INFORMATION TECHNOLOGY OF THE MULTIDIMENSIONAL BIOMEDIACL OBJECTS DIAGNOSTICS

L. Lomakina, Y. Gubernatorov

Nizhniy Novgorod State Technical University n.a. R.E. Alekseev, Nizhniy Novgorod, Russia, email: yaroslav.gubernatorov@gmail.com

Abstract

This article considers the problem of multidimensional biomedical objects diagnostics. We prove the relevance of the information technology using to solve this problem. We consider the basic model of the multidimensional object. We have human gastrointestinal tract research results. The multidimensional object is represented as a set of results. We need to diagnose the state of the multidimensional object. In our research we use the neural networks as diagnostics method. The neural network is used to diagnose the state of the multidimensional object. The neural network is based on the three-layer perceptron architecture. We use backward propagation of errors method for training our networks. We checked the operation of the neural network on real data. The results suggest that our method can classify the multidimensional biomedical object states. Also, we carry out the different architecture networks analysis.

Keywords: multidimensional object, neural networks, classification, dysbiosis

The classification and clustering of the multidimensional data is the actual problem. In medicine and biology, automated classification and clustering systems are used for the study of medical and biological data and diagnosis. Multidimensional objects are characterized by complex internal structure with nonlinear relationships between objects. Proper interpretation of biomedical research is of exceptional importance in deciding on the nature of the disease and related therapies. Existing methods of treatment of biomedical research and the diagnosis may be contradictory because of the strong influence of subjective factors. Nizhniy Novgorod Research Institute of Epidemiology and Microbiology named after I.N. Blokhina (NNIIEM) for several years, leading research activities to improve the effectiveness of objectification assessments of gastrointestinal microbiota of man. [1] To solve this problem, the diagnosis methods have been developed. This methods base on a comparison of the results of bacteriological studies with a reference group of healthy individuals selected by expert assessments and explored the possibility of using neural network technology and statistical methods. This article discusses the use of neural network

technology for the diagnosis of multi-dimensional medical and biological objects. As a set of multi-dimensional objects the results of medical and biological analysis are used.

The base model, which describes the state of the gastrointestinal tract, based on source data that are provided NNIIEM. The initial data are the results of bacteriological studies GI status separately for the sick and healthy people. The result of analysis of the individual patient represented as a vector $\varphi = \{\varphi_1, \varphi_2, ..., \varphi_n\}$ in n-dimensional Euclidean space whose coordinates are scalars, each of which is equal to the number of microorganisms of the species. [2] Thus, the model is n-dimensional feature spaces that are a priori divided into two classes, healthy and sick patients. Sick patients are considered with qualitative and quantitative composition of microflora in the digestive tract, are subject to disease dysbiosis. There are 3 degrees of dysbiosis. In this paper, we consider n = 29 signs that characterize the state of the microflora of the gastrointestinal tract by microorganisms 376 species from 70 genera. The given symptoms can be divided into four groups. The first group includes those microorganisms whose presence has definitely positive impact on microflora. Moreover, the effect increases with increasing number and variety of microorganisms. The second group includes those microorganisms whose presence in the microbiota has a limited negative impact. This means that their presence in amounts of less than 10^5, poses no threat to human health and even operation improves the first group of microorganisms. The presence of large quantities may complicate the course of the underlying disease. The third group includes microorganisms that have definitely a negative impact on the microflora, which increases with increasing number and diversity of this group. The fourth and last group consists of only two species of bacteria, but their impact on the microflora of the gastrointestinal tract of man is such that upon detection of at least one microorganism can be no further analysis to diagnose gastrointestinal disease of man, as they are the causative agents of diseases such as dysentery and salmonellosis. The study characterizing features, it has been observed that the species and quantity of the children and adult normal microflora has significant differences. 8 age groups have been allocated in this paper surveyed. This division is allowed to increase accuracy in separating people with normal microflora by those who have any deviation from the norm. In addition, it should be noted that all children aged up to 23 hours are healthy and they can be distinguished only norm and dysbiosis of first degree.

The task of diagnosing multi-dimensional objects is a classic classification task. Selection of neural networks as classifier due to the following features:

- Neural networks allow you to model complex patterns. Using the ability of learning on the set of examples, the neural network is able to solve problems in which the unknown patterns and the relationship between the input and output data.

- Neural networks are resistant to the input noise. Provides the ability to work with a large number of non-informative, the noise of the input signals. No need to do their initial screening, the neural network will detect them unsuitable for solving the problem and throw them.
- Easy to implement. The basic operation used in the simulation is a weighted sum.
- Allows the use of parallel processing computers in teaching and classification of objects.

To construct a neural network classifier diagnosis, you need to select the architecture of the neural network, the number of layers, the number of neurons in each layer, as well as the teaching method. We use multi-layer perceptron neural network architecture. Multilayer perceptrons have been successfully used to solve variety complex tasks. In this paper we propose to use a three-layer perceptron. The network has one input layer comprising 29 neurons at the number of medical research features and patient age, one hidden layer of 19 neurons, and an output layer consisting of four neurons that forming the network output vector. Since the characteristic values can range from 0 to 10^{12}, and the network operation should not depend on the order of inputs to the neural network classifier construction is necessary to perform normalization of the input parameters. [3] We propose the following algorithm:

- Pathogens are combined into a single parameter
- The age indication takes the values according to the number in age group
- From the values of other attributes is taken logarithm

The proposed algorithm leads to the input values of the same dimension, which will improve the speed and quality of the learning process designed neural network. Combining groups of pathogens (fourth group) in one parameter allowed adding as input the age of the patient, without increasing the number of input parameters (29 input parameters). Density distributions of input parameters were not known in advance, so as the network learning algorithm method used back propagation, is the consistent teaching of layers of neurons. This method is a method of supervised learning, i.e., neural network adjusts the weights using the training set of multidimensional data, as well as the value of an expert diagnosis. [4] After studying communications with weights less than a predetermined threshold ε were removed.

Table 1. Neural network output vector values

	First output neuron value	Second output neuron value	Third output neuron value	Fourth output neuron value
Norm	1	0	0	0
First degree dysbiosis	0	1	0	0
Second degree dysbiosis	0	0	1	0
Third degree dysbiosis	0	0	0	1

Table 2. Age groups

Age group	Input parameter Value
0 – 23 hours	0
24– 167 hours	1
7 – 29 days	2
1 – 11 months	3
1 – 6 years	4
7 – 17 years	5
18 – 59 years	6
60 years and over	7

In were selected and systematized the results of bacteriological studies of gastrointestinal microflora as 2576 man of the experimental data with 29 characteristic features. A comparative analysis of the proposed method with other neural network techniques was made. Since the composition of the microflora in people of different ages have significant qualitative and quantitative differences, it was decided to make the training of neural networks for some age groups separately. [5] According to the experiment results, the three-layer perceptron showed the best results among the examined neural network techniques in the diagnosis of dysbiosis among all age groups. Plus in the table indicates availability of the chosen method, and minus the inability to distinguish between normal from the pathology.

Table 3. Neural networks comparative analysis

Age Network	0 – 23 hours	1 - 6 days	7 - 29 days	1 -11 m months	1 - 6 years	7 - 17 years	18 - 59 years	60 > years and over
Kohonen network	-	-	+	-	+	+	-	-
Vord network	+	+	+	-	-	+	-	+
Three-layer perceptron	+	+	+	+	+	+	-	+

The results suggest that our method can successfully classify the multidimensional biomedical object states. Three-layer perceptron works for almost all groups. Further studies will be conducted to improve the accuracy of diagnosis and performance of neural network learning.

References

1. Lomakina L.S., Lomakin D.V., Pozhidaeva A.S., Blazhnov I.V. Models and algorithms for classification of states biocenoses based on structural properties of multidimensional data «*Sistemy upravleniya i informatsionnye tekhnologii»* [Proc. of the “Control systems and information technology”], 2015, no. 3(61), pp. 22-27.

2. Khlivnenko L.V. *Praktika neurosetevogo modelirovaniya* [The practice of neural network modeling]. Voronezh, Voronezhskiy Gosudarstvenny Tekhnicheskiy Universitet Publ., 2015.
3. Bolshakov A.A., Karimov R.N. Metody obrabotki mnogomernykh dannykh i vremennykh ryadov [Processing methods of multidimensional data and time series]. Moscow, Goryachaya liniya – Telekom Publ., 2007.
4. Lomakina L.S., Soloveva I.V., Zelentsov S.A., Pozhidaeva A.S. Models and algorithms of diagnosing the state ecological community based on a priori statistical data. *Nauchno-tekhnicheskiy vestnik Povolzhya «Modeli i algoritmy diagnostirovaniya sostoyaniy biotsenoza na osnove apriornykh statisticheskikh dannykh»* [Proc. of the Kazan NIKGU "Scientific and Technical messenger of Volga"], 2013, pp251-256. (In Russian).
5. Vasserman F. *Neirokompyuternaya tekhnika: teoriya i praktika* [Neurocomputer Techniques: Theory and Practice]. Moscow, MIR, 1992.