Mathematical support for solving the classification problem using neural network algorithms

D T Mukhamedieva[1], N A Egamberdiev[2],J Sh Zokirov[3]

[1]Science and innovation center for information and communication technologies at the Tashkent university of information technologies named after Muhammad Al-Khwarizmi
[2]Science and innovation center for information and communication technologies at the Tashkent university of information technologies named after Muhammad Al-Khwarizmi
[3]Tashkent university of information technologies named after Muhammad Al-Khwarizmi

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

Abstract. The article considers the problem of increasing the efficiency of solving the classification of an unknown object in a neural network algorithm in the case of a data set. In such cases, it is important to select the number of neurons in the neural network layers and adjust the parameters to redistribute the errors. The mathematical apparatus and algorithm of the neural network in solving the classification problem are given in detail, and the results are analyzed using model problems.

1. Introduction

In the twentieth century, the concept of artificial intelligence was formed and a number of definitions of this concept were proposed. One of the first definitions that is widely accepted and still popular today is: "A way to encourage computers to think like humans" [10]. The origin of this definition was a popular test proposed by Alan Turing, one of the first researchers in English mathematics and computer science. In this test, the experimenter tries to determine if the person or computer program is a "person" or is exchanging messages using a remote keyboard.

There are several areas of artificial intelligence, each with its own algorithm. This article discusses the neural network algorithm. The neural network algorithm is modeled like the neural networks in the human brain. It models and processes the nonlinear connections between the input, hidden, and output layers. These sequences are algorithmic and programs have been developed in several programming languages. This algorithm is difficult to learn and apply, but good results can be obtained. Neural networks are actively helping to master new technologies and develop existing ones. Today, there is programming for unmanned vehicles in which neural networks analyze the environment in real time. Year after year, he is discovering new areas, including medicine. Google has a whole department that deals directly with artificial intelligence. Neural networks are lagging behind in areas such as high-resolution image creation, human speech creation, and in-depth analysis of video streams. Working with characters and recursive structures is also difficult for neurosystems.

Scientists have not yet fully decided in which direction to develop neural networks. The industry is trying to bring both neural networks as close as possible to the human brain model and to create technologies and conceptual schemes.

Neural networks are actively helping to master new technologies and develop existing ones. Today, there is programming for unmanned vehicles in which neural networks analyze the environment in real time. Year after year, it is discovering new areas, including medicine. Google has a whole department that deals directly with artificial intelligence. Neural networks are lagging behind in areas such as high-resolution image creation, people's speech creation, and in-depth analysis of video streams. Working with characters and recursive structures is also difficult for neurosystems.

Scientists have not yet fully decided in which direction to develop neural networks. The industry is trying to bring both neural networks as close as possible to the human brain model and to create technologies and conceptual schemes.

2. Formulation of the problem of classification

The classification is one of the mechanical sections devoted to solving the following problem. There are many objects (situations) that are somehow divided into classes. Given a limited set of objects, it is known to which classes they belong. This set is called a study sample. The class of the rest is unknown. The initial set is required to build an algorithm capable of classifying any object.

Defining an object class means specifying the class name to which this object belongs. Defining an object class is the determination of the class name given by the classification algorithm as a result of its application to that object. In mathematical statistics, classification problems are also referred to as discriminant analysis problems. The issue of classification in machine learning belongs to the department of managed education. There is also uncontrolled learning when the division of objects into classes in the learning sample is not specified and classification of objects based solely on their similarity is required. In this case, it is customary to talk about cluster problems, and the classes are called clusters, respectively.

Suppose that the set of study options is $x_{p1}, x_{p2}, \ldots, x_{pm_p} \in X_p, p = \overline{1,r}$, which is expressed in the following form. Here x_{pi} - N is the space vector of the dimensional characters, each object $x_{pi} = \left(x_{pi}^1, x_{pi}^2, \ldots x_{pi}^N\right), i = \overline{1,m_p}$, N - considered in the space of dimensional characters, X_p, $p = \overline{1,r}$ indicating a set of classes, it m_p $x_{p1}, \ldots, x_{p_{m_p}}$ consists of objects.

Problem: to build a decisive rule for determining the unknown object, that is, which class it belongs to in the set of educational options.

3. Mathematical support of neural network algorithms

Despite the diversity of neural networks, they all have common features. Thus, they are all composed of neurons that mimic many elements of the same type - the neurons of the brain - just like the human brain. Figure 1 shows a schematic of the neuron.

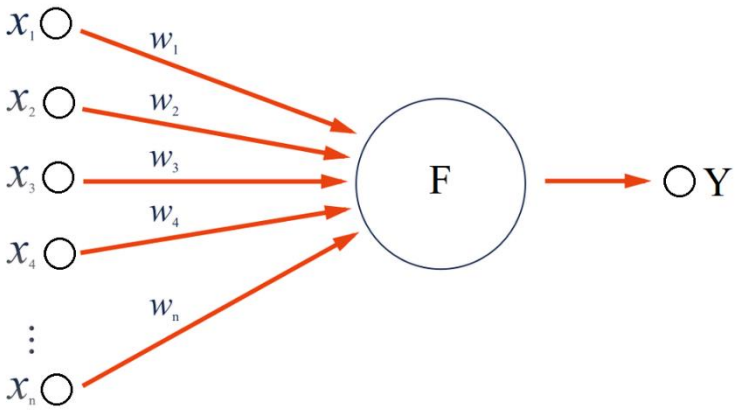

Figure 1. Schematic of neural networks.

As can be seen from the figure, the neurons in the input layer consist of synapses that connect with the nucleus, which processes the input signals and is the axon that connects the neuron to the neurons in the next layer. Each synapse has a weight that measures how much the input of the appropriate neuron affects its condition. The following state of the neuron is determined by the formula.

$$F = \sum_{i=1}^{n} x_i w_i \tag{1}$$

Here,

n - number of input neurons;

x_i - the value of i -input neurons;

w_i - i - synaptic weight.

The value of the neuron axon is then determined according to the following formula.

$$Y = f\left(S\right) \tag{2}$$

Here, f is a function called activation. Often the sigmoid is used as an activation function with the following form:

$$f\left(x\right) = \frac{1}{1+e^{-ax}} \tag{3}$$

The main advantage of this function is that it differs along the entire abscissa axis and has a very simple product:

$$f'(x) = \alpha f(x)\left(1 - f(x)\right) \tag{4}$$

As the parameter decreases, the sigmoid flattens out and becomes a horizontal line, at a = 0 at 0.5.

Redistribution of errors in neural networks is a powerful tool for forecasting and qualitative analysis. They got this name because of the algorithm used in the return networks, in which the error propagates from the output layer to the input layer, i.e. in the direction opposite to the direction of signal propagation during normal operation of the network. In the general case, the task of teaching neural networks is reduced to finding a specific functional relationship Y = F (X), where X is the input and Y is the output vectors. In the general case, such a problem with a limited set of access problems has an infinite set of solutions. During the training, the task is to minimize the objective function of the neural network error to limit the search area, which is found using the least squares method:

$$E(w) = \frac{1}{2}\sum_{j=1}^{p}(y_j - d_j)^2 \tag{5}$$

Here,

y_i - j - output neuron value;

d_i - j j is the ideal value of the output;

p - the number of neurons in the output layer.

The neural network is trained using the gradient drop method, i.e. the weight in each iteration is changed according to the following formula:

$$\square w_{ij} = -\eta \cdot \frac{\partial E}{\partial w_{ij}} \tag{6}$$

Where η is the parameter that determines the learning speed.

$$\partial_j^{(n)} = \left[\sum_k \partial_k^{(n+1)} \cdot w_{jk}^{(n+1)}\right] \cdot \frac{dy_j}{dS_j} \tag{7}$$

It is not difficult to find a neural network for the last layer because we know the target vector, i.e. the vector of the values that the neural network must create for a given set of input values.

$$\partial_j^{(N)} = \left(y_j^{(N)} - d_j\right) \cdot \frac{dy_j}{dS_j} \tag{8}$$

Finally, we write formula (6) in an expanded form.

$$\Delta w_{ij}^{(n)} = -\eta \cdot \delta_j^{(n)} \cdot x_i^n \tag{9}$$

Now consider the algorithm for studying a complete neural network:

Step 1: send to the input part of the neural network and determine the output values of the network neurons;

Step 2: Calculate for the output layer of the neural network according to formula (8) and calculate the change in the weights of n output layers according to formula (9);

Step 3: Formulas (7) and (9) for the remaining layers of the neural network, respectively and $\Delta w_{ij}^{(n)}$ calculate on, n = N-1..1;

Step 4: Adjust the weights of all neural networks;

$$w_{ij}^{(n)}(t) = w_{ij}^{(n)}(t-1) + \Delta w_{ij}^{(n)}(t) \tag{10}$$

Step 5: If the error is significant, go to step 1.

In step 2, the vectors from the training sequence are randomly presented to the grid.

4. Computing Experiment

Based on the above algorithm, kaggle.com (kaggle allows users to find and publish data sets, study and create models, work with scientists and mechanical engineers) is able to perform computational tests using the data sets at kaggle.com. 'tkazildi.

Two types of real databases were selected for the experiment (Table 1).

Table 1. Content of data sets.

Data set name	Attributs	Objects	Class
IRIS	4	150	3
DIABETES	8	768	2

The classification problem was solved using neural networks using an iris data set. To do this, the following scheme was selected for teaching the Iris data set using the above algorithm (Figure 2). There are 4 input neurons (corresponding to the number of attributes), 5 neurons in the latent layer, and 3 neurons at the output (corresponding to the number of classes).

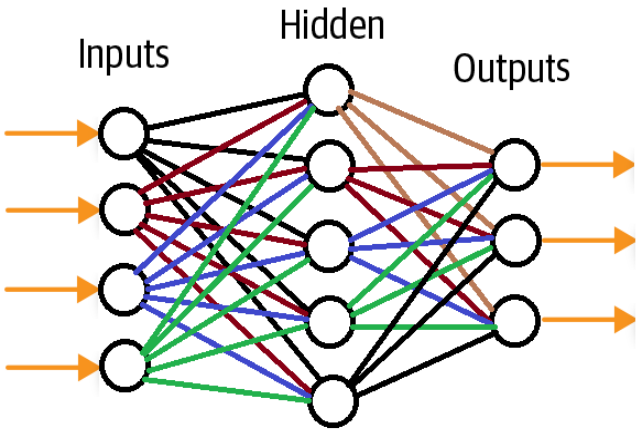

Figure 2. Schematic of neural networks for the iris data set.

Table 2 shows the results obtained using the proposed algorithm and weka software.

Table 2. Results.

Algorithm name	IRIS		DIABETES	
	Correctly	%	Correctly	%
Logistic	144	96	593	77
KStar	142	95	531	69
Decision tree (J48)	144	96	567	74
Decision table	139	93	547	71
Neural network	144	96	637	83

The experiments were performed on the basis of 10 cross-validations for each data set.

5. Conclusion

The article describes the solution of artificial intelligence problems based on neural network algorithms. The mathematical apparatus and algorithm of the neural network are given. Based on the proposed algorithm, a program was written in the Java programming language to analyze the Iris model problem, select the number of neurons in the neural network layers, and adjust the parameters for redistribution of errors and compare the results with several classification algorithms.

Artificial intelligence technologies can be used to solve problems in areas such as unmanned vehicle programming, medical diagnosis, image detection, human voice detection, and in-depth analysis of video streams.

References

A. Aliev R.A., Aliev R.R. 2001 The theory of intelligent systems and its application *Publishing House of the Chashyoglu* Baku 720
B. Muhamediyeva D.T. and Niyozmatova N.A. 2019 Problems of constructing models of intellectual analysis of states of weakly formalizable processes *Journal of Physics* **1210**
C. Kamilov M.M., Khujaev O.K., Egamberdiev N.A. 2019 The Method of Applying the Algorithm of Calculating Grades for Finding Similar Diagnostics in Medical Information Systems *International Journal of Innovative Technology and Exploring Engineering (IJITEE)* **8-6S** 722-724

D. Mukhamedieva D., Khasanov U., Egamberdiev N. 2020 Presentation of preferences in multi-criterional tasks of decision-making *IOP Conf. Series*:*Journal of Physics***1441**
E. Muhamediyeva D.T., Egamberdiev N.A. 2019 Algorithm and the Program of Construction of the Fuzzy Logical Model *International Conference on Information Science and Communications Technologies (ICISCT)* Tashkent 1-4.
F. Egamberdiyev N., Jo'rayev Z.Sh. 2019 Neyron to'rlari yordamida identifikatsiyalash *Materiali respublikanskoy nauchno-texnicheskoy konferentsii "Innovatsionnыe idei v oblaste IKT i programmnogo obespecheniya"* Samarkand 281-284.
G. Egamberdiyev N., Hasanov U. 2019 Neyron to'rlari yordamida boshqaruv *Materiali respublikanskoy nauchno-texnicheskoy konferentsii "Innovatsionnыe idei v oblaste IKT i programmnogo obespecheniya"* Samarkand 278-281.