



METHODS FOR ANALYZING MEDICAL AND PHARMACEUTICAL DATA PROVIDED IN NON-QUANTITATIVE FORM

Tursunov A.T.¹

Narzullayev D.Z.²

Rahmonov E.D.³

Tashkent Pharmaceutical Institute, Tashkent city, Republic of Uzbekistan

e-mail: tursunovr484z @ gmail.com.ru

<https://doi.org/10.5281/zenodo.17311693>

Relevance: The heuristic approach is based on the researcher's knowledge and intuition, which are difficult to formalize. In this approach, the researcher determines what information and how to use it to achieve the required level of recognition efficiency. Therefore, when analyzing medical and pharmaceutical data presented in quantitative form, it is relevant to convert them into a single type, that is, into a digital form.

Research Objective. The two fundamental methods of knowledge representation described above allow us to propose the following classification of pattern recognition methods. This requires new approaches to analyzing large volumes of medical and pharmaceutical data. To achieve the goal, we consider the following methods.

Method and Methods. Intensional methods of symbol recognition - methods based on performing operations with symbols. Extensional methods of symbol recognition - methods based on performing operations with objects.

At the same time, objects in these methods are not considered as whole information units, but rather as indicators for evaluating the interaction of their attributes and behavior. The group of intensional methods of pattern recognition is very wide, and its division into subclasses is, in a sense, conditional. These methods are used to determine the ratio of similarities in different areas of a multidimensional feature space. This group also includes a method for calculating the likelihood ratio for unrelated traits. This method does not require knowledge of the functional form of the distribution law, except for the assumption that the traits are unrelated (which is almost never true in reality). Therefore, it can be classified as nonparametric. Other nonparametric methods are of particular interest when the shape of the distribution density curve is unknown and no assumptions are made about its nature. These include the method of multidimensional histograms, the method of "k-nearest neighbors", the method of Euclidean distance, the method of potential functions, etc., a generalization of which is the method of "Parzen estimates". These methods formally work with objects as whole structures, but depending on the type of recognition problem, they can be either intensional or extensional forms.

Results. Nonparametric methods analyze the relative number of objects in a given multidimensional volume and use different distance functions between the objects in the training sample and the objects being recognized. For quantitative features, when their number is much smaller than the sample size, operations with objects play an intermediate role in estimating the local density of conditional probabilities.

Conclusions. Thus, the same technological operations of nonparametric methods, depending on the conditions of the problem, have the meaning of locally estimating the density of the probability distribution of character values or assessing the similarity and difference of objects. When classifying an unknown object, a given number (k) of other objects (nearest neighbors) known to belong to the classes being recognized are found in the feature space.