Normalization of Dataset for Measuring Compactness

Mukhammadiev Ravshan Abduvali ogli[1] and Mukhamadiev Abdivali Shukurovich[2]*

[1]National University of Uzbekistan, Tashkent, Uzbekistan

[2]Tashkent University of Information Technologies named by Muhammad Al- Khwarizmi, Tashkent, Uzbekistan

*Corresponding Author: Mukhamadiev Abdivali Shukurovich, Doctor of physical and mathematical science, professor Television and media technologies of Tashkent University of Information Technologies named after Muhammad Al-Khwarizmi, Tashkent, Uzbekistan.

Abstract

The article dedicated to finding measure of compactness of normalized dataset, and analyze dataset with using compactness. For finding measure of compactness, we search for hidden regularities from "jaw" database. In this way we need consider n-p full task and it will be 2^n combinations. To solve this task, we use heuristic way with using binary matrix. In the article, we use various methods which find normalization of dataset. Our experiments were taken from "jaw" and we obtained a result from this dataset.

Keywords: object; set; class; feature; nominal; quantitative; space; similarity matrix

Introduction

Normalization is a data preparation technique that is frequently used in data mining. The process of transforming the columns in a dataset to the same scale is referred to as normalization. Every dataset does not need to be normalized for data mining. It is only required when the ranges of characteristics are different. The problem of searching for hidden regularities by methods of data mining is considered. The presence of regularities is estimated according to the compactness of class objects and the sample as a whole. The measure of compactness is used for analysis of the structure of relations between objects according to the defined metrics and the method of data normalization. Rationing leads to a deformation of the feature space and an increase (decrease) in the generalizing ability of recognition algorithms. Generalizing ability applies to the main indicators characterizing the quality of recognition algorithms [1]. This ability is manifested in the ability to determine affiliation objects to classes that the algorithm did not see in the learning process. Hypothesis Verification compactness is the basis of many criteria and methods of the theory of pattern recognition. So, in [1] the compactness profile is described for calculating the generalizing ability of families of algorithms having of infinite capacity in the space VC (Vapnik – Chervonenkis) [2]. For determining affiliation the arbitrariness of an arbitrary admissible object to classes when using such families of algorithms move must store the entire sample in memory. Representative of the family with infinite capacity is the "nearest neighbor" algorithm(BS). For practical purposes, when calculating the generalizing ability, it is enough to use local properties (local restrictions) of object samples [1]. The local constraint in [3] can be considered the proposed N.G. Zagoruyko compactness index, determined by the number of objects-standards of minimum coverage, in which the recognition of class objects is fixed sampling was correct. In addition to the compactness indicator, candidates for inclusion in the set of local constraints are the number of noise objects, the dimension of the feature space, the set of objects around locks (subsets of boundary objects) of classes [4] by a given metrics. Interest is the limiting value of dimension, above which the compactness index [3] increases. The set of features that defines the limit value is considered as informative for the proximity measure used. Dimension above the limit leads to erosion of

---

the similarities between sampling objects.

There is a need to introduce a new measure of compactness with dimensional quantities with values in [0, 1]. The values of these quantities are required to analyze how the actual structure of the training sample is different(by a given measure of proximity) from the ideal for recognition. An ideal structure is considered in which the number of reference objects of the minimum coverage is equal to the number of classes. The measure of compactness can be used to compare metrics and transformations of attribute space in relation to "better" on fixed samples of objects. Analysis of the structure of samples is based on the use of the properties of this relationship. The analysis technique is focused on quantitative indicators calculated by the results of dividing class objects into disjoint groups [4]. The guarantee of the uniqueness of the partition by the number of groups and the composition of their objects is the stability of the algorithm used. The influence of noise objects on the indicators of the generalizing ability of algorithms has been repeatedly considered in scientific publications. An extensive list of works in [5] provides an overview of various 38 methods for detecting and removing noise objects. Most of these methods are focused on use of the BS rule. The recognition quality according to the BS rule substantially depends on the sensitivity of the metric to the dimension of the attribute space. The change in dimension is associated with the selection of informative features, and with the transition to the description of objects in space from latent features. In [4, 6], it was proposed to use two types of rules for hierarchical agglomerative grouping of initial features as a toolkit for transition to latent features. The first type is focused on the sequential combination of two signs into one by non-linear mapping of their values on the numerical axis. Grouping according to the rules of the second type is based on values of the stability criterion of objects by a given metric in a two class recognition problem. For each group of attributes, a generalized assessment of the object is calculated. Methods that implement two types of hierarchical grouping rules can be identified as non-linear and linear. Nonlinear methods are invariant to the scale of feature measurements. In linear methods, the invariance property is absent. The sequence of formation of groups and latent features based on them according to two types of rules determines the order by relation of the degree of informativeness. The informativeness of the feature is calculated as the extremum of the criterion dividing its (feature) values into disjoint intervals in the form of checking the degree of truth of a hypothesis: The sets of features values in the description of objects from different classes with the number of intervals equal to the number of classes do not intersect with each other.

In this paper, non-empty classes (sets) of metrics are considered, the cluster structures of training samples for which coincide (are equivalent) in the number of groups and the composition of the objects included in them. Information on the cluster structure allows sequential selection of objects of standards of minimum coverage, in each of which a local metric is defined. The method for calculating the weights of local metrics is similar to that used in the FRiS STOLP method [3]. To evaluate the generalizing ability of the BS method algorithms, it is proposed to use a criterion, the values of which are calculated depending on the dimension and composition of the set of features, the number of removed noise objects and the number of objects of the minimum coverage standards. Assessment by criterion was used to demonstrate the sustainability of the selection of informative features with cross validation method on random samples. Additionally, in this article we will mention our results.

Methodology. On splitting class objects into disjoint groups

The use of a partially trained sample (PTS) for setting grouping conditions is described in [7]. An example of a condition is the indication of a subset of pairs of sample objects that, when split, should not fall into one group. Belonging to disjoint classes serves a source of additional information for the study of cluster structure using various proximity measures. The main ideas of the method given below are presented in [4]. The aims of splitting class objects into disjoint groups are:

- Calculation of data normalization and convert from raw data;
- calculation and analysis of compactness values of class objects and the sample as a whole;
- search for the minimum coverage of the training sample with reference objects.

The recognition problem in the standard setting is considered. It is believed that the set of objects $E_0 = \{S_1, \ldots S_m\}$ divided by $l\,(l > 2)$ disjoint subsets (classes) $K_1, \ldots K_l$, $E_0 = \bigcup_{i=1}^{l} K_i$. Objects are described using a set of n heterogeneous attributes $(N) = \{S_1, \ldots S_m\}$, ξ value which is measured in interval scales, $(n - \xi)$ - in nominal. The metric $\rho(x, y)$ is given on the set of objects E_0.

Let $u_1^1, u_1^2 (u_2^1, u_2^2)$ be the number of difference values $|x_{\mu j} - x_{dj}|$, $j \in I$ from class $K_t (CK_t)$ in the intervals $[c_0, c_1]$ and $(c_1, c_2]$, $A = (a_0, a_1, a_2)$, $a_0 = 0$, $a_2 = m$, a_1 is the serial number of the element of the ascending sequence $r_{j_1}, \ldots, r_{j_m}$ of difference values $|x_{\mu j} - x_{dj}|$ for objects from E_0, defining the boundary of the interval $c_1 = r_{a_1}$, $m_t = |K_t \cap E_0|$.

$$\left(\frac{\sum_{p=1}^{2} \sum_{i=1}^{2} u_i^p \left(u_i^p - 1 \right)}{m_t \left(m_t - 1 \right) + \left(m - m_t \right) \left(m - m_t - 1 \right)} \right) \left(\frac{\sum_{p=1}^{2} u_1^p \left(m - m_t - u_2^p \right) + u_2^p \left(m_t - u_1^p \right)}{2 m_t \left(m - m_t \right)} \right) \to \max_{\{A\}} \quad (1)$$

Criterion allows you to calculate the optimal value of the interval boundary c_1 and use its (criterion) value as a parameter of a quantitative characteristic in the local metric of the object S_d. If within the boundaries of each of the two intervals the values of the differences, $|x_{\mu j} - x_{dj}|$, $j \in I$ of objects only from K_t or CK_t are placed, then criterion (2.1) takes a value equal to one. For $r_{j_1} = r_{j_2}, \ldots r_{j_{m-1}} = r_{j_m}$, the criterion is equal to 0. In all other cases, the maximum of criterion (2.1) takes a value from the interval (0,1). Independence of the measurement scale is ensured by normalizing the values of features in [0,1] according to the formula

$$x_{ji}^* = \frac{x_{ji} - x_{\min}^i}{x_{\max}^i - x_{\min}^i}$$

where $x_{\min}^i, x_{\max}^i$ is the minimum and maximum value of attribute $x_i, i \in I$, respectively. We define the local metric of object S_d as

$$\rho_d (x, y) = \sum_{j=1}^{n} w_{dj} \left| x_j - y_j \right| \quad (2)$$

where w_{dj} - is the value (1) for the j- attribute, $j \in I$. The weight of feature $x_j, j \in J$ for the local metric of object $S_d \in K_u$ is calculated as $w_{dj} = \frac{\left| \theta(S_d, j) \right|}{\left| CK_u \cap E_0 \right|}$, where $\theta(S_d, j) = \left\{ S_t \mid S_t \in CK_u \cap E_0, x_{tj} \neq x_{dj} \right\}$. Taking into account the weights of nominal features, the local metric takes the form

$$\rho_d (x, y) = \sum_{j \in I} w_{dj} \left| x_j - y_j \right| + \sum_{j \in J} \begin{cases} w_{dj}, x_j \neq y_j, \\ 0, x_j = y_j. \end{cases} \quad (3)$$

We denote by $\rho_d (x, y) = \sum_{j \in I} w_{dj} \left| x_j - y_j \right| + \sum_{j \in J} \begin{cases} w_{dj}, x_j \neq y_j, \\ 0, x_j = y_j. \end{cases} L(E_0, p)$ the subset of boundary objects of

classes defined on E_0 by the metric $\rho(x,y)$. Objects $S_i, S_j \in K_t$, $t = 1, \ldots, l$ are considered related $S_i \leftrightarrow S_j$, if $\{S \in L(E_0, \rho) \mid \rho(S, S_i) < r_i \, and \, \rho(S, S_j) < r_j\} \neq \varnothing$ where $r_i(r_j)$ distance to the nearest object from $S_i(S_j)$ from $CK_t(CK_t = E_0 \setminus K_t)$ by metric $\rho(x,y)$. The set $G_{tv} = \{S_{v1}, \ldots S_{vc}\}, c \geq 2, G_{tv} \subset K_t, v <\mid K_t \mid$ represents an area (group) with related objects in class K_t, if for any $S_{vi}, S_{vj} \in G_{tv}$ there exists a path $S_{vi} \leftrightarrow S_{vj} \leftrightarrow \ldots \leftrightarrow S_{vj}$. An object $S_i \in K_t, t = 1, \ldots l$ belongs to a group of one element and is considered unconnected if it does not exist paths $S_i \leftrightarrow S_j$ for no one object $S_j \neq S_i, S_j \in K_t$. It is required to determine the minimum number of disjoint groups of connected and unrelated objects for each class $K_t, t = 1, \ldots, l$. This problem can be considered in an alternative setting (without specifying features) if a square proximity matrix $\{a_{ij}\}, m \times m$ between m objects and a vector $F = \{f_{1, \ldots} f_m\}, f_i \in \{1, \ldots, l\}$ belonging of the objects to the $K_1 \ldots K_l$. Vector F serves as additional information for specifying grouping conditions. When determining the minimum number of groups of connected and unrelated class objects, $L(E_0, p)$ is used a subset of the boundary objects (shell) of classes by the given metric ρ and a description of objects in the new space from binary signs. To distinguish the class shell for each $S_i \in K_t, t = 1, \ldots, l$ a sequence ordered by $\rho(x,y)$ is constructed

$$S_{i_0}, S_{i_1}, \ldots, S_{i_{m-1}}, S_i = S_{i_0} \quad (4)$$

Let $S_{i\beta} \in CK_t$ be the object closest to S_i from (4), not belonging to the class K_t. Denote by $O(S_i)$ a neighborhood of radius $r_i = \rho(S_i, S_{i\beta})$ centered in S_i that includes all objects for which $\rho(S_i, S_\tau) < r_i, \tau = 1, \ldots, \beta - 1$. In $O(S_i)$ there is always a non-empty subset of objects

$$\Delta_i = \{S_{i_\alpha} \in O(S_i) \mid \rho(S_{i_\beta}, S_{i_\alpha}) = \min_{S_{i_\tau} \in O(S_i)} \rho(S_{i_\beta}, S_{i_\tau})\} \quad (5)$$

By (5), the belonging of objects to the class shell is defined as $L(E_0, \rho) = \bigcup_{i=1}^{m} \Delta_i$.

The set of shell objects from $CK_t \cap L(E_0, \rho)$ is denoted by $L_t(E_0, \rho) = \{S^1, \ldots, S^\pi\}, \pi \geq 1$. The value $\pi = 1$ uniquely determines the inclusion of all objects of the class in one group. For $\pi \geq 1$ we transform the description of each object $S_i \in S_\mu \vee S_\eta$ in $S_i = (y_{i1}, \ldots, y_{i\pi})$ where

$$y_{ij} = \begin{cases} 1, \rho(S_i, S^j) < r_i \\ 0, \rho(S_i, S^j) \geq r_i \end{cases} \quad (6)$$

Let, according to (6), a description of objects of the class K_t in a new (binary) attribute space be obtained, $\Omega = K_t$ and θ - is the number of disjoint groups of objects, $S_\mu \vee S_\eta, S_\mu \wedge S_\eta$ - accordingly, the operations of disjunction and conjunction by binary signs of objects $S_\mu, S_\eta \in K_t$ The step-by-step algorithm for splitting objects K_t into disjoint groups $G_1, \ldots G_\theta$ is as follows.

Step 1: $\theta = 0$

Step 2: Select object $S \in \Omega, \theta = \theta + 1, Z = S, G_\theta = \varnothing$

Step3: Make a choice $S = \Omega$ end $S \wedge Z = true$, $\Omega = \Omega \setminus S, G_\theta = G_\theta \cup S, Z = Z \vee S$ till $\{S \in \Omega \mid S \wedge Z = true\} \neq \varnothing$

Step 4: if $\Omega \neq \varnothing$ then goto step 2

Step 5: the end.

Partitioning of E_0 objects into disjoint groups according to the algorithm described above is used to find the minimum coverage [4] of the training sample by objects by standards. Denote by $R_s = \rho(S, \overline{S})$ the distance from the object $S \in K_t$ to the nearest object $\overline{S}$ from the class $(\overline{S} \in CK_t)$ opposite to K_t, by δ is the minimum number of disjoint groups of connected and unconnected objects of classes on E_0.

We order the objects of each group $G_u \cap K_t, u = 1,...\delta, t = 1,...l$ according to the set of values $\{R_s\}_{s \in G_u}$ As a measure of proximity between $S \in G_u, u = 1,...\delta$, and an arbitrary admissible object S', a weighted distance over the local metric $(S, S') = \rho(S, S') / R_s$. The decision on whether S' belongs to one of the classes $K_1,...K_i$ is made according to the rule: $S' \in K_t$ if

$$d(S_\mu, S') = \min_{S_j \in E_0} d(S_j, S') \, and \, S_\mu \in K_t \, and \, d(S_\mu, S') \neq \min_{S_j \in CK_t} d(S_j, S') \ (7).$$

According to the principle of *sequential exclusion* used in the search for coverage, the sample E_0 is divided into two subsets: the set of standards E_{ed} and the control set $E_k, E_0 = E_{ed} \cup E_k$. At the beginning of the process $E_{ed} = E_0, E_k = \varnothing$. Sorting by values from $\{R_s\}_{s \in G_u}, u = 1,...\delta$ is used to determine the candidate for removal from the number of reference objects by the group G_u. The idea of selection consists to the find the minimum number of standards at which the recognition algorithm by (7) remains correct (without errors recognizing objects) on E_0. We assume that the numbering of groups of objects reflects the order $|G_1| \geq ... \geq |G_\delta|$ and the group $G_p, p = 1,..., \delta$ δ did not select reference objects. E_{ed} deletion candidates are sequentially selected starting from $S \in G_p$ with a minimum R_s value. If the inclusion of S in E_k violates the correctness of the decision rule (4), then S returns to the set E_{ed}.

Realization of the concept. On compactness measures in recognition problems with a teacher

Compactness measures are claimed for evaluating the generalizing ability of recognition algorithms. When calculating the estimates, the results of the search and removal of noise objects, the selection of informative feature sets, the number of reference objects of the minimum coverage of training samples are used. Consider the method of forming a set of noise objects whose power depends on the verification of the conditions proposed below.

Let be $S_k \in K_i, \rho(S_k, S_r) = \min_{S_j \in CK_i} \rho(S_k, S_j)$ and $Z = |\{S_\mu \in K_i \mid \rho(S_k, S_\mu) < \rho(S_k, S_r)\}|$. Denote by $D_i(D_i \in CK_i)$ the set of noise objects of class K_i. The object $S_r \in CK_i$ is included in D_i and is considered as noise if the condition is satisfied:

$$\frac{ZZ - \lambda}{|K_i|} > \frac{1}{m - |K_i|} \ (8)$$

where $ZZ = |\{S_\mu \in K_i \mid \rho(S_r, S_k) < \rho(S_p, S_k) < \rho(S_\eta, S_k)\}|, |\lambda| < \min_{1 \leq i \leq l} |K_i|, \rho(S_\eta, S_k) = \min_{S_j \in CK_{i\{S_r\}}} \rho(S_j, S_k)$. The values Z and Z + ZZ can be considered as the number of representatives of the class K_i in the hypersphere centered in $S_k \in K_i$ respectively, before and after the removal of the noise object S_r. The selection of objects of training samples, under certain restrictions, helps to increase the generalizing ability of recognition algorithms. It is believed that the generalizing ability of the algorithm is improved if it is allowed to make mistakes on the determined objects of the sample. In our case as such objects from $\bigcup_{i=1}^{l} D_i$.

Let the representatives of the class be divided by the minimum number μ of disjoint groups of objects according to the algorithm from above mentioned, $\min m_{ij} =| G_{ij} |, j = 1,...,\mu, \sum_{j=1}^{\mu} m_{ij} = m_i$. To analyze the results of dividing the class K_i into disjoint groups, taking into account their number, representativeness (by the number of objects) and removal of noise objects, it is proposed to use such a structural characteristic as compactness estimation:

$$\Theta_i = \frac{\sum_{j=1}^{\mu} m_{ij}^2}{m_i^2} \quad (9)$$

Obviously, the set of admissible values of Θ_i by (6) lies in the interval $\left[\frac{1}{m_i}, 1\right]$. If the group G_{i1} contains all objects from $K_i \cap (E_0 \setminus \bigcup_{j=1}^{l} D_j)$ then $\Theta = 1$. The average estimate of the compactness of the training sample as a whole is made taking into account the fraction of $\left(\frac{E_0 \setminus \bigcup_{i=1}^{l} D_j}{m}\right)$ noise objects excluded from consideration according to (8) as

$$R(E_0, \rho) = \left(\frac{\sum_{i=1}^{l} D_j}{m}\right) \frac{\sum_{i=1}^{l} m_i \Theta_i}{E_0 \setminus \bigcup_{i=1}^{l} D_j} = \frac{\sum_{i=1}^{l} m_i \Theta_i}{m} \quad (10)$$

Values (9) and (10) indirectly indicate the homogeneity (heterogeneity) of the structure of the training sample. The closer the similarity of the groups in terms of the number of class objects included in them, the closer the value of (9) – to $\frac{1}{m_i}$, and (10)- to $\frac{1}{m}$.

Obviously, the number and composition of noise objects depend both on the value of parameter λ in (5) and on sets of features in the description of objects. The problem of implementing computational procedures is to coordinate the selection of informative features and the removal of noise objects. Let the structure of class objects in the sample E_0 be calculated by the grouping algorithm from above mentioned. We denote by $Sh(\lambda, X(k))$ the number of noise objects E_0 determined depending on the value of λ according to (8) on the set of features $X(k) \subset X(n), CF$ is the number of reference objects of the minimum coverage of the training sample, from which $Sh(\lambda, X(k))$ noise objects are removed. Since it is impossible to obtain an exact solution to the problem of selecting informative features without enumerating all their combinations taking into account the removal of noise objects, in practice it is recommended to use various heuristic methods. Regardless of the methods used, the quality of the selection of informative features is proposed to be determined by checking two conditions:

- When removing noise objects from , the indicator of the minimum coverage of the sample with reference objects

$$F(X(k), \lambda) = \left(\frac{m - Sh(\lambda, X(k))}{m}\right)\left(\frac{m - Sh(\lambda, X(k))}{CF}\right) \quad (11)$$

tends to the maximum allowable value $\frac{m}{l}$;

- Product of the number of reference objects of the minimum coverage by the dimension of the attribute space

$$\frac{k * CF}{m - Sh(\lambda, X(k))} \to \min_{E_0} \quad (12)$$

The first condition (11) is necessary to assess the compactness of the coverage of the sample with standard objects, the second (12) - to assess the complexity of the calculations.

To search for informative sets $\{X(k) \mid X(k) \subset X(n)\}$ two criteria are proposed. Both criteria clearly do not use the number of reference objects of the minimum CF coverage. The number of noise objects $Sh(\lambda, X(k))$ according to (8) is calculated by a fixed value of λ. Such λ for all sets $X(k) \subset X(n), k \geq 2$, is defined as

$$\lambda = \arg\max_{0 \leq |\eta| < \min_{1 \leq i \leq l} |K_i|} F(X(n), \eta) \quad (13)$$

The use of (10) is based on the assumption that the probability of selecting informative feature sets with a higher compactness in (11) is close to zero for λ other than (13). In the first (in the order of presentation) criteria, the results of covering sample objects with hyperspheres are taken into account, taking into account the removal of noise objects, in the second, compactness estimates are used according to (10) based on the property of connectivity over objects of class shells. Let $O(S_i, X(k))(1 \leq k < n)$ be the neighborhood of the object, $S_i \in E_0 \cap K_j, j = 1, ..., l$ defined as $O(S_i, X(k)) = \{S \in K_j \mid \rho(S, \overline{S}_l)\}$, where $\overline{S}_l \in CK_j$ is the object closest to S_i by the metric $\rho(x, y)$ from the complement to the class K_j by the set of attributes $X(k)$. We define the estimate $S_i \in E_0$ on $X(k)$ as

$$Z(S_i, X(k)) = \max_{S_i \in O(S, X(k))} |O(S, X(k))| \quad (14)$$

The attribute $x_d \in X(n)$ is a candidate for inclusion in the set $X(k)$ if

$$\sum_{S_i} Z(S_i, X(k+1)) > \sum_{S_i \in T} Z(S_i, X(k)) \quad (15)$$

where $X(k+1) = X(k) \cup \{x_d\}, T \subset E_0$.

Computational experiment. For a given training set $E_0 = \{S_1, ..., S_m\}$, consisting of representatives l disjoint classes of objects. It is believed that each object $S_i \in E_0$ is described by n nominal features.

We choose a one wolf and dog("jaw") dataset which determines jaws of dog and wolf. The features of dataset are the parts of jaws. The values of feature means the values of jaw parts. In our dataset there are 42 objects and 6 features. The last feature in our dataset is class of object.

First of all, we determined shell objects, which is subset of boundary objects. As a result, 19 objects from 42 objects in dataset were taken as a shell object, after that, we built y_{ij} binary matrix. After that, we grouped objects of classes. Groups of first class:

{0,1,2,4,5,22,}, {3,18,}, {6,8,11,15,19,25,12,13,14,17,9,28,20,21,}, {7,10,26,}, {16,}, {23,27,29,24,}– number of groups=6.

Groups of second class:

{30,}, {31,40,41,}, {32,33,34,38,}, {35,}, {36,}, {37,}, {39,}– number of groups=7.

Finally, we found compactness of both classes: 0,2911111 and 0,2083333 respectively. It can be seen that, compactness of classes are not acceptable in raw dataset. So, we converted metrics to another one with the using of normalization. Then we determined shell objects and 9 objects from 42 objects were shells. Then we built again y_ij binary matrix. After that we grouped objects of classes

{0,2,3,4,5,7,10,11,14,18,20,21,22,25,26,1,6,8,12,13,15,17,19,9,24,29,16,23,27,28,}- 1st class groups.

{30,35,36,38,39,40,41,31,32,33,34,37,}- 2nd class groups.

Finally, measurement of compactnesses of classes were 1. It can be seen that, normalization gave us excellent result.

References

1. Vetrov DP and Ryazanov VV. "On minimizing the feature space in recognition problems". Tr. international conf. Mathematical methods of pattern recognition: M (2001): 22-24.
2. Ignatiev NA. "Extraction of explicit knowledge from different types of data using neural networks". Computing technologies. Novosibirsk 8.2 (2003): 69-73.
3. Ignatiev NA and Madrakhimov ShF. "About some ways to increase the transparency of neural networks". Computing technologies.- Novosibirsk 8.6 (2003): 31-37.
4. Ignatiev NA, Madrakhimov ShF and Yuldashev RU. Artificial neural networks with minimal configuration and forecasting problems in meteorology Uzbek Journal "Problems of Informatics and Energy". 1 (2009): 16-21.
5. Belkin AR and Levin MS. "Decision-making: combinatorial models of information approximation". Moscow: Nauka (1990): 160.
6. Bernatovich AS. "An active experiment in the identification of functional systems for the operational implementation of simulation-type models". Cybernetics 1 (1983): 99-104.
7. Ayvazyan SA., et al. "Applied statistics. Classification and dimension reduction". Moscow: Finance and Statistics (1989): 608.
8. Asuncion A and Newman DJ. "UCI Machine Learning Repository. University of California". Irvine (2007).
9. Aliyev FA, Khujaev JI and Ravshanov ZN. Differential-difference method for solving one-dimensional equations of parabolic type under boundary conditions first and second births Scientific Bulletin of Andijan National university (Andijan) 4 (2017): 5-10.
10. Anderson D, Tannehill G and Pletcher R. Computational fluid mechanics and loop exchange (Moscow: World) Moscow: Nauka) (1990).
11. N Sedova., et al. "Automated Stationary Obstacle Avoidance When Navigating a Marine Craft". 2019 International Multi-Conference on Engineering, Computer and Information Sciences, SIBIRCON 2019; Novosibirsk; Russian Federation (2019).
12. Beknazarova S, Mukhamadiyev ASh and Jaumitbayeva MK. "Processing color images, brightness and color conversion". International Conference on Information Science and Communications Technologies ICISCT 2019 Applications, Trends and Opportunities. Tashkent (2019).
13. Beknazarova S., et al. "The Mask of Objects in Intellectual Irrigation Systems". International Conference on Information Science and Communications Technologies ICISCT 2020 Applications, Trends and Opportunities. Tashkent (2020).
14. Beknazarova S and Mukhamadiyev ASh. "Iterative methods for solving managed problems of the television image processing". Journal of Physics: Conference Series 2032.1 (2021): 012002.
15. Aleksandrovich IN., et al. "Extraction informative features for making decision". International Conference on Information Science and Communications Technologies: Applications, Trends and Opportunities, ICISCT 2021 (2021).