
CLASSIFICATION PROBLEMS IN THE PRELIMINARY DATA PROCESSING

Djuraev Gulomjon Primovich

Associate Professor Of The Department “Methodology Of Exact And Natural Sciences” Of The National Center For Education Teachers Of Kashkadarya Region To New Methods, Candidate Of Pedagogical Sciences, Uzbekistan

ABSTRACT: The main purpose of symbol recognition is to create a rule, and according to this rule, it is determined which class the objects belong to. In this sense, a class is understood as a certain set (subset) of objects with similar properties. Classification methods are widely used in practical systems of various issues, namely, medical and technical diagnostics, geology and mineral exploration, aerospace methods of research, robotics, etc.[1-3].

KEYWORDS: Practical approaches, diagnosis due to speed, accuracy and reliability.

INTRODUCTION

The usage of classification methods in medical diagnosis and the increase in the number of signs analyzed in the creation of computer systems based on them allows to change the efficiency of diagnosis due to speed, accuracy and reliability.

Each practical system of diagnosis requires the creation of a mathematical model designed for the classification of objects and phenomena in a specific field of science.

We will consider the main issues that arise in the process of designing classification systems based on the theoretical and practical approaches proposed by the authors[4-5] in solving problems in the field of identifying medical symbols.

1) The first problem is to form a complete list of symbols describing objects or events. These symbols can be divided into deterministic, probabilistic, logical and structural signs[6-7].

Deterministic signs are symbols that take specific numerical values, and when considering them, measurement errors are not taken into account.

Probabilistic signs are symbols whose random values are distributed over all classes of objects. At the same time, the decision about whether the recognized object belongs to one or another class can be reached only on the basis of the exact value of the signs of this object obtained as a result of conducting relevant experiments. It is necessary to consider probabilistically even when the numerical values of the signs are measured, with errors that make it impossible to say exactly what numerical value this value received according to the results of measurements.

Logical signs are simple considerations that accept two values (existing (true), not available (false)) according to their existing properties. First of all, it is worth saying that all signs that do not have a quantitative value are logical signs. For example, if the patient has an increase

in body temperature or pain in the heart area, etc., it can be considered as logical signs. Also, not the value itself, but only the fact that it matches or does not match the specified interval is important, and can be included in the logical signs.

Structural (linguistic) signs are understood to be expressed in the form of words, signs or images according to the properties of objects. It can be seen from this that if the objects in the classes are expressed in the form of words and signs, then when classifying the control object under consideration, the words and symbols describing the control object are compared with the objects consisting of each word in the classes. As a result, it is determined to which class the considered control object belongs.

2) The second problem is to form a predetermined set of signs. Taking into account the results of the solution of the first problem above, only signs are added to this set, with the help of these signs, the previously obtained information necessary for describing the classes can be entered.

3) The third problem is to describe all class objects in the language of signs.

For example, suppose we are given a collection of $X = \{x\}$ objects. Let each $x \in X$ object be represented by a N sign.

Hence, $\forall x \in X$ uchun $\exists(x^1, x^2, \dots, x^N)$ there exists a character set for such that appropriate $x \leftrightarrow (x^1, x^2, \dots, x^N)$.

Here, the sign x^j can take different values:

$x^j \in \{0,1\}$; $x^j \in \{0,1,-\}$; $x^j \in \{1,2, \dots, K\}$; $x^j \in (a, b) \subset R$ or

$x^j \in [a, b] \subset R$; $x^j \in \{\mu\}$, $j = \overline{1, N}$; - a set of probability measures, etc.

Let's define the studied j –object by a set of sign values D_j ($j = \overline{1, N}$). Then the space of characters describing objects $D = D_1 \times D_2 \times \dots \times D_N$

constitutes, here $\dim(D) = N$.

Definition 1. a set is called D_j ($j = \overline{1, N}$) an alphabet of characters.

Definition 2. $x \in X$ is called an admissible object if $x^j \in D_j$ ($j = \overline{1, N}$).

Let's assume that the objects of the educational selection are given in the following form:

$$X_p = \{x_{p1}, x_{p2}, \dots, x_{pm_p}\}, x_{pi} = (x_{pi}^1, x_{pi}^2, \dots, x_{pi}^N) \subset D^N, i = \overline{1, m_p};$$

where denotes the class $p = \overline{1, r}$; the number of objects in its $m_p - X_p$ class, the dimensional $D^N - N$ sign space. To work on the character space, we introduce a vector $\lambda = (\lambda^1, \lambda^2, \dots, \lambda^N)$ describing the following objects $\lambda = (\lambda^1, \lambda^2, \dots, \lambda^N)$, where $x_{pi} = (x_{pi}^1, x_{pi}^2, \dots, x_{pi}^N)$ means that the components of the vector equal to one participate in the signs corresponding to the components of the object, and the components equal to zero do not participate, i.e.

$$x_{pi}|_\lambda = (\lambda^1 x_{pi}^1, \lambda^2 x_{pi}^2, \dots, \lambda^N x_{pi}^N).$$

So, $X_p|_\lambda = \{x \in X_p|_\lambda, x = (\lambda^1 x^1, \lambda^2 x^2, \dots, \lambda^N x^N)\}$ we can express the classes in the same way.

Definition 3. If

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

space. In addition to the geometric interpretation of the task, there is also an algebraic interpretation.

It is required to create a distribution function $F_p(x_i^1, x_i^2, \dots, x_i^N)$, $p = \overline{1, r}$; for functions with the following properties: if $(x_i^1, x_i^2, \dots, x_i^N)$ the object i – with the signs belongs to the class X_p , the value of $F_p(x_i^1, x_i^2, \dots, x_i^N)$ must be the largest. If $x_q X_k$ denotes a vector of signs of objects belonging to the class, then $F_k(x_q) > F_l(x_q)$, $k, l = \overline{1, r}$; $k \neq l$.

In this way, it is represented by the boundary equation $F_k(x) - F_l(x) = 0$, which is called the determinant between the areas corresponding to the classes of D_p , in the space of X_p signs.

5) The fifth issue is the choice of algorithms for constructing a decisive classification rule (recognition) that ensures that the classified object or phenomenon belongs to a particular class[11-12]. Classification algorithms are based on comparing criteria of proximity or similarity of the classified object with each class. At the same time, if X_p , $p = \overline{1, r}$; a specific object is; if the selected criterion of proximity to a particular class is greater than the criterion of proximity to another class, a decision is made on whether this object belongs to the X_p class.

REFERENCES

1. Vapnik V.N., Chervonenkis A.Ya. Theory of pattern recognition.-M.: Nauka, 1974.-416 p.
2. Verhagen K., Dein R., Grud F., Josten J., Verbeek P. Pattern recognition. State and prospects.-M.: "Radio and communications", 1985.-103 p.
3. J. Tu, Gonzalez R. Principles of pattern recognition//Translated from English. Moscow: Mir Publishing House. 1978.-c.410.
4. Zhuravlev Yu.I., Ryazanov V.V., Senko O.V. Recognition. Mathematical methods. The software system. Practical applications. M.: Fazis, 2005.-113 p.
5. Stepanova M.D., Samodumkin S.A., Lemesheva T.L. Mathematical diagnostic methods in medical intelligent systems//Educational and methodical manual.-Minsk, 2001.-44 p.
6. Gorelik A.L., Skripkin V.A. Recognition methods//Moscow: Higher School, 1977.-222 p.
7. Bekmuratov Q.A., Mamaraufov O.A, Bekmuratov D.Q. Symbol recognition systems//Methodological manual. Samarkand: TATU Samarkand branch, 2016. 380 p.
8. Журавлев Ю.И., Камилов М.М., Туляганов Ш.Е. Алгоритмы вычисления оценок и применение.-Ташкент: ФАН, 1974.-119 с.
9. Nishanov A.Kh., Djurayev, G.P., Khasanova, M.A. Classification and feature selection in medical data preprocessing//Compusoft: An International Journal of Advanced Computer Technology, 9(6), June-2020.-pp. 3725-3732.
10. Nishanov A.Kh., Juraev G.P., Maksudova M.Kh., Khasanova M.Akh. Application and analysis of preliminary processing of medical images for solution of practical problems//Journal of Cardiovascular Disease Research, ISSN: 0975-3583, 0976-2833 Vol. 12, ISSUE 03, 2021.-pp. 36-50.
11. Nishanov A.Kh., Djurayev G.P., Kasanova M.Kh. Improved algorithms for calculating evaluations in processing medical data//Compusoft: An International Journal of Advanced Computer Technology, 8(6), June-2019.-pp. 3158-3165.

12. Nishanov A.Kh., Akbaraliev B.B., Juraev G.P., Khasanova M.A., Maksudova M.Kh., Umarova Z.F. The algorithm for selection of symptom complex of ischemic heart diseases based on flexible search//Journal of Cardiovascular Disease Research, Vol. 11(2), 2020.-pp. 218-223.