

12-28-2019

Algebraic Correction Of Algorithms For Recognition And Identification Of Biological Objects In Data Mining Problems

Mirzoyan Kamilov

Scientific And Innovation Center Of Information And Communication Technologies, Tashkent, Uzbekistan,
kamilov_m@mail.ru

Mirzaakbar Hudayberdiev

Scientific And Innovation Center Of Information And Communication Technologies, Tashkent, Uzbekistan,
mirzaakbarhh@gmail.com

Follow this and additional works at: <https://uzjournals.edu.uz/ijctcm>

 Part of the [Engineering Commons](#)

Recommended Citation

Kamilov, Mirzoyan and Hudayberdiev, Mirzaakbar (2019) "Algebraic Correction Of Algorithms For Recognition And Identification Of Biological Objects In Data Mining Problems," *Chemical Technology, Control and Management*: Vol. 2019 : Iss. 5 , Article 4.

Available at: <https://uzjournals.edu.uz/ijctcm/vol2019/iss5/4>

This Article is brought to you for free and open access by 2030 Uzbekistan Research Online. It has been accepted for inclusion in Chemical Technology, Control and Management by an authorized editor of 2030 Uzbekistan Research Online. For more information, please contact sh.erkinov@edu.uz.

Algebraic Correction Of Algorithms For Recognition And Identification Of Biological Objects In Data Mining Problems

Cover Page Footnote

Tashkent State Technical University, SSC «UZSTROYMATERIALY», SSC «UZKIMYOSANOAT», JV «SOVPLASTITAL», Agency on Intellectual Property of the Republic of Uzbekistan

Erratum

?????



UDC 004.93

ALGEBRAIC CORRECTION OF ALGORITHMS FOR RECOGNITION AND IDENTIFICATION OF BIOLOGICAL OBJECTS IN DATA MINING PROBLEMS

Kamilov Mirzoyan¹, Hudayberdiev Mirzaakbar²

¹Scientific And Innovation Center Of Information And Communication Technologies, Tashkent, Uzbekistan
E-mail: kamilov_m@mail.ru;

²Scientific And Innovation Center Of Information And Communication Technologies, Tashkent, Uzbekistan
E-mail: mirzaakbarhh@gmail.com

Abstract: The integration of information technologies with biology is considered, enabling specialists to further study the subject area. The object of the study is the identification of the information structure and parameters of Insects Orthoptera of Uzbekistan. For this purpose, an information and identification software complex were created and the analysis of the variety of orthoptera having a certain internal structure and, accordingly, the formed structure of the relational database of the Insecta Orthoptera species was carried out. At the same time, attention will also be paid to the study of a number of scientific issues, one way or another related to the algebraic approach in pattern recognition, such as the choice of optimization procedures for algebraization of algorithms, the formation of a training sample of biological objects, etc. there are also some problems due to the heterogeneity of data on the erect insects of Uzbekistan (Insecta Orthoptera) as biological objects. In such cases, there is a need to improve the existing algorithmic base.

Keywords — identification, algebraic correction algorithms, Uzbekistan Orthoptera insects, morphological analysis of taxon information, database, Data Mining.

Аннотация: Ахборот технологияларини биология билан интеграциясини қарши орқали соҳа мутахассислари предметни янада чуқурроқ ўрганишга имконият яратиши масаласига қаралган. Тадқиқот объекти сифатида Ўзбекистон Insecta Orthoptera ахборот тизими тузилмаси ва параметрларини идентификациялаш олинган. Бунинг учун ахборот-идентификацион дастурий мажмуа ишлаб чиқилган ва orthoptera биохилма-хиллиги таҳлил қилинган ҳамда мавжуд ички тузилмаси аниқланиб, мос ҳолда Insecta Orthoptera тури учун маълумотлар базасининг реляцион тузилмаси шакллантирилган. Шунингдек, иш жараёнида тимсолларни аниқлашда алгебраик ёндашув билан боғлиқ бир қатор илмий масалаларни тадқиқига ҳам эътибор қаратилган, масалан, биообъектлар бўйича ўқув танлангани шакллантириши, алгоритмларни алгебралаштиришида оптималлаштириши амалларини танлаш каби масалалар. Биообъект сифатида қаралаётган Ўзбекистон тўғри қанотли ҳашоратлари (Insecta Orthoptera) ҳақидаги турли шкалланик масалалари ҳам вужудга келади. Бундай ҳолларда мавжуд алгоритмик базани такомиллаштиришининг зарурияти юзага келади.

Таянч сўзлар: идентификация, алгоритмларни алгебраик коррекциялаш, Ўзбекистон ҳашоратлари Orthopterаси, таксонлар ҳақидаги ахборотларни морфологик таҳлил қилиш, маълумотлар базаси, Data Mining.

Аннотация: Рассматривается интеграция информационных технологий с биологией, дающих возможность специалистам дополнительно изучить предметную область. Объектом исследования являются идентификация информационной структуры и параметры Insecta Orthoptera Узбекистана. Для этого создан информационно-идентификационный программный комплекс и проведён анализ разнообразия orthoptera, имеющих определённую внутреннюю структуру и, соответственно, сформированной структуры реляционной базы данных вида Insecta Orthoptera. Вместе с тем при выполнении работы будет уделено внимание также исследованию ряда научных вопросов, так или иначе связанных с алгебраическим подходом в распознавании образов, как, например, выбор оптимизационных процедур при алгебраизации алгоритмов, формирование обучающей выборки биообъектов и т.д. Возникают также некоторые проблемы из-за разношкальности данных о прямокрылых насекомых Узбекистана (Insecta Orthoptera) как биообъектов. В таких случаях возникает необходимость усовершенствовать существующие алгоритмические базы.

Ключевые слова: идентификация, алгебраической коррекции алгоритмов, Orthoptera насекомых Узбекистана, морфологический анализ информации о таксонах, база данных, Data Mining.

INTRODUCTION

Currently, the situation in biology, as well as in many other areas, for example, in medicine, health, etc., is characterized by the availability in the various institutions and relevant institutions of extensive information material associated with the accumulation of data during their long-term operation and operational activities. In many cases, the information in such materials can be presented in the form of tables, rows of which correspond to the descriptions of observations of certain similar objects (bioobjects), and the columns to the characteristic values in terms of which this description is given. The formation of such precedent samples is the most important part of the research work, which makes it possible to use a wide arsenal of mathematical and software analysis tools in order to extract hidden dependencies, evaluate various characteristics and precedents (objects), calculate the so-called "basic properties". In the terminology of this work, the "basic property" is an indication of the belonging of the selected (control) bioobject to a particular genus, species or family, i.e. the result of solving the problem of identification (definition) of the bioobject, and in terms of pattern recognition - the number of the class (taxon) to which the object under study belongs. If the most important hidden characteristic (the "main property") takes a finite number of values, then the problem of creating an algorithm for its calculation from the given values of the characteristics can be solved in the setting of the standard recognition problem by precedents (partial precedents). In the present paper, it is proposed to use the so-called algebraic approach to solving the problem of the synthesis of correct (that is, accurate precedents) recognition algorithms proposed by Yu.I.Zhuravlev, whose essence lies in the fact that certain algorithms are selected in a certain way from the existing recognition models and, using suitable algebraic operations on them (algebraic correction), models are constructed for solving specific problems (for example, identification)[1].

The information and identification system consists of three modules

- a module that accumulates information about bio objects (in a database);
- a module for identifying the state of an bio objects;
- a module that is the identification via a property of bio objects.

The first two modules served for a pattern recognition module. The considered first and second modules are built on relational and recursive principles. From a mathematical point of view, a relational database is a finite set of finite relations of arbitrary arity over predetermined sets of elementary data, in other words, a relational database (more precisely, any of its states) is a finite model (for given relations) in the sense of mathematical logic. Above relations of a model it is possible to carry out various operations, the study of which becomes the field of applications of mathematical programming and modern algebra. All data, according to the relational model, are treated as stored in tables, in which each row has the same format. Each row in the table represents a real-world object or a ratio between objects [1, 2]. The relational approach to representing the data model developed by Codd in the second half of the last century, gained immense fame thanks to the simplicity of the basic ideas and the strict formal theoretical foundation.

Currently, there are several main areas of research in the relational approach. One direction is to expand its scope by modifying the theoretical base of relational theory, which allows the use of new properties and concepts to create a more complete and natural description of the subject area. Traditional Codd algebra allows you to manipulate ordinary flat tables. However, users often have to deal with meaningful recursive concepts for which Codd's algebra does not work, therefore, there is a need to define this algebra. An object whose formal definition contains a link to itself is called recursive. This is an equicomponent of a special kind. Note that a tuple cannot refer to itself either directly or indirectly.

METHODS

Note that, despite the study of objects of biology for a long time, the integration of information technology with biology has allowed us to further explore and explore biological objects. It is known

that bioobjects consist of large quantitative systematic units (for example, species, genus, family, etc.). Identification of biological objects is performed on the basis of these systematic units.

As a result of the use of information technologies in biology, naturally, huge data arrays are being formed, which are accumulated in data storages (Data Center, Bio Data). In turn, when processing information on large databases, such targeted processing of information contained in the data of these databases required (rather stimulated) the development and development of new methods and research principles, resulting in the emergence of such new scientific directions as Bioinformatics, Data Mining, WebMining, Data Sciences, Big Data. Therefore, the development of information systems for the description, identification and classification of biological objects is one of the urgent tasks in various applied fields of biology.

It is known that systematic analysis is carried out with the help of definitive keys, the main task of which is to ensure the distinction between different taxa of certain genera of biological objects based on their morphological characteristics, i.e. the task is somehow connected with the task of identifying bioobjects. It should be emphasized that, although the task of identifying bioobjects was given attention even Zh.B. Lamarck (Lamarck, 1778) in the book "Flora of France", in a special section "The Problem of Identification", still had a serious interest in it in the second half of the last century, as a result of which a new scientific direction "biological diagnostics" was formed, the subject of which is the development principles and methods of using the keys (signs) to identify bioobjects. A great contribution to the development of this scientific direction was made by a number of outstanding scientists and researchers [3-5].

Currently, the intensive development of theoretical and applied informatics, data mining methods, the emergence of an ever-growing range of unique computer systems with broad functionality and relevant information technologies, and, most importantly, the desire of researchers to maximize these opportunities to solve problems leads to significant intensification and a dramatic improvement in the qualities of both the research process itself and its final result. In particular, the use of pattern recognition algorithms (classification and clustering) to solve the problems of systematization of biological objects led to the formation of the concept of "computer identification" of biological objects among specialists in this field.

RELATED WORKS

There are several software systems in the field of "Computer Identification", such as BIKEY 8 (Biological Identification KEYS), PICKEY 8 (Picture Identification KEYS), Linnaeus II, Xper3, BioBASE for Windows 7.0, Bioimages, Gene data data, Metabolic pathways, Gene network inference, Protein-protein interaction data, Metabolic network analysis, Tulipa Recognition 1.0, etc. These software systems for computer identification are designed for different operating systems or platforms [2,3].

The analysis shows that most of the mentioned software systems with varying degrees of success use various models and algorithms for pattern recognition, in particular, classification and clustering algorithms, which are currently the most effective means of data mining (DM). True, some of the algorithms used in the noted complexes existed in scientific everyday life even before the design of the DM as an independent scientific field. Apparently, the choice of one or another algorithm in the complexes was most likely associated with the characteristic features of the tasks to be solved, identifiable objects or, perhaps, with the scientific preferences of researchers. In this paper, biological objects (formation, structure identification and generation of information about objects) are defined as the object base for the information-identification system. Moreover, for the formation of the final (acting when identifying bioobjects) algorithm, a special operation will be used - an algebraic correction over some selected initial subset of the primary algorithms of the models.

The above considerations determine the relevance and necessity of theoretical and applied research and future developments in the framework of the so-called algebraic approach to the formation of algorithms. At the same time, in carrying out the work, attention will also be paid to a

number of scientific issues one way or another connected with the algebraic approach in pattern recognition, such as, for example, the choice of optimization procedures for algebraization of algorithms, the formation of a training sample of biological objects, etc. There are also some problems due to the disparate data on the orthoptera of Uzbekistan (InsectaOrthoptera) as bioobjects. In such cases, there is a need to improve the existing algorithmic framework.

PUPOSE OF WORK

The purpose of this work are consists of two tasks. The first one to develop a relational structure of a database of orthoptera diversity, having a specific internal structure (organized hierarchically) and identification based on the morphological features of biological objects orthoptera insects of Uzbekistan. And second of the design of optimization procedures for algebraization of pattern recognition algorithms.

IDENTIFICATION OF THE DATABASES

There is a need for two different approaches to creating a relational database structure for a selected area of biological objects:

- determining the type of database for the field;
- defining database table fields using orthopteroidea super ordo and implementing a relational relation.

As for the first approach, the analysis of literary sources has shown that a new electronic resource in the field of zoology has appeared, consisting of several databases [3], namely:

- electronic library form database;
- entomology database;
- geographic database;
- systematic database;
- cytogenetic database;
- bibliographic database.

The second approach aims to create an entomological database and systematic database of the Insecta Orthoptera family of crickets, according to the purpose of this article.

FORMATION OF RELATIONAL STRUCTURES OF THE INSECTA DATABASE AND GENERATION

For the experiments we use the insect dataset of Entomological collection Institute of the Zoology of the Uzbek Academy of sciences. Specialists on the subject proposed genus of the Supersea Insecta Orthopteroidea, because it is good investigated subject area by Uzbek specialists. Clarification of the hierarchy of the Supersea Insecta Orthopteroidea [11]. The next task is the optimal placement of the Insecta Orthoptera super-endo-entomological database table and the establishment of recursive-relational relationships between the tables. A recursive talitsa - is a fixed set of tuples that meet the following conditions:

- a tuple type, where all attributes - the a_i names - are pairwise distinct ($a_i \in A_i, i = \overline{1, n}$), is called the main tuple;
- table T, compiled by a finite set of pairwise different basic tuples — values, is called the main table;
- A generalized table of the main table T is a finite set of recursive tuples - the values of the main table.

Analysis of real information systems shows that often table domains have a large variety of values. For example, the attributes: "Kingdom", "Kingdom", "Type", "SuperClass", "Class", "SubClass", "InfraClass", "Nadotryad", "Order", "Suborder", "Family", etc. [1]. Therefore, we can propose the following idea: each tuple of a Cartesian product of sets with degree n is assigned a number, and instead of a tuple this number is stored in the database. To do this, set the display:

$$F: A_1 \times A_2 \times \dots \times A_n \rightarrow N_n,$$

where $A_1 \times A_2 \times \dots \times A_n$ is the Cartesian product of sets; N_n is the set of numbers $\overline{0, n}$.

If F is bijective, then you can set the inverse mapping:

$$F^{-1}: N_n \rightarrow A_1 \times A_2 \times \dots \times A_n.$$

Thus, the bijective mapping F specifies the algorithm for identifying the tuple of a Cartesian product (1), and the mapping F^{-1} - the algorithm for generating the value of the tuple by number (2):

$$\begin{cases} num = Input(D, a), & (1) \\ a = Output(D, num), & (2) \end{cases}$$

where $a \in A_1 \times A_2 \times \dots \times A_n$, $num \in N_n$, D is the description of the sets of the Cartesian product $A_1 \times A_2 \times \dots \times A_n$.

Then the relation $R \subset A_1 \times A_2 \times \dots \times A_n, num \in N_n$ can be uniquely represented by a subset of integers $NUM \subset N_n$. The presence of optimization problems (1) and (2), for the solution of which the same numerical method is suitable, naturally leads to the combined formulation of a single optimization problem. The execution of the first or second task can be parameterized with the numeric parameter $\lambda \in [0, 1]$:

$$Q_\lambda(A_i) = (1 - \lambda)Q(A_i) + \lambda Q(F(A_1 \times A_2 \times \dots \times A_n)). \quad (3)$$

Consider a method for constructing descriptions of sets of domain values D , the Input identification algorithm, and Output generation. As such a procedure, we suggest using decision trees [1]. If the tree describes a set, then the variant describes one element of the set. Then the identification of nodes and the generation of variants of the tree is expressed by the following formula:

$$q(v) = \begin{cases} \bigcap_{i=1}^n A_i = \emptyset, \\ \text{if } \lambda = 0, \text{ to } \bigvee_{i=1}^n (\bigcap_{v=1}^m q(s_i^v)), \\ \text{if } \lambda = 1, \text{ to } \bigwedge_{i=1}^n (\bigcap_{v=1}^m q(s_i^v)), \\ \lambda \in [0, 1], \end{cases} \quad (4)$$

Where v is the considered tree node; s_i^v is the set of sons of the node v ; n is the number of sons.

The algorithm for generating variants works on the basis of the stack principle.

THE SELECTED MODEL TO SOLVE THE PROBLEM

Choosing recognition model and constructing structure of APP

We would like to suggest applying the algorithms of partial precedents (APP) to problems of systematically analysis of tulips via morphological features. The basic model of APP presented below is defined by specifying the stages [1,2,4]. Recognition task: Given $\mathbb{M} = \bigcup_{j=1, \overline{\ell}} \mathcal{K}_j$ - set of disjoint classes of objects. Given the initial information I_0 (training) and a description of some object. The task of recognition consists in determining the inclusion of a given object S in classes \mathcal{K}_j . In the case of APP based on deadlock tests, the initial information I_0 is given by the table:

$I_0 = \|a_{ij}\|_{m \times n}$ - table of attributes of objects in the training sample;

$I(S_i) = (a_{i1}, \dots, a_{in})$ - description of the object from the training sample;

$S_{m_{j-1}+1}, S_{j+2}, \dots, S_j \in \mathcal{K}_j, j = \overline{1, \ell}, m_0 = 0, \dots, m_\ell = m$ - an expression that defines the inclusion of objects in classes;

Algorithm of recognition $A(I_0, S) = \alpha(S)$, где $\alpha(S) = \alpha_1(I_0, S), \dots, \alpha_\ell(I_0, S)$.

$$\alpha_i(S) = \begin{cases} 1, & S \in \mathcal{K}_i, \\ 0, & S \notin \mathcal{K}_i, \\ \Delta, & \alpha \text{ doesn't accept } S. \end{cases}$$

The structure of an APP based on deadlock tests

$\Omega = \{\omega \mid \omega \subseteq \{1, \dots, n\}\}$ - support set system;

A proximity function is introduced for two objects along the reference set $\omega = \{j_1, \dots, j_r\}$:

$B_\omega(S, S') = \bigwedge_{v \in \omega} [\rho_v(S, S') \leq \varepsilon_v]$, where ε_v non-negative numbers, called thresholds, $v =$

$\overline{1, n}$;

An estimate is made of the closeness of the object to the class Γ_i ;

The APP according to the rule: $1 > \delta_1 \geq 1/\ell, \delta_2 \geq 0$ - thresholds of caution.

The algebraically approach to the problem

Every pattern recognition algorithm consists of a recognition operator \mathcal{B}_ω and a superposition of the decisive rules \mathbb{C} [1,2]: $\mathbb{A} = \mathcal{B}_\omega \otimes \mathbb{C}$. In the algebraic approach, the focus is on improving the activities of the recognition operator. At the same time the recognition operator \mathcal{B}_ω will consists of several recognition operators,

$$\mathcal{B}_\omega = (\mathcal{B}_\omega^1, \dots, \mathcal{B}_\omega^r).$$

As mentioned above, all future work is done on the operator \mathcal{B}_ω and below are the mathematical tools needed to handle it. The \mathcal{B}_ω recognition operator performs the basic process of converting the initial information in recognition problems to a standard size matrix.

$$\mathcal{B}(I, K^m, s^q) = \|a_{ji}\|_{q \times m}. \quad (5)$$

If there is a possibility of disable a decision rule from the models algorithm of APP and we could to select recognition operators. The standard-sized matrices can be considered as reflections and can be defined by adding, multiplying, and scaling multiplication. Let us assume that , \mathcal{B}_1 and \mathcal{B}_2 are the operators:

$$\mathcal{B}_1(I, K^m, s^q) = \|a_{ji}^1\|_{q \times m}, \quad \mathcal{B}_2(I, K^m, s^q) = \|a_{ji}^2\|_{q \times m}, \quad (6)$$

The relationship between \mathcal{B}_1 and \mathcal{B}_2 -recognition operators is defined as $\mathcal{B}_\leq = \mathcal{B}_1 \leq \mathcal{B}_2$ and expressed as:

$$\mathcal{B}_\leq(I, K^m, s^q) = \|a_{ji}\|_{q \times m} \leq \|a_{ji}^2\|_{q \times m}. \quad (7)$$

(7) is principally called the linear interconnection of $\{\mathcal{B}_\leq\}$ recognition operators and denoted by $L\{\mathcal{B}_\leq\}$:

$$L\{\mathcal{B}_\leq\} = L\{\mathcal{B}_\leq\} * \{\mathbb{C}\}. \quad (8)$$

The formula for algorithm (8) is called linear closure of the family $\{\mathbb{A}\}$. We denote the linear coupling of the recognition operators by the \mathcal{B}_\leq^g operators, which are $\{\mathcal{B}_\leq\}$,

$$a_1 \mathcal{B}_\leq^1 \cup a_2 \mathcal{B}_\leq^2 \cup \dots \cup a_g \mathcal{B}_\leq^g \cup \dots \cup a_k \mathcal{B}_\leq^k, \quad (9)$$

where $a_g \in R$ is a real number, $g = \overline{1, k}$.

In (7) - (8), the $\{\mathcal{B}_\leq\}$ recognition is a linear closure of the family of operators in the APP, which is called the algebraic gradient of operation. The closure of the algorithm family $\{\mathbb{A}\}$ is expressed as follows,

$$\tilde{L}\{\mathbb{A}\} = L(L\{\mathcal{B}_\leq\} * \{\mathbb{C}\}). \quad (10)$$

(10) The bond is known as algebraic idempotent closure.

EXPERIMENTAL EVALUATION

It consists of 6 types of the Supersea Insecta Orthopteroidea, 72 objects in data base and 20 features. We are called types of Insecta Orthopteroidea with taxon's or precedents. In this experiment an algorithm was trained to measure k-fixing votes from $n = 20$ characters (table 1). In addition, 72 of the 72 objects are randomly assigned as control objects, and the rest are left as reference objects. This choice was made 10 times per k, and APP training was carried out. The result is shown in table 1.

TABLE 1.

TRAINING RESULTS DATABASE OF "ORTHOPTERA COLLECTIONS"

| k | result of training (%) | Time (by second) |
|----------|------------------------|------------------|
| <i>I</i> | <i>2</i> | <i>3</i> |
| 1 | 74,13 | 0,193 |
| 2 | 74,14 | 0,662 |
| 3 | 77,46 | 0,626 |
| 4 | 82,44 | 2,356 |

| | | |
|----|-------|---------|
| 5 | 86,63 | 7,485 |
| 6 | 100 | 18,858 |
| 7 | 100 | 39,391 |
| 8 | 100 | 65,355 |
| 9 | 83,29 | 90,087 |
| 10 | 89,13 | 101,757 |
| 11 | 79,14 | 94,568 |
| 12 | 89,14 | 78,175 |
| 13 | 89,97 | 47,349 |
| 14 | 87,47 | 32,392 |
| 15 | 88,29 | 10,367 |
| 16 | 80,79 | 3,501 |
| 17 | 88,29 | 0,915 |
| 18 | 94,15 | 0,175 |
| 19 | 86,63 | 0,027 |
| 20 | 89,12 | 0,023 |

CONCLUSIONS

The result of this work is an analysis of the state of the theory of information and identification of biological objects and the formation of a relational database, including tables with a recursive relationship, which store data on species and interspecific taxonomy of biological objects for identification of orthopteran insects in Uzbekistan. On the basis of the proposed theory a database of "Orthoptera Info" and "Orthoptera collections" were developed. An analysis of the results shows that arbitrary recursive relations are transitively closed into finite convex sets [3].

We selected informative features and etalon objects for training set. The procedure for selecting reference objects for metric classification algorithms has been developed, which makes it possible to improve the quality of classification, reduce the time of the classification and the amount of stored data. The investigation realized by means of program-recognition complex "Orthoptera Recognition", which is based on APP.

ACKNOWLEDGEMENT

This work was supported by Grant BV-ATEX-2018-(22) of the Ministry of Innovation Development of the Republic of Uzbekistan.

REFERENCES:

1. Zhuravlev Yu.I., Ryazanov V.V., Senko O.V. «Recognition». Mathematical methods. Software system. Practical applications. - Moscow: Publishing «FAZIS», 2006. - P.176. (in Russian).
2. Kamilov M., Hudayberdiev M., Khamroev A. Algorithm for the Development of a Training Set that Best Describes the Objects of Recognition // Procedia Computer Science 150 (2019). - Pp. 116-122. <https://doi.org/10.1016/j.procs.2019.02.024>.
3. Hudayberdiev M. Formation of relative structures of information-identification models for insecta orthoptera bio object. Descendants of Mohammed Al-Khwarizmi. Scientific-practical and information-analytical journal. Volume 4(10), - 2019, -Pp. 31-36.
4. Khamroev Alisher. An algorithm for constructing feature relations between the classes in the training set. Procedia Computer Science. -Volume 103. -2017. -Pp. 244-247. <http://dx.doi.org/10.1016/j.procs.2017.01.094>.
5. Hudayberdiev M. Kh., Akhatov A.R., Hamroev A.Sh. On a Model of Forming the Optimal Parameters of the Recognition Algorithms. International journal of KIMICS, vol.9, no.5. – Korea, Oktober 2011. – Pp. 607-609.