Gradient algorithms for artificial neuron network teaching

1st Erkin Uljaev
faculty of Electronics and Automation
Tashkent State Technical University
named after Islam Karimov
Tashkent, Uzbekistan
e.uljaev@mail.ru

2nd Shohrukh Narzullayev
faculty of Electronics and Automation
Tashkent State Technical University
named after Islam Karimov
Tashkent, Uzbekistan
narzullayevsh@tdtu.uz

3rd Utkirjon Ubaydullaev
faculty of Electronics and Automation
Tashkent State Technical University
named after Islam Karimov
Tashkent, Uzbekistan
utkir2005@mail.ru

4th Elyor Khudoyberdiyev
faculty of Electronics and Automation
Tashkent State Technical University
named after Islam Karimov
Tashkent, Uzbekistan
elyor125rqt11@gmail.com

Abstract

This article is devoted to the study and application of the basic features of algorithms for "teaching" multilayer artificial neural networks, in particular, gradient algorithms. The article states that scientific research is being carried out on the development of intelligent measuring devices for controlling the moisture content of scattering materials based on functions such as automatic adjustment of measuring range, self-calibration, linear static characteristics, high measurement accuracy and reliability, data processing, decision making. As a result of the research, it was considered that the requirements for measuring the moisture content of scattering materials by control and management systems can be met by high-precision moisture meters developed using intelligent technologies. Today, a wide range of new opportunities for the construction of intelligent systems for the control of technological processes are developing through the use of artificial neural networks. Symbols are widely used in issues such as recognition, prediction and diagnostics, optimization, signal processing under the influence of noise. The authors have built an artificial neural network for an intelligent device that measures the moisture content of scattering materials. The main parameters of the artificial neural network were determined and "training" was carried out on the basis of gradient algorithms. In an intelligent device that measures the moisture content of scattering materials, the factors that affect the measurement accuracy are minimized as a result of gaining knowledge base through experiments. Based on the above results, it is stated that an artificial neural network has been used in an intelligent device.

Keywords: Artificial neural network, gradient algorithm, humidity, weight coefficients, constant shift (bias).

1. INTRODUCTION

In the era of globalization, where science and information and communication technologies are rapidly developing and leading to the fourth industrial revolution, public and social governance, economics, industry, social protection, education, science, medical diagnostics, agriculture, defense and security, tourism in developed countries and the widespread use of modern information technology and artificial intelligence in many areas.

The solution of technical problems such as management, control and measurement of technological processes in all sectors of industrial production in the world is being intelligent. The intellectualization of such technological processes requires the expansion of the functional capabilities of technical means of management, control, and measurement, the improvement of technical characteristics, the solution of new scientific and technical problems of efficient use of energy resources. Therefore, a class of intelligent devices and sensors is emerging. They have the ability to self-adapt to the operating principle, images of input signals, technical condition and external factors affecting the measuring system due to changes in control and measurement conditions [1].

Intelligent devices and sensors differ from other analogues of this type in that they have functions such as automatic adjustment of the measuring range, self-calibration, data processing, decision making [2,3].

Nowadays, measuring instruments that measure temperature, humidity, pressure and other quantities are rapidly being developed by specialists in various fields. However, the accuracy of measuring instruments used in many industries is insufficient, and the measurement error differs from the specified "range". By optimizing the controlled and controlled parameters, it is possible to improve the quality of manufactured products and reduce defects. This makes it necessary to create new measuring instruments.

2. PURPOSE OF THE WORK

Construction of an artificial neural network based on gradient algorithms for an intelligent device that measures the moisture content of scattering materials.

3. PROBLEM STATEMENT

It is known that moisture is one of the most important technological parameters, which directly affects the cost, technological structure, and other properties of substances and materials [4-8]. High-precision moisture meters, which are involved in moving technological processes, are used directly to organize the quality control of moisture in the process of spraying materials (wheat, barley, rice, millet, etc.). Therefore, the selection of measuring instruments for each control system, the study of their principles of operation, design, and other parameters is of great importance [9-12].

IEECP'21, July 29-30, 2021, Silicon Valley, San Francisco, CA – USA

DAI : https://sci-index.com/DAI/2021.99101/IEECP/14904897

The following requirements can be set for devices used to measure the moisture content of spray materials:

• measuring device - measures moisture in the technological process and does not damage the measured substance;

• the absence of elements in the design of the measuring device that affect the accuracy of the measurement of scattering materials;

• measuring device - low sensitivity to the external environment and a stable measuring system;

• Measuring device - should be an intelligent device with a microprocessor that measures with high accuracy on the basis of modern technologies.

An analysis of the available moisture metering devices shows that the above requirements for measuring the moisture content of spray materials by control and management systems can be met by moisture metering devices developed using intelligent technology.

Today, scientific research is being carried out on the development of intelligent measuring devices for automatic control of the measuring range, calibration, linear static characteristics, high measurement accuracy and reliability, data processing, decision-making to control the moisture content of dispersed materials [11,13- 15].

Factors that directly affect the measuring accuracy of the device for measuring the moisture content of scattering materials can include changes in the density of the material under study, dielectric constant, temperature, the chemical composition of the substance, and other parameters. These factors cause errors in the measurement process. Therefore, these errors should be minimized as much as possible.

Extensive new possibilities for building intelligent process control systems are being developed through the use of artificial neural networks, and they are widely used in issues such as image recognition, prediction, and diagnostics, optimization, signal processing under the influence of noise [16]. In an intelligent device based on artificial neural networks that measure the moisture content of scattering materials, the factors that affect measurement accuracy are brought closer to the minimum value as a result of gaining a knowledge base through experiments.

3.1. SOLUTION METHOD OF THE PROBLEM

Properly distributed artificial neural networks are now widely used to solve problems in intelligent systems. Therefore, we use properly distributed cross-linked neural networks in an intelligent device that measures the moisture content of scattering materials. Because the system is static, all neurons in the given vectors of the signals of the neurons in the input layer form a single value of the signal vector in the output layer. A neural network consists of a number of inputs and outputs, consisting of a set of interconnected neurons, and performs a nonlinear change. A model of an artificial neural network performing the given problem is shown in Figure 1 [17-20].

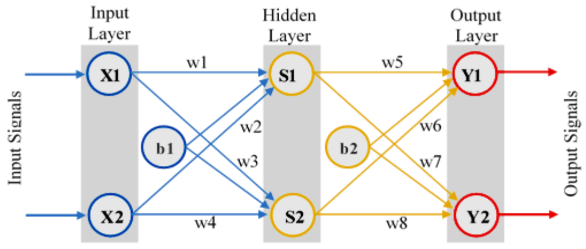

Figure 1. The structure of an artificial neural network

Several algorithms for teaching multilayer artificial neural networks have been developed. We can divide these algorithms into two groups [16-18, 21-22]:

- gradient algorithms;

- stochastic algorithms.

The algorithms of the first group include the calculation of the product of the error function and the neural networks that correct the weights according to the found product. Each subsequent step is directed towards the antigradient of the error function. The basis of these algorithms is the algorithm of inverse distribution of errors. This algorithm serves to minimize the deviation of the actual values of the neurons in the neural network output layer from the required value. Errors are found using the quadratic error function for each output neuron, and they give the total error of the output neuron. Examples of this algorithm are moment study, autonomous gradient algorithm, and second-order methods.

In stochastic algorithms, however, random changes in weight and shear (bias) coefficients are made while maintaining the previous changes. This leads to an improvement in neural network parameters. Stochastic algorithms include random directional search algorithms, Cauchy learning algorithms, Boltsman learning algorithms, evolutionary algorithms. The main disadvantages of this algorithm include the high demand for technical means, the long duration of learning time, and so on. Nevertheless, these algorithms are used because they provide global optimization. In gradient algorithms, however, local minimums of the error function can be found. Currently, hybrid algorithms based on the combined use of the above two algorithms are also used.

3.2. RESEARCH METHODS AND THE RECEIVED RESULTS

Frequency input signals describing changes in temperature and capacitance (dielectric constant) of the scattering materials under study (e.g., wheat) were taken. The actual values of the quantities received as the input signal vary in frequency and temperature of range through the corresponding functions. To perform artificial neural network training, we interact with input layer neurons through and input signals equal to and.

Each of the neurons in the hidden layer is associated with neurons in the input layer that have a certain weight. It describes the interdependence of the elements of the layers and determines the magnitude of the communication efficiency of the elements. We know that each element of the neural network operates in a discrete-time and forms the resulting signal based on the received signal. Table 1 shows the random weight coefficients of the input and hidden layer neurons and the values of the constant shifts (bias).

According to Figure 1, the input signals coming to the hidden layer neurons are the cumulative signal by their respective weights [17, 23-26].

$$S = w \times X + b \qquad (1)$$

Where - weight coefficients of neurons; - signal given by neurons of the input layer; - constant displacement (bias); - the sum of input signals.

We calculate the sum of the input signals using the random weight coefficients given in Table 1.

$$S_1 = w_1 \cdot X_1 + w_2 \cdot X_2 + b_1 ,$$
$$S_1 = 0{,}20 \cdot 0{,}10 + 0{,}25 \cdot 0{,}70 + 0{,}60 = 0{,}795 ,$$
$$S_2 = w_3 \cdot X_1 + w_4 \cdot X_2 + b_1 ,$$
$$S_2 = 0{,}30 \cdot 0{,}10 + 0{,}35 \cdot 0{,}70 + 0{,}60 = 0{,}875$$

Table 1.

Weight coefficients of layers of the artificial neural network, constant shifts (bias), and values of required output signals

X_1	X_2	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	b_1	b_2	Y_1^*	Y_2^*
0,10	0,70	0,20	0,25	0,30	0,35	0,40	0,45	0,50	0,55	0,60	0,65	0,35	0,65

Using activation functions, we convert the results in hidden layer neurons into output signals. Several types of activation functions are currently used in practice [27,28]. We use the activation function in the sigmoid view. The activation function in the sigmoid view is nonlinear and the given multidimensional function can be polished with arbitrary precision per unit cross-section [27,29,30]. As a result, the output layer also has a nonlinear characteristic. It has the following appearance:

$$f(S)=\frac{1}{1+e^{-S}}$$

From the above expression, we calculate the output value of each neuron in the hidden layer.

$$f(S_1)=\frac{1}{1+e^{-S_1}}=\frac{1}{1+e^{-0,795}}=0{,}6889039179\ 7695\ ,$$

$$f(S_2)=\frac{1}{1+e^{-S_2}}=\frac{1}{1+e^{-0,875}}=0{,}7057850278\ 3701$$

The resulting values are the output value for the hidden layer neurons and the input signals for the output neurons in the second layer.

We also determine the values of the neurons in the next layer using the random weight coefficients in Table 1 using expression (1).

$$S_3=w_5\cdot f(S_1)+w_6\cdot f(S_2)+b_2,$$

$$S_3=0{,}40\cdot 0{,}6889039179\ 7695+0{,}45\cdot 0{,}7057850278\ 3701+\ ,$$
$$+0{,}65=1{,}2431648297\ 1744$$

$$S_4=w_7\cdot f(S_1)+w_8\cdot f(S_2)+b_2,$$

$$S_4=0{,}50\cdot 0{,}6889039179\ 7695+0{,}55\cdot 0{,}7057850278\ 3701+$$
$$+0{,}65=1{,}3826337242\ 9883$$

Equation (2) calculates the values of the output signals from the expression. We find the true values of the output signals using the activation function in the sigmoid view. We know that the activation function in the sigmoid view is both a negative and a monotonous growth function. Therefore, the values of the output signals also consist of positive results.

$$Y_1=f(S_3)=\frac{1}{1+e^{-S_3}}=\frac{1}{1+e^{-1,24316482971744}}=0{,}7761144182\ 1532$$

,

$$Y_2=f(S_4)=\frac{1}{1+e^{-S_4}}=\frac{1}{1+e^{-1,38263372429883}}=0{,}7994136548\ 4440$$

These values are the output signals of the artificial neural network we are looking for.

The most complex process in building multilayer artificial neural networks is “training” (Figure 2). The unknown parameters of "training" neurons are the identification of weight and bias coefficients. In the teaching of multilayer neural networks, gradient search methods are used to minimize the criterion functions that depend on the parameters of the neurons.

This process is iterative, and in each iteration, all the coefficients of the network are found first for the output layer, then for the hidden layer, and for the input layer (the method of inverse distribution of errors) [16-17, 31-33]. Other forms of minimum search, such as genetic algorithms and the least-squares method, are also used.

We calculate the errors for each output neuron using the error inverse distribution method.

$$E_\Sigma=\sum\frac{1}{2}(Y*-Y)^2 \qquad (3)$$

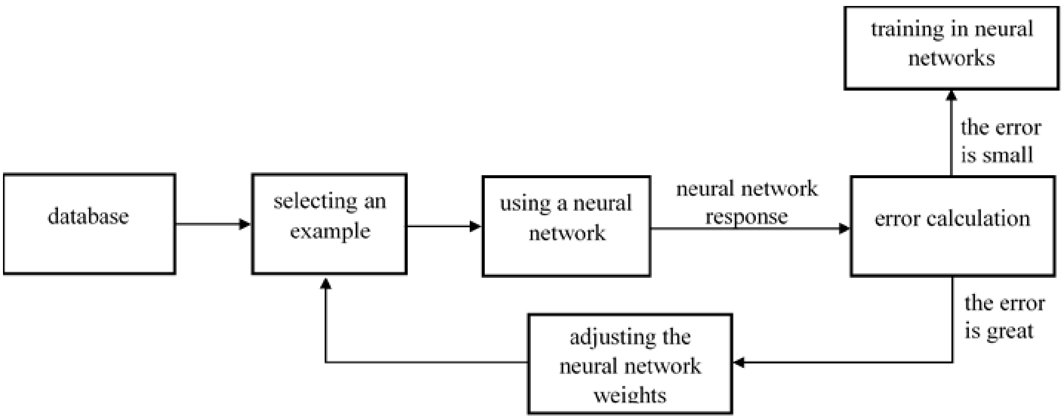

Figure 2. The process of "training" the neural network

Where, and are the required and actual output values of the neural network accordingly. The sum of the errors is performed on the output layer and the whole set of neurons in the neural network training.

We calculate the errors of the neurons at the output of the neural network through the expression (3).

$$E_1=\frac{1}{2}(Y_1^*-Y_1)^2=\frac{1}{2}(0{,}35-0{,}77611441821532)^2=0{,}09078674870549$$

,

$$E_2=\frac{1}{2}(Y_2^*-Y_2)^2=\frac{1}{2}(0{,}65-0{,}79941365484440)^2=0{,}01116222012698$$

The sum of these errors gives a total error.

$$E_\Sigma=E_1+E_2=0{,}09078674870549+0{,}01116222012698=0{,}10194896883247$$

The purpose of using the inverse error distribution method is to update each weighting coefficient and bias in the neural network. This is because we can minimize the total error in each output neuron and by bringing the actual values of the neurons in the output layer of the neural network closer to the required values.

Now we determine that each weighting coefficient in the output layer of the artificial neural network affects the total error (Fig. 3.).

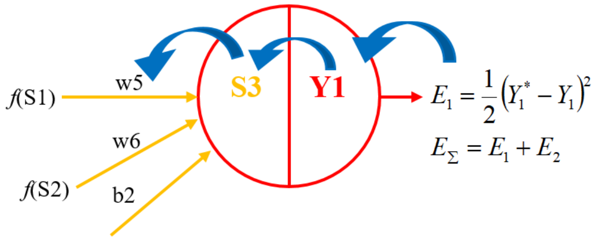

Figure 3. Structural view of the effect of weighting coefficients on the output of the total error

consider the effect of the weight coefficient w_5 on the overall error.

$$\frac{\partial E_\Sigma}{\partial w_5} = \frac{\partial E_\Sigma}{\partial \mathrm{Y}_1} \cdot \frac{\partial Y_1}{\partial S_3} \cdot \frac{\partial S_3}{\partial w_5} \quad (4)$$

We can calculate this expression by dividing it into parts, such as the complex rule product (chain rule). First, we determine the effect of the output signal on the overall error.

$$E_\Sigma = E_1 + E_2 = \frac{1}{2}\left(Y_1^* - Y_1\right)^2 + \frac{1}{2}\left(Y_2^* - Y_2\right)^2,$$

$$\frac{\partial E_\Sigma}{\partial Y_1} = -2 \cdot \frac{1}{2}\left(Y_1^* - Y_1\right)^{2-1} + 0 = -\left(Y_1^* - Y_1\right) = -(0{,}35 - 0{,}77611441821532) =$$
$$= 0{,}42611441821532$$

We express the effect of the input signal S_3 on the output signal Y_1 using the following product

$$Y_1 = f\left(\mathrm{S}_3\right) = \frac{1}{1+e^{-S_3}},$$

$$\frac{\partial Y_1}{\partial S_3} = Y_1\left(1 - Y_1\right) = 0{,}7761144182\ 1532 \cdot \left(1 - 0{,}7761144182\ 1532\right) = 0{,}1737608280\ 5362$$

Finally, we consider the effect of the weight coefficient on the overall expression.

$$S_3 = w_5 \cdot f\left(S_1\right) + w_6 \cdot f\left(S_2\right) + b_2,$$

$$\frac{\partial S_3}{\partial w_5} = f\left(S1\right) + 0 + 0 = 0{,}6889039179\ 7695$$

Based on expression (4), we combine the results found above.

$$\frac{\partial E_\Sigma}{\partial w_5} = 0{,}42611441821532 \cdot 0{,}17376082805362 \cdot$$
$$\cdot 0{,}68890391797695 = 0{,}05100781986798$$

We use the following expression to find the new weight coefficient.

$$w_5^* = w_5 - \eta \cdot \frac{\partial E_\Sigma}{\partial w_5}$$

where, η is the training coefficient. η is a gradient that allows you to control the amount of weight coefficient in each iteration, which is one of the parameters of the neural network training algorithm [16-17, 34-37]. Its value $\left(0 \le \eta \le 1\right)$ varies in the range. Selecting a zero value is useless because in this case, the value of the weight coefficients does not change. Learning algorithm runs faster (less time is required to minimize error function) in the interval of $0{,}7 \le \eta \le 1$, but the accuracy of neural network error function minimization can be reduced, which exacerbates the training error.

At small values of the training coefficient, the $0{,}1 \le \eta \le 0{,}3$ number of cycles of finding the extremum increases, but the accuracy of minimizing the error function is higher. This in turn can lead to a reduction in teaching error. In practice, the value of the training coefficient is usually chosen experimentally or arbitrarily. Assume that the training coefficient is equal to $\eta = 0{,}5$.

$$w_5^* = w_5 - \eta \cdot \frac{\partial E_\Sigma}{\partial w_5} = 0{,}40 - 0{,}5 \cdot 0{,}05100781986798 =$$
$$= 0{,}37449609006601$$

We also perform the above sequence of actions for the remaining neurons w_6, w_7, w_8 in the output layer.

The effect of the weight coefficient w_6 on the total error is as follows.

$$\frac{\partial E_\Sigma}{\partial w_6} = \frac{\partial E_\Sigma}{\partial \mathrm{Y}_1} \cdot \frac{\partial Y_1}{\partial S_3} \cdot \frac{\partial S_3}{\partial w_6},$$

$$\frac{\partial E_\Sigma}{\partial w_6} = 0{,}05225773090557$$

From this, we calculate a new weight coefficient w_6^*.

$$w_6^* = w_6 - \eta \cdot \frac{\partial E_\Sigma}{\partial w_6} = 0{,}42387113454722$$

For the weight coefficient w_7, we also perform the above sequence of actions.

$$\frac{\partial E_\Sigma}{\partial w_7} = \frac{\partial E_\Sigma}{\partial \mathrm{Y}_2} \cdot \frac{\partial Y_2}{\partial S_4} \cdot \frac{\partial S_4}{\partial w_7},$$

$$\frac{\partial E_\Sigma}{\partial w_7} = 0{,}01650524105286$$

The new value of the weight coefficient w_7^* is as follows:

$$w_7^* = w_7 - \eta \cdot \frac{\partial E_\Sigma}{\partial w_7} = 0{,}49174737947357$$

Determine the final weight coefficient w_8 in the output layer:

$$\frac{\partial E_\Sigma}{\partial w_8} = \frac{\partial E_\Sigma}{\partial \mathrm{Y}_2} \cdot \frac{\partial Y_2}{\partial S_4} \cdot \frac{\partial S_4}{\partial w_8},$$

$$\frac{\partial E_\Sigma}{\partial w_8} = 0{,}01690969046912$$

Based on the results found, we calculate the new weight coefficient .

$$w_8^* = w_8 - \eta \cdot \frac{\partial E_\Sigma}{\partial w_8} = 0{,}54154515476544$$

We perform network updates after we have calculated for neurons in the hidden layer as well. When finding new values of the weight coefficients in the hidden layer, we use the values of the initial weight coefficients and shifts (bias), rather than the values found above. Figure 4 shows the effect of each weighting coefficient and bias on the total error in the hidden layer of the artificial neural network.

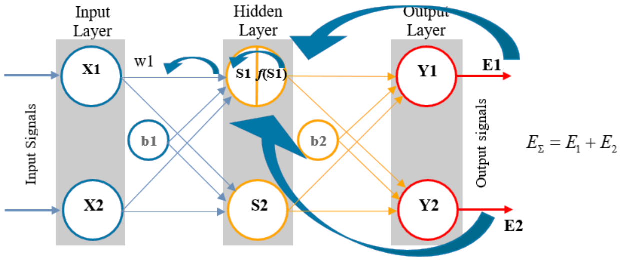

Figure 4. Structural view of the effect of hidden layer weight coefficients on the total error

Through the calculations, we found new values of weight coefficients representing the connections of neurons in the hidden and output layers. Now we also need to calculate the new values of the weight coefficients that represent the connections between the input and hidden layer neurons.

The hidden layer weight coefficients affect all output signals. Therefore, their effect on all neurons in the output layer should be taken into account [17-18,23].

$$\frac{\partial E_\Sigma}{\partial w_1} = \frac{\partial E_\Sigma}{\partial f(S_1)} \cdot \frac{\partial f(S_1)}{\partial S_1} \cdot \frac{\partial S_1}{\partial w_1} \quad (5)$$

We know that $f(S_1)$ affects both errors in the output signals. Therefore, the effect of the expression $\frac{\partial E_\Sigma}{\partial f(S_1)}$ on the error in both output signals must be taken into account.

$$\frac{\partial E_\Sigma}{\partial f(S_1)} = \frac{\partial E_1}{\partial f(S_1)} + \frac{\partial E_2}{\partial f(S_1)} \quad (6)$$

We calculate the product of E_1 to.

$$\frac{\partial E_1}{\partial f(S_1)} = \frac{\partial E_1}{\partial S_3} \cdot \frac{\partial S_3}{\partial f(S_1)} \quad (7)$$

we find the value of $\frac{\partial E_1}{\partial S_3}$ by the results calculated above

$$\frac{\partial E_1}{\partial S_3} = \frac{\partial E_1}{\partial Y_1} \cdot \frac{\partial Y_1}{\partial S_3} = 0{,}4261144182\;1532 \cdot 0{,}1737608280\;5362 = 0{,}0740419941\;5468$$

The value of $\frac{\partial S_3}{\partial f(S_1)}$ the expression is calculated as follows.

$$S_3 = w_5 \cdot f(S_1) + w_6 \cdot f(S_2) + b_2,$$

$$\frac{\partial S_3}{\partial f(S_1)} = w_5 = 0{,}40\ .$$

We combine the results found in expression (7).

$$\frac{\partial E_1}{\partial f(S_1)} = 0{,}0740419941\;5468 \cdot 0{,}40 = 0{,}0296167976\;6187$$

Now we calculate the second component of expression (6).

$$\frac{\partial E_2}{\partial f(S_1)} = \frac{\partial E_2}{\partial S_4} \cdot \frac{\partial S_4}{\partial f(S_1)} \quad (8)$$

For this expression, too, we perform the sequence of actions in expression (7).

$$\frac{\partial E_2}{\partial S_4} = \frac{\partial E_2}{\partial Y_2} \cdot \frac{\partial Y_2}{\partial S_4} = 0{,}1494136548\;4440 \cdot 0{,}1603514632\;9272 = 0{,}0239586981\;9021$$

$$S_4 = w_7 \cdot f(S_1) + w_8 \cdot f(S_2) + b_2,$$

$$\frac{\partial S_4}{\partial f(S_1)} = w_7 = 0{,}50\ .$$

Determine the value of the expression (8).

$$\frac{\partial E_2}{\partial f(S_1)} = 0{,}0239586981\;9021 \cdot 0{,}50 = 0{,}0119793490\;9511$$

We put these results in expression (6) and calculate the value of $\frac{\partial E_\Sigma}{\partial f(S_1)}$.

$$\frac{\partial E_\Sigma}{\partial f(S_1)} = 0{,}02961679766187 + 0{,}01197934909511 = 0{,}04159614675698\cdot$$

From the activation function in the form of sigmoid in expression (5) we find the product of the sum of the input signals S_1.

$$f(S_1) = \frac{1}{1+e^{-S_1}},$$

$$\frac{\partial f(S_1)}{\partial S_1} = f(S_1)\cdot(1 - f(S_1)) = 0{,}68890391797695 \cdot$$

$$\cdot(1 - 0{,}68890391797695) = 0{,}21431530977296$$

Here, we calculate the effect w_1 on the total error using the sum of the input signals.

$$S_1 = w_1 \cdot X_1 + w_2 \cdot X_2 + b_1,$$

$$\frac{\partial S_1}{\partial w_1} = X_1 = 0{,}10\ .$$

In that case, the value $\frac{\partial E_\Sigma}{\partial w_1}$ is equal to the following.

$$\frac{\partial E_\Sigma}{\partial w_1} = 0{,}04159614675698 \cdot 0{,}21431530977296 \cdot 0{,}10 = 0{,}00089146910775\cdot$$

From the results obtained we find the new weight coefficient of w_1.

$$w_1^* = w_1 - \eta \cdot \frac{\partial E_\Sigma}{\partial w_1} = 0{,}20 - 0{,}5 \cdot 0{,}00089146910775 = 0{,}19955426544612$$

.

We also perform the above steps for the weight coefficients w_2, w_3 and w_4.

$$\frac{\partial E_\Sigma}{\partial w_2} = \frac{\partial E_\Sigma}{\partial f(S_1)} \cdot \frac{\partial f(S_1)}{\partial S_1} \cdot \frac{\partial S_1}{\partial w_2} = 0{,}0062402837\;5431$$

$$w_2^* = w_2 - \eta \cdot \frac{\partial E_\Sigma}{\partial w_2} = 0{,}25 - 0{,}5 \cdot 0{,}00624028375431 = 0{,}24687985812285$$

$$\frac{\partial E_\Sigma}{\partial w_3} = \frac{\partial E_\Sigma}{\partial f(S_2)} \cdot \frac{\partial f(S_2)}{\partial S_2} \cdot \frac{\partial S_2}{\partial w_3} = 0{,}00086375447928,$$

$$w_3^* = w_3 - \eta \cdot \frac{\partial E_\Sigma}{\partial w_3} = 0{,}30 - 0{,}5 \cdot 0{,}00086375447928 = 0{,}29956812276036\cdot$$

$$\frac{\partial E_\Sigma}{\partial w_4} = \frac{\partial E_\Sigma}{\partial f(S_2)} \cdot \frac{\partial f(S_2)}{\partial S_2} \cdot \frac{\partial S_2}{\partial w_4} = 0{,}00604628135496,$$

$$w_4^* = w_4 - \eta \cdot \frac{\partial E_\Sigma}{\partial w_4} = 0{,}35 - 0{,}5 \cdot = 0{,}00604628135496 = 0{,}34697685932252$$

We found new values of all weight coefficients. Now we can make updates. Initially as input signals $X_1 = 0{,}10$ and $X_2 = 0{,}70$ when we gave the values, the error in the neural network was 0.1019489683247. After we updated the weight coefficients, the error was 0.1019458256412. We continue the process to the allowable value of the error.

4. Conclusions

Studies have shown that the use of artificial neural networks in an intelligent device that measures the moisture content of scattering materials can significantly increase the accuracy and timing of measurements. The layers of the artificial neural network and the basic parameters of the neurons were found separately for each period, and an activation function in the form of a sigmoid was used. In the creation of multilayer artificial neural networks, training based on gradient algorithms with a "teacher" was used. As a result, neuro model data generalization and training time was reduced.

REFERENCES

[1] Michael T., McGrath, Cliodhna Ni Scanaill. Sensor Technologies. *Apress Media*. LCC, New York (NY). USA, 2014. – 321 p.

[2] Gaskarov D.V. Intellektualnыe informatsionnыe sistemы[Intellectual information system].- M.: Izd-vo: Vыssh. shk., 2003 g. - 432 s.

[3] Randy Frank. Understanding Smart Sensors. *Artech House*, Boston London, MA, USA, 2000. – 389 p.

[4] M.A.Berliner, Izmereniya vlazhnosti [Moisture measurements]. Izd. 2-e, pererab. Moskva: Energiya, 1973. 400 p. (in Russian).

[5] Theory and practice of rapid control of humidity of solid and liquid materials. Ed. E.S. Krichevsky. - M.: Energy, 1980.

[6] M.V. Kulakov, Tekhnologicheskie izmereniya i pribory dlya khimicheskikh proizvodstv [Technological measurements and devices for chemical production]. Izd. 3-e, pererab. i dop. Moskva: Mashinostroenie, 1983. 424 p. (in Russian).

[7] E.Uljaev, U.M.Ubaydullaev, Sh.N.Narzullaev, E.F.Xudoyberdiev, F.G.Haydarov, "Classification of Detectors of Capacitive Humidity Transducers of Bulk Materials", *International Journal of Advanced Science and Technology*, vol. 29(11s), pp.1949 – 1953, 2020. Retrieved from http://sersc.org/journals/index.php/IJAST/article/view/21858.

[8] E.Uljaev, Sh.N.Narzullaev, F.G.Xaydarov, "K voprosu razrabotki intellektualnogo kombinirovannogo ustroystva kontrolya vlajnosti sipuchix materialov"[On the issue of developing an intelligent combined device for controlling the moisture content of bulk materials], *Materiali mejdunarodnoy nauchnoy konferensii «Innovatsionnie resheniya injenerno-texnologicheskix problem sovremennogo proizvodstva»*, vol. 2, 14-16 noyabrya 2019 g, Buxara-2019, pp.204-206. (in Russian).

[9] Arbuzov V.P., Mishina M.A., Belыnseva P.N. Sistematizatsiya sposobov fazovogo razdeleniya kanala izmeritelnыx sepey yemkostnыx datchikov // Izvestiya vыsshix uchebnыx zavedeniy. Povoljskiy region. Texnicheskie nauki, 2017. № 1(41). S. 85-95.

[10] Frayden Dj. Sovremennыe datchiki: spravochnik. –M.: Texnosfera, 2005 g. -592 s.

[11] E. Uljaev, U.M. Ubaydullaev, and Sh.N. Narzullaev, "Capacity transformer of coaxial and cylindrical form of humidity meter", *Chemical Technology, Control and Management*. 4(4):23-30, 2020. DOI: https://doi.org/10.34920/2020.4.23-30.

[12] Umarov E.O., Mardonov U.T., Shoazimova U.Kh. Influence of the Magnetic Field on the Viscosity Coefficient of Lubricoolant that is used in the Cutting Process. *International Journal of Mechatronics and Applied Mechanics*,8(2):144-149 . https://www.doi.org/10.17683/ijomam/issue8.50.

[13] Ernest O.D. Measurement Systems: Application and design. McGraw - Hill. Higher Education. - New York, 2004 – 1078 p.

[14] E.Uljayev, U.M.Ubaydullayev, Sh.N.Narzullayev, and L.N.Nasimxonov, "Optimization of the sizes of the cylindrical measuring transducer", *Chemical Technology, Control and Management*: Vol. 2020: Iss. 5, Article 5, 2020. DOI: https://doi.org/10.34920/2020.5-6.29-32.

[15] E.Ulzhaev, Sh.N.Narzullaev, & O.N.Norboev, "Substantiation of application of artificial neural networks for creation of humidity measuring devices". *International Virtual Conference On Innovative Thoughts, Research Ideas and Inventions in Sciences*, no.1(1), New York, USA, January, 2021, pp.86-91.

[16] N.R.Yusupbekov, R.A.Aliev, R.R.Aliev, A.N.Yusupbekov, *Intellektual'ny'e sistemy' upravleniya i prinyatiya resheniy [Intelligent management and decision-making systems]*. Tashkent: Uzbek-skaya nacional'naya e`nciklopediya, 2014, 490 p. (in Russian).

[17] M. A. Nielsen, Neural networks and deep learning. Determination Press, 2015.

[18] V.G. Spitsyn, Yu.R. Tsoj, *Predstavlenie znanij v informatsionnykh sistemakh [Knowledge representation in information systems]*. Uchebnoe posobie, Tomsk: Izd-vo TPU, 2006, 146 p. (in Russian).

[19] P.P.Mal'tsev, K.M.Ponomarev, Yu.I.Stepanov, "«Umnaya pyl'» na osnove mikrosistemnoj tekhniki" ["Smart dust" based on microsystem technology], *«Intellektual'nye robototekhnicheskie sistemy – 2021»: materialy molodezh. nauch. shkoly*, Taranrog.: Izd-vo TRTU, 2001, pp. 220-232. (in Russian).

[20] A.V.Gavrilov, *Gibridnye intellektual'nye sistemy [Hybrid intelligent systems]*. Novosibirsk: Izd-vo NGTU, 2003, 164 p. (in Russian).

[21] Dias M.B., Stentz A. A Free Market Architecture for Distributed Control of a Multirobot System. Proceedings of the 6th International Conference on Intelligent Autonomous Systems (IAS), Venice, Italy, July, 2000, pp. 115 - 122.

[22] Stentz A., Dias M.B. A Free Market Architecture for Coordinating Multiple Robots tech. report CMU-RI-TR-99-42, Robitics Institute, Carnegie Mellon University, December, 1999.

[23] Aliyev R.A., Vahidov R.M., Aliyev R.R. Artificial Neural Networks: Theory and Practice. –Tabriz: Tabriz University Press, 1993, 193 p.

[24] Uljayev E., Ubaydullayev U.M., Tadjitdinov G.T., Narzullayev S. (2021) Development of Criteria for Synthesis of the Optimal Structure of Monitoring and Control Systems. In: Aliev R.A., Yusupbekov N.R., Kacprzyk J., Pedrycz W., Sadikoglu F.M. (eds) 11th World Conference "Intelligent System for Industrial Automation" (WCIS-2020). WCIS 2020. Advances in Intelligent Systems and Computing, vol 1323. Springer, Cham. https://doi.org/10.1007/978-3-030-68004-6_73E.

[25] A.R.Marakhimov, Kh.Z.Igamberdiev, A.N.Yusupbekov, I.Kh.Siddikov, *Nechetko-mnozhestvennye modeli i intellektual'noe upravlenie tekhnologicheskimi protsessami [Fuzzy-set models and intelligent process control]*. Tashkent: TashGTU, 2014, 240 p. (in Russian).

[26] Yu. A.Kravchenko, "Postroenie prognoznykh modelej dinamicheskikh sistem na osnove integratsii nejronnykh setej

i geneticheskikh algoritmov" [Building predictive models of dynamic systems based on the integration of neural networks and genetic algorithms], *Izvestiya Taganrog. gos. radiotekh. un-ta*, vol. 64, no. 9, pp. 103–104, 2006. (in Russian).

[27] Szandała T. (2021) Review and Comparison of Commonly Used Activation Functions for Deep Neural Networks. In: Bhoi A., Mallick P., Liu CM., Balas V. (eds) Bio-inspired Neurocomputing. Studies in Computational Intelligence, vol 903. Springer, Singapore. https://doi.org/10.1007/978-981-15-5495-7_11

[28] Chigozie Nwankpa, Winifred Ijomah, Anthony Gachagan, Stephen Marshall, Activation functions: Comparison of trends in practice and research for deep learning. Machine Learning (cs.LG); Computer Vision and Pattern Recognition (cs.CV), arXiv preprint arXiv:1811.03378, 2018.

[29] A. Hertz, Introduction to the theory of neural computation. CRC Press, 2018.

[30] Y. LeCun, L. Bottou, G.B. Orr, and K.R. Müller, "Efficient backprop," in Neural networks: Tricks of the trade, Berlin, Heidelberg: Springer Berlin Heidelberg, 1998, pp. 9–50.

[31] T.A.Gavrilova, V.F. Khoroshevskij, *Bazy znanij intellektual'nykh sistem [Knowledge bases of intelligent systems]*. Sankt-Peterburg: Piter, 2000, 382 p. (in Russian).

[32] E. Ulzhaev, Sh.N. Narzullaev, E.F. Khudoyberdiev, Sh.A. Sulaimonova, Algorithm of the process of creation of neural networks for measuring qualitative parameters of bulk materials. Proceeding of International Conference on Scientific Endeavors and opportunities, Telavi, Georgia, March, 2021, pp. 78-81.

[33] Robbins, H., Monro, S.: A stochastic approximation method. Ann. Math. Statist.22, pp. 400–407.

[34] Bengio, Y. Practical recommendations for gradient-based training of deep architectures. arXiv 2012, arXiv:1206.5533.

[35] Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. Nature 1986, 323, 533–536.

[36] Kelley, C.T. Iterative methods for linear and nonlinear equations. In Frontiers in Applied Mathematics; SIAM: Philadelphia, PA, USA, 1995.

[37] Kelley, C.T. Iterative Methods for Optimization. In Frontiers in Applied Mathematics; SIAM: Philadelphia, PA, USA, 1999.