarova. 2020. The algorithm for selection of symptom complex of ischemic heart diseases based on flexible search. Journal of Cardiovascular Disease Research, 11(2): 218-223
I. Nishanov A.Kh., B.B. Akbaraliev,B.S. Samandarov,O.K. Akhmedov,and S.K. Tajibaev. 2020. An algorithm for classification, localization and selection of informative features in the space of politypic data. Webology, 17(1): 341-364.
J. Nishanov, A.Kh., O.B. Ruzibaev, J.C. Chedjou, K. Kyamakya, G.P. Kolli Abhiram, D. Perumadura De Silva, M.A. Khasanova. 2020. Algorithm for the selection of informative symptoms in the classification of medical data. World Scientific Proceedings Series on Computer Engineering and Information ScienceDevelopments of Artificial Intelligence Technologies in Computation and Robotics: 647-658.
K. Nishanov, A.Kh., B.B. Akbaraliev, Sh.Kh.Tajibaev. 2020. About one feature selection algorithm method in pattern recognition. Eleventh World Conference on Intelligent Systems for Industrial Automation – WCIS. November, Tashkent.
L. Nishanov, A.Kh., B.B. Akbaraliev G.P.Juraev. 2020. A symptom selection algorithm based on classification errors//International Conference on Information Science and Communications Technologies ICISCT 2020 Applications, Trends and Opportunities, Tashkent, November 2020 (accepted).
M. Nishanov, A.Kh., A.T.Rakhmanov, O.B.Ruzibaev, M.E.Shaazizova. 2020. On one Method for Solving the Multi-class Classification Problem//International Conference on Information Science and Communications Technologies ICISCT 2020 Applications, Trends and Opportunities, Tashkent, November 2020 (accepted).
N. Nishanov A.Kh., Sh.N. Saidrasulov,E.S. Babadjanov,U.E. Mamasaidov,Kh.I.Toliev. 2020.Mathematical Statement of Dynamic Factors Affecting the Development of Electron Government. International Journal of Engineering Research and Technology, 13(12):5240-5246.