International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Neural Network Model and Decision Support Methodology Based on Machine Learning for Time Series Forecasting

Odilbek Askaraliyev [1,a)], Jonibek Usmonov [1,b)], Malika Normatova [2,c)], Sherzod Sharipov[1,d)], Qosimjon Mamatkulov [1,e)], Shokirjon Madaminov [1,f)]

[1]Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, Tashkent 100084, Amir Temur avenue 108, Uzbekistan.

[2]Gulistan State University, Gulistan city 4 microdistrict, Uzbekistan.

[a)] oasqaraliyev77@gmail.com, orcid.org/0009-0003-7363-4485
[b)] jonibek.usmonov@gmail.com, orcid.org/0009-0002-5131-9893
[c)] normatova.2989@gmail.com, orcid.org/0009-0009-9246-5596
[d)] sharipovsherzodbek@gmail.com, orcid.org/0009-0005-7629-3430
[e)] qosimjon.mamatqulov97@gmail.com, orcid.org/0009-0005-7869-6313
[f)] mkshokirojn@gmail.com, orcid.org/0009-0006-3391-0459
Corresponding Author: oasqaraliyev77@gmail.com

Annotation.

This article considers the issue of building an optimized neural network model for time series forecasting. In this case, the training sample of size NxM using machine learning methods is based on the solution of the correlation-regression problem for training the training sample. The research is based on the analysis of tax authority data. In this case, a neural network-based model and a decision-making method are proposed based on the analysis of tax revenue types and states. In addition, an algorithm for intelligentizing decision-making is proposed.

Keywords: neural network, decision support methodology, tax authority data, time series, forecasting, correlation, regression.

INTRODUCTION

In automated information systems of control systems, the problem of predicting dependent and independent characteristics based on determining the regression between characteristics using machine learning methods is relevant. Problems in the field of science and technology require solutions based on optimal algorithms and models. Intellectualized forecasting is one of the most pressing issues today.

To solve this problem, generalized parametric estimates and assessments of experts in the selected subject area are used as training data. Expert assessments are formed based on the opinions of a certain number of experts in order to support subsequent decisions and have two types of notation: numerical and verbal. In the process of expert assessment, the

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

numerical, that is, type is determined in the range [0:1] or on a percentage scale. The corresponding verbal assessments of experts are justified by the proposed solutions to improve the state of the control object.

In this study, tax authorities, which are considered to be the economic and social sphere, were selected as the object of study, and based on relevant statistical data, a generalized parametric assessment of the tax base of legal entities, the tax base of individual entrepreneurs and large taxpayers was presented. In this case, indicators were determined by the field of activity. 3 calculation methods were used for the general assessment of the fields of activity: the first one was without taking into account weighting coefficients and two methods were used to determine the weighting coefficients of indicators. As a result, the results of a parametric generalized assessment of the tax base of legal entities, the tax base of individual entrepreneurs and large taxpayers using 3 different methods of determining the weight of indicators were obtained. The results of the generalized assessment of these areas of activity are presented in Appendix 3.

MAIN PART

The appropriate solution proposed by the expert for a new object is considered, and models and algorithms based on machine learning are considered to solve the problem of predicting tax revenues. Thus, based on machine learning algorithms, adaptive functions of the dependence of expert decisions on generalized parametric estimates of objects are determined. Based on this function, the appropriate expert decision (recommendation) for a new object is determined based on its generalized parametric estimates.

The chosen formulation of the machine learning problem is expressed as follows. Given a set of objects X consisting of bodies from the statistical data of regional tax authorities, and for which Y expert assessments, as well as a training sample, are applied as follows:

$$X = \{X_1, X_2, \dots, X_j, \dots, X_n\}, \quad (1)$$

$$Y = \{Y_1, Y_2, \dots, Y_j, \dots, Y_n\}, \quad (2)$$

Also, many well-known expert assessments of the answers:

$$Y_i = \{Y'(X_i), i \in [1, .., N]\} \quad (3)$$

Here $Y'(X_i)$ -is a function that represents the expert's estimate for a new object that is passed to it, a previously unknown relationship. Therefore, we can assume that $Y'(X_i)$ is the value that needs to be found. The statistics and recommended expert estimates given for the objects refer to the training sample. Each object X_j is able to take values from a certain set:

$$X_j = \{x_{j1}, x_{j2}, \dots, x_{jh}, \dots, x_m, y_j\}$$

For the input data of the training sample, you need to build the following type of model or algorithm:

$$f: X \to Y. \quad (4)$$

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Here f $-$ is the calculated rule (decision function) - the approximation of - $y'(x)$ to the entire set X. This is the result of inductive reasoning, which allows us to obtain some estimates of expert assessments for them.

As a rule, to solve the problem of constructing the function $f_S: X \to Y$, a certain learning model is selected from the training sample S, which includes two components:

First, a certain function is defined

$$f: X^*\beta \to Y, \qquad (5)$$

here is a set of so-called β $-$coefficients. The required function f_S has the following form:

$$f_S\{X\} = f(X, \beta), \qquad (6)$$

Function (6) is also called a prediction model.

Secondly, an algorithm is developed to determine the values of the function coefficients, that is, the value of β is calculated, for which the function f_S, defined by equality (3), must satisfy the given optimality conditions.

Next, models and algorithms (predictions) of supervised machine learning are considered. This is done on the basis of a machine learning model using multiple linear regression.

The function with multiple regression corresponding to formula 5 looks like this:

$$Y' = f(x, \beta) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_j x_j + \cdots + \beta_n x_n, \quad (7)$$

where β_j sets are regression coefficients; x_j sets are regressors (factors); n $-$ is the number of factors.

The parameter β_j, which does not include a factor in the formula, is often called a constant. Formally, this is the value of the function for which all factors are zero. For convenience, we can consider a constant as a parameter with a coefficient equal to $x_0 = 1$.

To formulate a regression equation, it is necessary to determine the values of its coefficients. The sum of squared errors is used to estimate the regression coefficients. The method of least squares allows us to obtain such estimates of the coefficients in which the sum of the squared deviations of the target (actual) values of the characteristic Y from the theoretical Y' is minimal, i.e.:

$$E = \sum (Y' - Y)^2 \to min. \qquad (8)$$

In the problem posed, the number of independent variables is three, so according to formula (7), we obtain the following expression:

$$Y_i' = f(x, \beta) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3}, \qquad (9)$$

Linear regression uses the mean square error. In the given problem, there is a sample containing 3 features. To find the values of the coefficients (or weights) of the function, the following optimization problem is solved:

Received: August 14, 2025

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

$$E(\beta) = \sum_{i=1}^{n}(y'_i - y_i)^2 = \sum_{i=1}^{n}(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} - y_i)^2 \rightarrow min, \qquad (10)$$

where x_{i1}, x_{i2}, x_{i3} - are the values of the attributes of the i −th object, y_i − is the value of the target attribute of the i − th object (or the numerical value of the given assessment or the solution proposed by the expert). i −th object, n − is the number of objects in the training set (dataset).

The coefficients β_0, β_1, β_2 and β_3 in formulas (9) and (10) can be found numerically using a combination of iterative gradient and stochastic gradient descent algorithms. At each step, the vector of coefficients changes in the direction of decrease, that is, it is the antigradient of the objective function. In this case, the gradient step (learning step) of changing the values of the coefficients will be:

$$\beta_i = \beta_i - \eta \nabla E(\beta) = \beta_i - \eta \frac{1}{n} \sum_{i=0}^{n} \nabla E_i\,(\beta). \qquad (11)$$

In a single object, formula (11) takes the following form:

$$\beta_{ij} = \beta_{ij} - 2\eta x_{ij}\left(\sum_{j=0}^{k}\left(\beta_{ij} x_{ij}\right) - y_i\right), \qquad (12)$$

where j − is the number of indicators of the object.

In the gradient descent algorithm, the coefficient values are changed based on the mean square error value. In contrast, we change the parameter values based on the squared error for each set of training examples, as in the stochastic gradient descent algorithm. Then, the expanded expression for calculating the step change of each coefficient corresponding to formulas (11) and (12) takes the following form:

$$\begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\\ \beta_3\end{bmatrix} = \begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\\ \beta_3\end{bmatrix} - \eta \begin{bmatrix} \frac{\partial}{\partial \beta_0}(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} - y_i)^2 \\ \frac{\partial}{\partial \beta_1}(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} - y_i)^2 \\ \frac{\partial}{\partial \beta_2}(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} - y_i)^2 \\ \frac{\partial}{\partial \beta_3}(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} - y_i)^2 \end{bmatrix} == \begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\\ \beta_3\end{bmatrix} -$$

$$2\eta \begin{bmatrix} \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} - y_i \\ x_{i1}(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} - y_i) \\ x_{i2}(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} - y_i) \\ x_{i3}(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} - y_i) \end{bmatrix}. \qquad (13)$$

The conditions for stopping the change in the value of the coefficients are expressed by the following formulas:

$$\left\|\beta^{[l]} - \beta^{[l-1]}\right\| \leq \varepsilon; \qquad (14)$$

$$\left|\frac{1}{n}\sum |y' - y|^{2[l]}\right| \leq \varepsilon. \qquad (15)$$

Thus, the conditions of the process of finding the minimum are checked after each period, that is, after each complete iteration.

Received: August 14, 2025

International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Based on these formulas, the step-by-step view of the algorithm developed to solve the problem of learning generalized estimates and expert estimates in numerical form by linear regression is as follows:

Step 1. The generalized estimates of taxpayers in the information category (training source) according to the tax base and the numerical values of the solutions proposed by experts for each object are loaded into the arrays $X[i,j]$ and $Y[i]$, that is, independent and dependent factor matrices corresponding to formulas (1) and (2) are constructed.

Step 2. The process of selecting the prediction function, setting the initial values of the coefficients β_1, the learning step η, the calculation accuracy ε_1 and ε_2, and the number of epochs (Epoches).

Step 3: l epoch, repeated:

1. iteration t: is calculated for the next indicator sample.

The output value of the function y' and the error E according to formulas (7) and (10): $y' = f(x_i, \beta), E_i = f(x_i, \beta) - y_i$..

2. Based on the value of the tax burden of the function for the next indicator sample, the new values of the coefficients β are calculated according to formula (13): $\beta = \beta - 2\eta x_i(\beta x_i - y_i)$.

Step 4. Using formulas (13) and (14), the stopping condition of the algorithm is checked: if $\left\|\beta^{[l]} - \beta^{[l-1]}\right\| \leq \varepsilon_1$ and $\left|\frac{1}{n}\sum|y' - y|^{2[l]}\right| \leq \varepsilon_2$ or $= Epochs$ then proceed to the next step, otherwise $l = l + 1$ and proceed to step 3.

Step 5. Derivation of β coefficients Based on β coefficients and generalized estimates, y' is calculated for a new object and the management decision of the corresponding object is considered from among the objects given as a decision. In this case, the value of y differs minimally from the value of y' and the algorithm ends.

The parameter values for each method of determining the weighting coefficients of indicators when calculating generalized estimates (training sources) of the studied tax authorities' areas of activity have the following form (Tables 1-3):

$$f_1(x,\beta) = 0{,}564 + 0{,}409x_1 + 0{,}213x_2 - 0{,}315x_3. \quad (16)$$

$$f_2(x,\beta) = 0{,}684 + 0{,}085x_1 + 0{,}108x_2 - 0{,}407x_3. \quad (17)$$

$$f_3(x,\beta) = 0{,}623 + 0{,}239x_1 + 0{,}137x_2 - 0{,}44x_3. \quad (18)$$

Table 1.

Results of training with training resources calculated without taking into account the weighting coefficients of statistical indicators.

№	Implementation	Coefficient

Received: August 14, 2025

International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

1	Assessment of the tax base of taxpayer legal entities x_1	$\beta_1 =$	0,409
2	Assessment of the tax base of individual entrepreneurs x_2	$\beta_2 =$	0,213
3	Assessment of the tax base of large taxpayers x_3	$\beta_3 =$	0,315
4	Unexpected error	$\beta_0 =$	0,564

Table 2.

Results of training with a training complex in which the weight of indicators in the calculation is determined by the method of assigning points

№	**Implementation**	**Coefficient**	
1	Assessment of the tax base of taxpayer legal entities x_1	$\beta_1 =$	0,085
2	Assessment of the tax base of individual entrepreneurs x_2	$\beta_2 =$	-0,108
3	Assessment of the tax base of large taxpayers x_3	$\beta_3 =$	-0,407
4	Unexpected error	$\beta_0 =$	0,684

Table 3.

The results of training with the training complex, the weight of indicators in their calculation, are determined by the rating method.

№	**Implementation**	**Coefficient**	
1	Assessment of the tax base of taxpayer legal entities x_1	$\beta_1 =$	0,239
2	Assessment of the tax base of individual entrepreneurs x_2	$\beta_2 =$	-0,237
3	Assessment of the tax base of large taxpayers x_3	$\beta_3 =$	-0,440
4	Unexpected error	$\beta_0 =$	0,623

A comparative analysis in the form of a graph of the values of the trained linear functions and the actual values of the expert estimates based on the test results is presented at the end of the article in the results section.

The next predictive function under consideration is a multivariate second-order polynomial function. According to formula (5) with multivariate polynomial regression, the function for three independent variables looks like this:

$$Y_i' = f(X, \beta) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i1}^2 + \beta_5 x_{i3}^2 + \beta_6 x_{i3}^2. \quad (19)$$

As with many function training, constructing a regression equation involves estimating its coefficients using a least-squares error minimization algorithm:

International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

$$S = \sum(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i1}^2 + \beta_5 x_{i3}^2 + \beta_6 x_{i3}^2 - y_i) \to min. \qquad (20)$$

Then, the expression for calculating the step change for each of the coefficients β_0, β_1, β_2, β_3, β_4, β_5 and β_6 corresponding to formulas (11) and (12) will look like this:

$$\begin{bmatrix}\beta_0\\\beta_1\\\beta_2\\\beta_3\\\beta_4\\\beta_5\\\beta_6\end{bmatrix} = \begin{bmatrix}\beta_0\\\beta_1\\\beta_2\\\beta_3\\\beta_4\\\beta_5\\\beta_6\end{bmatrix} - \eta \begin{bmatrix}\frac{\partial}{\partial\beta_0}(y_i' - y_i)^2\\ \frac{\partial}{\partial\beta_1}(y_i' - y_i)^2\\ \frac{\partial}{\partial\beta_2}(y_i' - y_i)^2\\ \frac{\partial}{\partial\beta_3}(y_i' - y_i)^2\\ \frac{\partial}{\partial\beta_4}(y_i' - y_i)^2\\ \frac{\partial}{\partial\beta_5}(y_i' - y_i)^2\\ \frac{\partial}{\partial\beta_6}(y_i' - y_i)^2\end{bmatrix} == \begin{bmatrix}\beta_0\\\beta_1\\\beta_2\\\beta_3\\\beta_4\\\beta_5\\\beta_6\end{bmatrix} -$$

$$2\eta \begin{bmatrix}(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i1}^2 + \beta_5 x_{i3}^2 + \beta_6 x_{i3}^2 - y_i)\\ x_{i1}(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i1}^2 + \beta_5 x_{i3}^2 + \beta_6 x_{i3}^2 - y_i)\\ x_{i2}(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i1}^2 + \beta_5 x_{i3}^2 + \beta_6 x_{i3}^2 - y_i)\\ x_{i3}(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i1}^2 + \beta_5 x_{i3}^2 + \beta_6 x_{i3}^2 - y_i)\\ x_{i1}^2(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i1}^2 + \beta_5 x_{i3}^2 + \beta_6 x_{i3}^2 - y_i)\\ x_{i2}^2(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i1}^2 + \beta_5 x_{i3}^2 + \beta_6 x_{i3}^2 - y_i)\\ x_{i3}^2(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i1}^2 + \beta_5 x_{i3}^2 + \beta_6 x_{i3}^2 - y_i)\end{bmatrix} \qquad (21)$$

The coefficient values are determined using the above algorithm. The values of the function coefficients for each set of generalized assessments of the areas of activity of regional tax authorities for monitoring and forecasting tax revenues are presented in Table 4.

$$f_4(x,\beta) = 0{,}593 + 0{,}257x_1 + 0{,}064x_2 - 0{,}215x_3 + 0{,}147x_{i1}^2 + 0{,}094x_{i2}^2 + 0{,}136x_{i3}^2.$$

$$f_5(x,\beta) = 0{,}68 + 0{,}08x_1 + 0{,}086x_2 - 0{,}338x_3 + 0{,}002x_{i1}^2 + 0{,}013x_{i2}^2 + 0{,}154x_{i3}^2.$$

$$f_5(x,\beta) = 0{,}618 + 0{,}222x_1 + 0{,}117x_2 - 0{,}378x_3 + 0{,}07x_{i1}^2 + 0{,}016x_{i2}^2 + 0{,}149x_{i3}^2.$$

Table 4.Results of training with the training complex calculated without taking into account the weighting coefficients of the indicators.

№	Implementation	Coefficient	
1	Assessment of the tax base of taxpayer legal entities x_1	$\beta_1 =$	0,257
2	Assessment of the tax base of individual entrepreneurs x_2	$\beta_2 =$	0,064
3	Assessment of the tax base of large taxpayers x_3	$\beta_3 =$	-0,215
4	$x_{i4} = x_{i1}^2$	$\beta_4 =$	0,147
5	$x_{i5} = x_{i2}^2$	$\beta_5 =$	0,094

Received: August 14, 2025

International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

6	$x_{i6} = x_{i3}^2$	$\beta_6 =$	-0,136
7	Unexpected error	$\beta_0 =$	0,593

Table 5.The results of the study on the training sample, the values of the weighting coefficients of the indicators in its calculation are determined by the scoring method.

№	Implementation	Coefficient	
1	Assessment of the tax base of taxpayer legal entities x_1	$\beta_1 =$	0,08
2	Assessment of the tax base of individual entrepreneurs x_2	$\beta_2 =$	-0,086
3	Assessment of the tax base of large taxpayers x_3	$\beta_3 =$	-0,338
4	$x_{i4} = x_{i1}^2$	$\beta_4 =$	0,002
5	$x_{i5} = x_{i2}^2$	$\beta_5 =$	-0,013
6	$x_{i6} = x_{i3}^2$	$\beta_6 =$	-0,154
7	Unexpected error	$\beta_0 =$	0,679

Table 6.The results of training with the training complex, the weighting coefficients of the indicators in their calculation are determined by the rating method.

№	Implementation	Coefficient	
1	Assessment of the tax base of taxpayer legal entities x_1	$\beta_1 =$	0,222
2	Assessment of the tax base of individual entrepreneurs x_2	$\beta_2 =$	-0,117
3	Assessment of the tax base of large taxpayers x_3	$\beta_3 =$	-0,378
4	$x_{i4} = x_{i1}^2$	$\beta_4 =$	0,070
5	$x_{i5} = x_{i2}^2$	$\beta_5 =$	-0,149
6	$x_{i6} = x_{i3}^2$	$\beta_6 =$	-0,618
7	Unexpected error	$\beta_0 =$	0,679

Decision support model using neural networks in tax revenue forecasting

The problem of classification and prediction is considered using a multilayer feedforward neural network. Using the classification and prediction model, a set of rules is formed by which any new object can be classified into one of the classes. The rules are determined based on previously obtained information about the objects.

In the classification and forecasting problem, it is necessary to determine the value of the dependent variable based on the values of other independent variables characterizing the object of study. The forecasting problem can be described as follows. A large number of statistical objects (data) are given:

$$X = \{X_1, X_2, \dots, X_j, \dots, X_n\},$$

where X_j is the object under study. In this research work, the objects themselves are considered to be the tax base in the tax authorities.

Each object is characterized by a set of variables, that is, each territorial tax authority is characterized by generalized parametric estimates obtained from three.

Areas of activity (tax base of legal entities, tax base of individual entrepreneurs and the status of the base of large taxpayers):

$$X_j = \{x_{j1}, Xx_{j2}, \dots, x_{jh}, \dots, x_m, y_j\};$$

Here x_{jh} − these are independent variables whose values are known and on the basis of which the value of the dependent variable y is determined. In this case, the independent variables are generalized parametric estimates for the field of activity. Here y_j, j-is the dependent variable, which contains the numerical values of expert assessments.

As is known, machine learning can be unsupervised or supervised. In the first case, the teacher evaluates the behavior of the system and monitors its subsequent changes. Training a multilayer perceptron neural network is considered supervised training.

To train a neural network, a perceptron neural network with three inputs (for inputting generalized assessments for three fields of activity) and one output (expert decisions) was created (Fig.1). The neural network has two hidden layers of neurons.

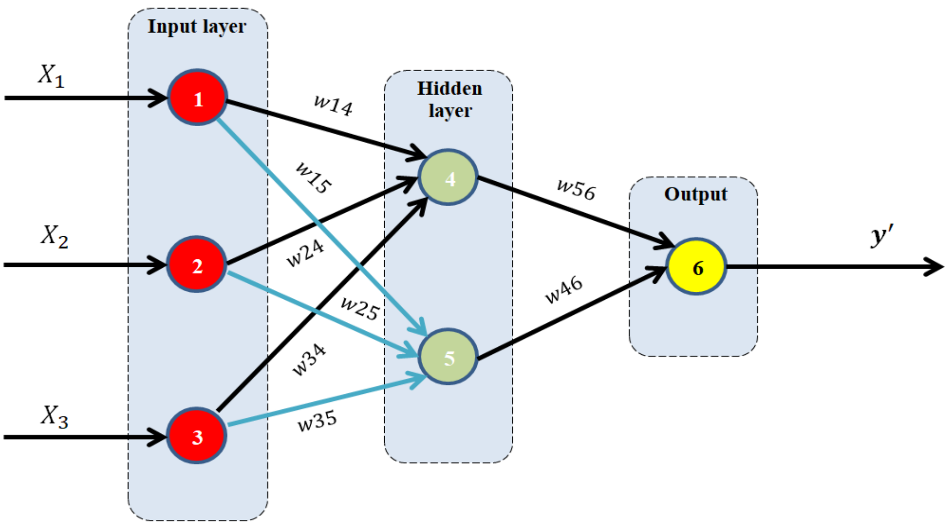


Fig.1. Overview of a perceptron neural network model with three inputs.

The number of neurons in the hidden layers of homogeneous multilayer neural networks with a sigmoidal activation function can be calculated based on an experimental formula to estimate the required number of synaptic weights N_w. That is, the required number of

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

neurons in the hidden layers of a perceptron can be determined by the formula, which is a consequence of the *Arnold-Kolmogorov-HextNielsen* theorems:

$$\frac{N_y N_p}{1+log_2 N_y} \leq N_w \cdot \left(\frac{N_p}{N_x}+1\right) \cdot \left(N_x + N_y = 1\right) + N_y, \quad (22)$$

where $N_y -$ the size of the output signal; $N_x -$the size of the input signal; $N_p -$the number of elements of the training set.

For each neuron, $x_1, \dots, x_n -$ are input signals from other neurons or inputs, $w_1, \dots, w_n -$are the synaptic weights of the inputs. The current state of the neuron is defined as the sum of the multiplied input signals and the corresponding connection weights:

$$S_j = net_j = \sum_{i=1}^{n} x_{ij} w_{ij} - b.$$

Here n $-$is the number of inputs to the j $-$th neuron; x_{ij} $-$is the i $-$input value of the $j-$ neuron$w_{ij} -$ is the weight of the i $-$synapse input to the j $-$neuron; b $-$is the threshold of the neuron. The threshold of a neuron can be considered as the equivalent synaptic weight corresponding to the estimated input x_0. If we tentatively choose $w_0 = b$, $x_0 = -1$, then we can consider a neuron with a threshold of 0 and one additional input:

$$S_j = net_j = \sum_{i=1}^{n} x_{ij} w_{ij} - b = \sum_{i=1}^{n} x_{ij} w_{ij} - x_{0j} w_{0j} = \sum_{i=1}^{n} x_{ij} w_{ij}\,. \quad (23)$$

The task of an artificial neural network is to transform the input training data into an output vector. The transformation is determined by the weights of the neural network. The training process consists of adjusting the weights of the neurons. The goal of training is to find the state of the weights that minimizes the network output error on the training and test samples.

Then, the state of the neurons is changed using the activation function f to determine the output value. There are different types of neural activation functions: threshold, sign, sigmoid, semilinear, linear, radial basis, semilinear with saturation, linear with saturation, hyperbolic tangent, and triangular.

All neurons in our neural network use the most common activation function - sigmoid:

$$f(x) = \frac{1}{1+e^{-\alpha x}}. \quad (24)$$

One of the important properties of the sigmoid function is the simple expression for its derivative:

$$f'(x) = \propto \cdot\ f(x)\big(1 - f(x)\big). \quad (25)$$

By substituting the input signal values (3.n) for the argument x in formula (3.n), we obtain expression (3.n) of the following form:

$$f(S_j) = \frac{1}{1+e^{-S}} = \frac{1}{1+e^{-\sum_{i=1}^{n} x_{ij} w_{ij}}} \quad (26)$$

Received: August 14, 2025

Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

When one of the training samples x_k is given as input to a neural network, we get a certain value y'_k at the output. Then the error of this iteration can be considered:

$$E = \|y_k - y'_k\|. \qquad (27)$$

The main goal of training a neural network is to minimize the error function. The error function is determined by the method of least squares:

$$E = \frac{1}{2}\sum_{i=1}^{n}(y_i - y'_i)^2 \rightarrow min. \qquad (28)$$

where y'_i −the value of the output signal when the i −th set of training samples is applied to the inputs of the network neuron; y_i −the required (specified) value of the output neuron for the i −th set of training samples for training.

The process of training a neural network with sigmoidal activation functions is considered. There are many multilayer training algorithms, the most common of which is the back-propagation algorithm. This algorithm is used to minimize the deviation of the actual values of the neural network output signals from the desired one. This rule of machine learning is called the delta rule and can be used to train both linear perceptrons and perceptrons with a threshold activation function.

There are two main approaches to back-propagating the error of a perceptron neural network: sequential (online learning) and group (ensemble learning). We use sequential training, which is suitable for the requirements of the research problem, in which the connection weights are changed after each training source is assigned to the neural network inputs. In this case, the squared error of the neural network is used for one k −th input training example:

$$E = \frac{1}{2}(y_i - y'_i)^2. \qquad (29)$$

When the error rolls back, the weights are adjusted according to the expression

$$\Delta w_{ij} = \eta\delta_j x_{ij}; \qquad (30)$$

where η −is the learning rate coefficient specified by the researcher $0 < \eta < 1$,, x_{ij} − is the input value of the j −th neuron. $\delta_j = (y_i - y'_i)$ - errors at the output of hidden neurons are not directly related to the output error. The weights in hidden neurons should vary according to (3.n) depending on their contribution to the error value of the next layer.

Thus, the error value of a neuron is determined by its position in the network. For output layer neurons:

$$\delta_j = \left(wf_{j,k}(S)\right)'\left(f_j(S) - y_j\right) = y'_{j,k}\left(1 - y'_{j,k}\right)\left(y_{j,k} - y'_{j,k}\right), \qquad (31)$$

where $y_{j,k}$ - is given, i.e. a predetermined value; $y'_{j,k}$– the calculated value of the output signal of the k −neuron for the j −th data set from the training source. Since our neural network (see Fig. 1) has only one output, the index j can be omitted in formula (32).

If the neuron belongs to the hidden layer, then

International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

$$\delta_j = \left(f_j(S)\right)' \sum_k w_{j,k}\delta_k = y_j'\left(1 - y_j'\right)\sum_k w_{j,k}\delta_k, \qquad (32)$$

where, δ_j − is the error of the j −th neuron in the hidden layer, δ_k is the error of the k − output neuron; $w_{j,k}$ – is the weight of the connections connecting the output neurons; $\left(f_j(S)\right)'$ – is the value of the derivative of the activation function of the jth neuron.

If the value of the change in the weight of the i −connection is known, then the value of the new weight can be calculated using the formula.

$$w_{j,k}(k+1) = w_{j,k}(k) + \Delta w_{ij.} \qquad (33)$$

Here k − is the number of training iterations; i − is the input number; j − is the number of neurons.

Decision support algorithm using neural networks for time series forecasting

Training a multilayer feedforward neural network using the backpropagation method and the sequential learning method (online training) consists of 3 main steps:

1) direct distribution of the input training sample;

2) error detection and backpropagation of this error;

3) correction of weight values according to the error.

Let us consider in more detail the stages of the training algorithm for a multilayer neural network, which consists of the following steps:

Step 1. Set the learning rate η $(0 < \eta < 1)$, the minimum total root-mean-square error of the network E_m, and the period values - the number of runs of the training sets.

Step 2. The weight coefficients w_1, w_2,..., w_m of the neural network are chosen randomly, for example, all weights can be set to 0.

Step 3. The next indicator sample is fed to the inputs of the neural network and the states of the neural network are calculated, that is, the values of the output neuron are calculated.

Step 4. According to the formulas (31) and (32), the error δ_j and $\Delta w_i(b)$ and the values for adjusting the weight of the output neuron are calculated, respectively.

Step 5. The weight coefficients of the output neurons of the neural network are changed according to the formula (34).

Step 6. According to the formulas (31) and (32), the error δ_j, the weight adjustment values Δw_i and the hidden layer weight adjustment values are calculated, respectively.

Step 7. The weight coefficients of the neurons in the hidden layer of the neural network are changed according to the formula (34).

Step 8. If the iteration is completed, then the next step is performed, otherwise the transition to step 3 is performed.

International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

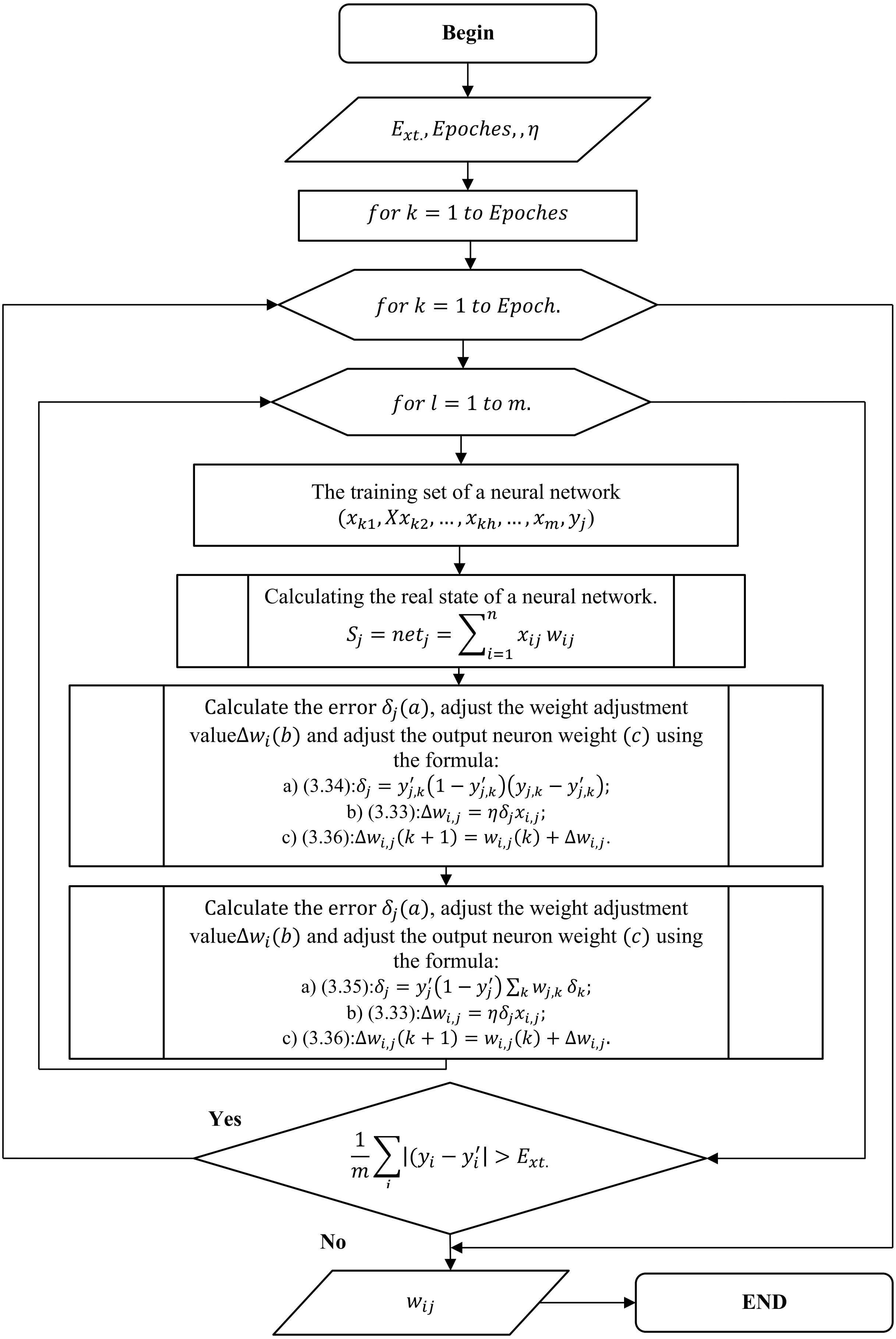
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Step 9. If the total root mean square error of the network is less than the specified $E \leq E_m$ or the end of the period, it goes to the next step, otherwise it goes to step 3.

Step 10. Output the values of the weight coefficients w_1, w_2,..., w_m of the neural network, calculate and output the value of y' for the new object, and as the resulting decision, the given objects for the corresponding object are proposed, in which the value of y differs minimally from the value of y', and the algorithm ends. The block diagram of the training algorithm for a multilayer feedforward neural network using the backpropagation method and the sequential learning approach is shown in Fig. 2.

Received: August 14, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)



Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Fig. 2. Block diagram of the multilayer neural network training algorithm.

Then, a neural network built on the basis of the proposed algorithm is trained on three training samples (general calculations, in which three methods of determining the weighting coefficients of indicators of the spheres of activity are used separately). As a result of training the neural network, the weighting coefficients of the neurons are determined. The weighting coefficients of the neural network when trained with training sources calculated without taking into account the weighting coefficients of indicators of the sphere of activity are presented in Fig. 3.

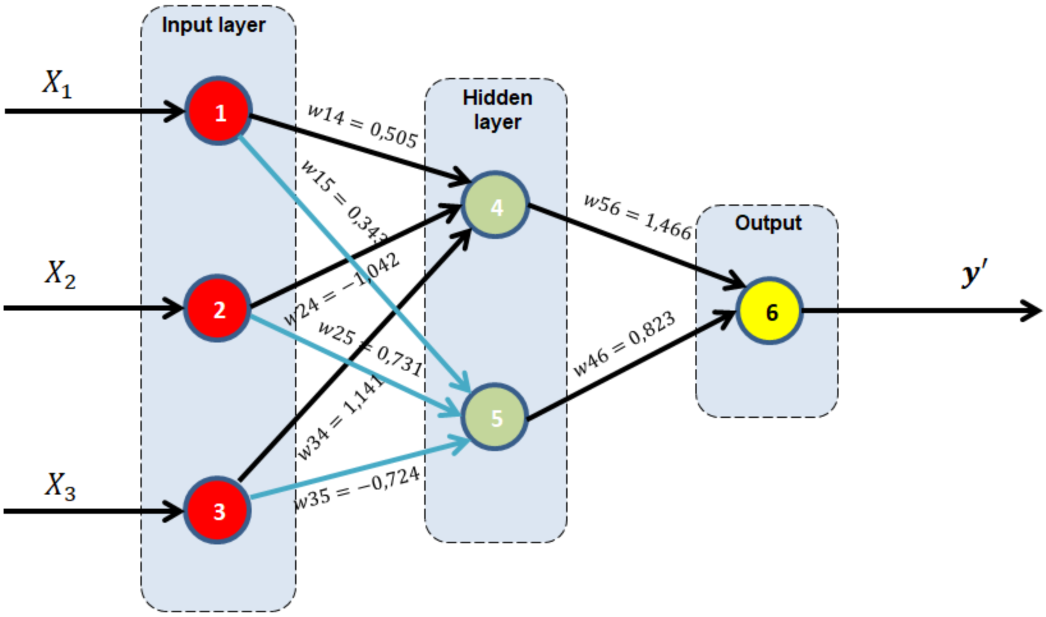


Fig. 3. Neural network weighting coefficients in an experiment with a training set calculated without taking into account the weighting coefficients of the activity indicators.

When trained with a sample calculated without taking into account the weighting coefficients of the indicators of the field of activity, the function of the neural network has the following form:

S_4 =0,505x_1 −1,042x_2 −1,141x_3 +0,28;

$$S_4' = \frac{1}{1 - e^{-0,505x_1 - 1,042x_2 - 1,141x_3 + 0,28}};$$

$$S_5 = 0,343x_1 - 1,731x_2 - 0,724x_3 + 0,135;$$

$$S_5' = \frac{1}{1 - e^{-0,343x_1 - 1,731x_2 - 0,724x_3 + 0,135}};$$

$$S_6 = 1,466S_4' - 0,823S_5' + 1,047;$$

$$y' = S_6' = \frac{1}{1 + e^{-1,466S_4' - 0,823S_5' + 1,047}}.$$

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The weighting coefficients of the neural network when learning with a training set calculated using the scoring method when determining the weighting coefficients of the indicators of the field of activity are shown in Fig. 4.

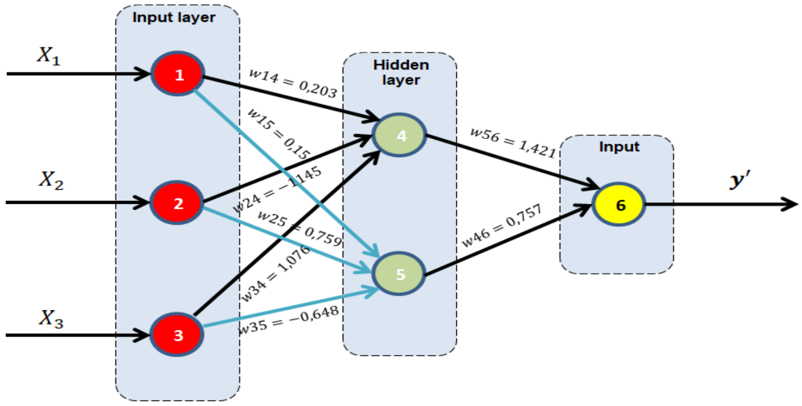


Fig. 4. Neural network weighting coefficients in an experiment with a training set calculated using the scoring method for determining the weighting coefficients of indicators.

In the process of learning with a training set, where the rating score method is used to determine the weighting coefficients of indicators in the calculation, the function of the neural network has the following form:

S_4 =0,203x_1 −1,145x_2 −1,076x_3 +0,273;

$$S_4' = \frac{1}{1 - e^{-0,203x_1 - 1,145x_2 - 1,076x_3 + 0,273}};$$

$$S_5 = 0,15x_1 - 1,759x_2 - 0,648x_3 + 0,116;$$

$$S_5' = \frac{1}{1 - e^{-0,15x_1 - 1,759x_2 - 0,648x_3 + 0,116}};$$

$$S_6 = 1,421S_4' - 0,757S_5' + 1,971;$$

$$y' = S_6' = \frac{1}{1 + e^{-1,421S_4' - 0,757S_5' + 1,971}}.$$

The weighting coefficients of the neural network when learning with a training set calculated by the rating method when determining the weighting coefficients of the indicators of the field of activity are shown in Fig. 5.

Received: August 14, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

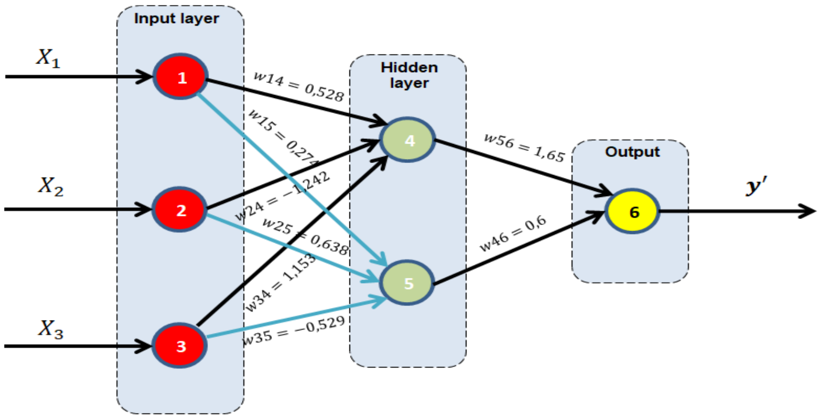


Fig.5. Weighting coefficients of a neural network when training with a training set calculated using the rating score method when determining the weighting coefficients of indicators of the field of activity.

The graph of the actual values of the expert assessments based on the values of the weighted neural network and the test sample is presented in the results of the study.

RESULS

The system is automatically coordinated during the process of taxpayers' registration for “value added tax”. To form information for the organization of value added tax control and cameral control, it is enough to integrate the information of the information system integrated between management systems. Fig. 7 shows the algorithm of this process. To analyze the results of the application of the algorithms developed during the study, generalized assessments for the areas of activity of the research object were calculated based on three types of training samples, namely: without taking into account weighting coefficients, using indicators of the area of activity, using the ranking method and using the scoring method.

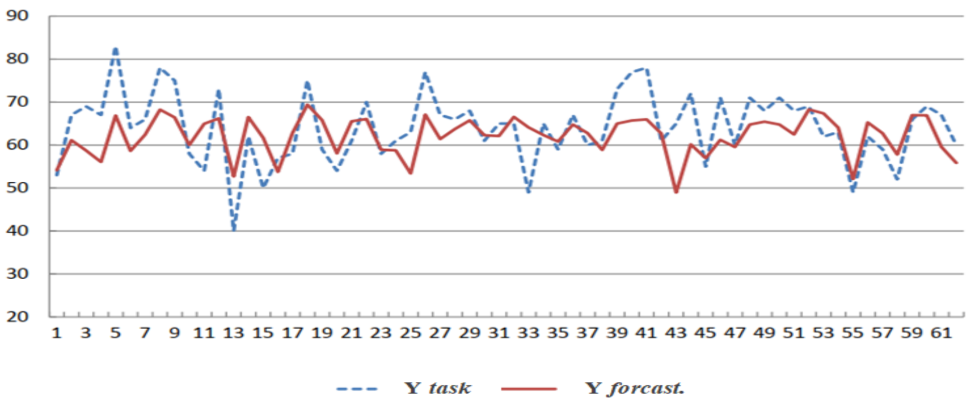


Fig. 6. Comparison graph of predictions based on the multilinear regression function (16) based on the test sample calculated without taking into account the weighting coefficients of the indicators.

Figs. 9-10 show graphs constructed based on the test results of the forecasting function using formulas (16) - (18) and (22) - (24), respectively.

Received: August 14, 2025

International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

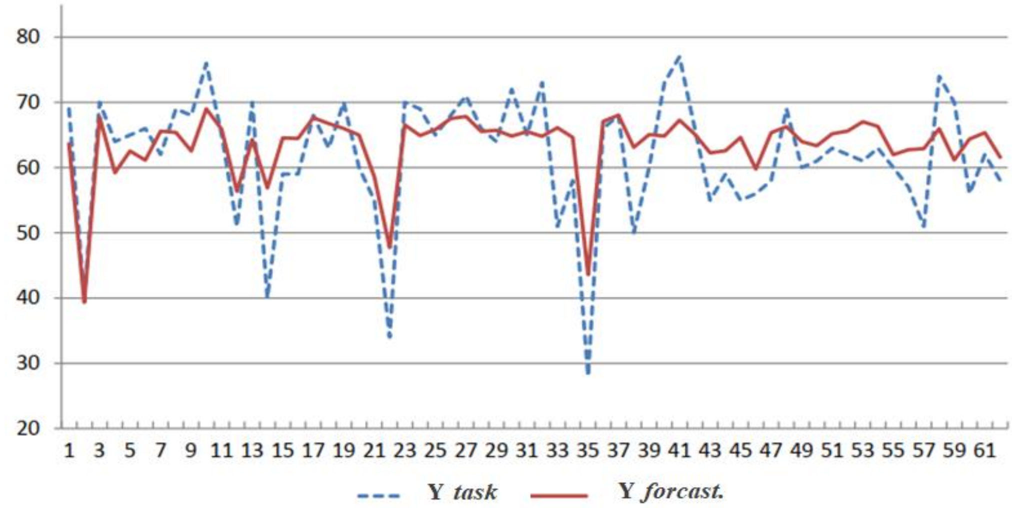


Fig. 7. Comparison graph of the prediction function (17) based on the test result of multiple linear regression.

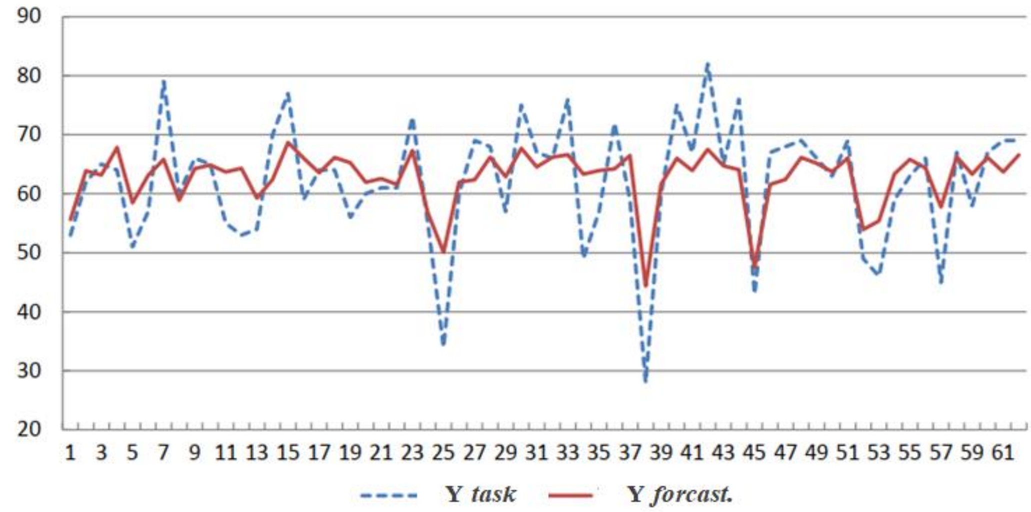


Fig. 8. Comparison graph of the prediction function (24) based on the test sample polynomial regression.

Figs. 8–9 show graphs of the neural network performance for the same input test sample used to test the regression functions. The first test sample (Appendix 3) was calculated without taking into account the weighting coefficients of the indicators of the areas of activity performed by experts.

The second test sample, calculated taking into account the weighting coefficients of the indicators of the activity areas using the ranking method, is fed to the inputs of the neuron. In this case, the network training stage is $\eta = 0.001$ epochs, $Epochs = 500$, the initial values of the weighting coefficients of all neurons are equal to 0.

The third test sample (Appendix 3) is calculated taking into account the weighting coefficients of the indicators of the activity areas using the scoring method. We feed this data

Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

to the neuron inputs. In this case, the network training stage is $\eta = 0.001$, $Epochs = 500$, the initial values of the weighting coefficients of all neurons are equal to 0. A graph comparing the results of the neural network performance with the given one is shown in Figure 4.20.

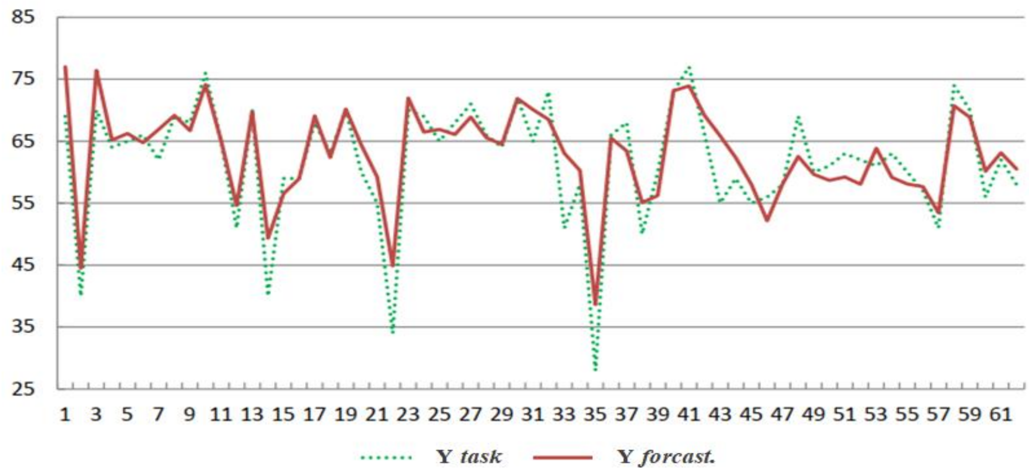


Fig. 10. Graph comparing the performance of a neural network with the given data.

In general, based on the developed methods and algorithms, a generalized software component was developed for parametric assessment of the state of the organization and forecasting of decisions to be made using machine learning of the indicators of tax authorities.

CONCLUSION

It was found that the most convenient way to characterize automated management is to include in the arithmetic values of the characteristics characterizing digital tax authorities (indicators of the standard of living and quality of life of the population). Analysis of average indicators allows us to assess the nature of tax authorities included in a certain cluster. As a result of numerical solutions of the developed mathematical models, graphs were constructed comparing the results of forecasting using a neural network built on the basis of a test sample.

A neural network method was developed for the intelligent forecasting process. This allowed us to reduce the amount of time spent on the forecasting process in the applied field by up to 23%. In addition, an intelligent methodology for decision support was created.

The effectiveness of using machine learning methods in forecasting based on automated information systems of management systems has been confirmed. In forecasting based mainly on time series, the problem of predicting dependent and independent characteristics based on determining the regression between characteristics is relevant. Expert assessments are formed based on the opinions of a certain number of specialists in order to support subsequent decisions. In this study, tax authorities, which are considered an economic and social sphere, were selected as the object of study and experiments were conducted.

Acknowledgments

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The authors would like to express their sincere gratitude to everyone who provided support in carrying out this study. We are grateful to our colleagues for their valuable advice and suggestions at various stages of our work. Special thanks to the communications organization that provided data for this work.

REFERENCES

[1]. A. Saidov, A. Dustmukhammedov, F. Khakimova. Problematic issues in the organization of customs control related to the application of artificial intelligence methods. *International scientific and practical conference on "modern problems of applied mathematics and information technology (*MPAMIT2022*)". AIP Conf. Proc.* 3004, 050003 (2024). https://doi.org/10.1063/12.0025282

[2]. O.U.Askaraliyev, G.X.Muxtarova, N.T Malikova., R.H. Yuldashev. Practice of applying models of clustering of incoming semi-structured data in management systems. *V International Scientific Conference "Construction Mechanics, Hydraulics and Water Resources Engineering". (CONMECHYDRO - 2023),* https://doi.org/10.1051/e3sconf/202340105063

[3]. O.U.Askaraliev, N.T.Malikova, N.A. Naim, Sh.Sharma, Zh.Chohan, R.Kumar, S.Singh. Intelligent Decision Support Systems in Integrated Management Systems – A Novel Multicriteria Computational Control System. Advancements in Civil Engineering & Technology, *AIP Conference Proceedings 2558, 020001 (2023). 03 February 2023.* https://doi.org/10.1063/5.0121301

[4]. I.Siddikov, A.Amdumalikov, O.Turakulov, A.Tojiyev. Algorithms and hardware and software for improving energy efficiency in telecommunications facilities. *AIP Conf. Proc.* 3244, 060020 (2024). https://doi.org/10.1063/5.0242040

[5]. Collier M., Beel J. Implementing neural turing machines. *Lecture Notes in Computer Science*, 2018, vol. 11141, pp. 94–104. https://doi:10.1007/978-3-030-01424-7_10

[6]. T. A. Plate. Holographic reduced representations. *IEEE Transactions on Neural Networks.* no. 3, pp. 623–641. https://doi:10.1109/72.377968

[7]. P. Kanerva. Hyperdimensional computing: An introduction to computing in distributed representation with high-dimensional random vectors. *Computation*, 2009, no. 1, pp. 139–159. https://doi:10.1007/s12559-009-9009-8

[8]. E.Avery, T.Kelley, D.Davani. Tectures to improve robot control: Integrating production systems, semantic networks, and sub-symbolic processing. *Proc. 15 th Annual Conference on Behavioral Representation in Modeling and Simulation (BRIMS)*, 2006, pp. 190–198.

[9]. T.Kelley, E.Avery, L.Long, E.Dimperio. Symbolic and sub-symbolic intelligent system for mobile robots. *Proc. AIAA Infotech@Aerospace Conference*, 2009, p. 1976. https://doi:10.2514/6.2009-1976

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[10]. J.Parvar, D.Lowe, M.Emsley, R.Duff, A.Akintoye. Neural networks as a decision support system for the decision to bid process. *Proc. 16th Annual ARCOM Conference*, 2000, pp. 209–217.

[11]. Delen D., Sharda R. Artificial neural networks in decision support systems. In: *Handbook on decision support systems 1*. Springer, 2008. Pp. 557–580. https://doi:10.1007/978-3-540-48713-5_26

[12]. D. Golmohammadi. Neural network application for fuzzy multi-criteria decision making problems. *International Journal of Production Economics*, 2011, vol. 131, no. 2, pp. 490–504. https://doi:10.1016/j.ijpe. 2011.01.015

[13]. J.Figueira, S.Greco, M.Ehrgott. *Multiple criteria decision analysis: State of the art surveys*. New York, Springer-Verlag Berlin Heidelberg, Springer Publications, 2005. 1267 p. https://doi:10.1007/978-1-4939-3094-4

[14]. A.V.Demidovskij, E. A. Babkin. Developing a distributed linguistic decision making system. *Business Informatics*, 2019, vol. 13, no 1, pp. 18–32. https://doi:10.17323/1998-0663.2019.1.18.32

[15]. F.Herrera, L.Martinez. A 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Transactions on Fuzzy Systems*, 2000, vol. 8, no. 6, pp. 746–752. https://doi:10.1109/91.890332

[16]. W. Hendricks, “Review of NoSQL Data Stores: Using a reactive threetier application for software developers to achieve a high availability application design architecture,” In 2019 Open Innovations (OI), 2019, pp. 71-77.

[17]. Mahruqi, R. S. A., Alalfi, M. H., & Dean, T. R. (2019, November). A semi-automated framework for migrating web applications from SQL to document oriented NoSQL database. In Proceyedings of the 29th Annual International Conference on Compyuter Science and Software Engineering (pp. 44-53).

[18]. P. Kolyasnikov et al., “Analysis of software tools for longitudinal studies in psychology,” International Journal of Advanced Compyuter Science and Applications, vol. 10, no. 8, pp. 21– 33, 2019.

[19]. Roy-Hubara, N., Shoval, P., & Sturm, A. “A Method for Database Model Selection,” In Enterprise, Business-Process and Information Systems Modeling, 2017, pp. 261-275.

[20]. Chopade, R., & Pachghare, V. (2020). MongoDB Indexing for Performance Improvement. In ICT Systems and Sustainability (pp. 529-539). Springer, Singapore.

[21]. Alessandro P., Statistical methods for biomedical signal analysis and processing. Ph.D. Thesis of, European doctorate program in information technology, CYCLE XXI - ING-INF/01, January 2006 - March 2009.

[22]. Rogachev, A.F. Manufacturing and Consumption of Agricultural Products as a Tool of Food Security Management in Russia / Rogachev, A.F., Shokhnekh, A.V., Mazaeva, T.I. // REVISTA GALEGA DE ECONOMÍA (RGE) – Vol. 25-2. July-December 2016.

[23]. J. K. Chen and W. Z. Lee, “An Introduction of NoSQL Databases based on their categories and application industries,” Algorithms, vol. 12, no. 5, p. 106, 2019

Received: August 14, 2025

[24]. Tormozov V.S., Zolkin A.L., Vasilenko K.A. Structural optimization of the perceptron using the evolutionary modeling apparatus for the problem of predicting the values of time series // Industrial ACS and controllers. 2020; 8. -P.40-45. (In Russ.).

[25]. Kurchenkov K.B. Electronic document management. Criteria for the development of electronic document management systems. [Вестник Воронежского института високых текнологий] Bulletin of the Voronezh Institute of High Technologies. 2014, 12. –P. 102-106. (In Russ.).

[26]. A.K.Nikhath and K.Subrahmanyam Feature selection, optimization and clustering strategies of text documents // International Journal of Electrical and Computer Engineering (IJECE). Vol. 9, №2, Apr. 2019. –P. 1313. https://doi: 10.11591/ijece.v9i2.– P.1313-1320

[27]. R. Delima, H. B. Santosa, and J. Purwadi. Development of Dutatani Website Using Rapid Application Development // IJITEE (International Journal of Information Technology and Electrical Engineering). Vol. 1, № 2, 2017. – P. 36-44.

[28]. T.V.Khronusova, S.V.Kruchinin and E.V.Bagrova. Implementation of Electronic Document Management in Russian Education. Quality Assessment // 2019 International Conference "Quality Management, Transport and Information Security, Information Technologies" (IT&QM&IS). 2019. –P. 608–610. doi: 10.1109/ITQMIS.2019.8928356.)

[29]. R.Waszkowski, M.Kiedrowicz, T.Nowicki and K.Worwa. Access control management in administrative office with RFID tagged documents // Proceedings - 2017 4th International Conference on Mathematics and Computers in Sciences and in Industry, MCSI 2017. Vol. 2018-Janua, 2018. – P. 244–251. doi: 10.1109/MCSI.2017.48.

[30]. M.K.Ugale, S.J.Patil and V.B.Musande. Document management system: A notion towards paperless office // 2017 1st International Conference on Intelligent Systems and Information Management (ICISIM). 2017. – P. 217–224. doi: 10.1109/ICISIM.2017.8122176.

[31]. T.Mandicak, P.Mesaros, M.Spisakova and A.Kanalikova. Exploitation of document management systems in construction industry // IOP Conference Series: Materials Science and Engineering. Vol. 1252, №1, 2022. –P. 012077. doi: 10.1088/1757-899x/1252/1/012077.

[32]. S.V.Koptyakova, E.G.Zinovyeva and T.V.Maiorova Development and deployment of automated electronic document management system in municipal units // Journal of Physics: Conference Series. Vol. 1333, №.7, 2019. doi: 10.1088/1742-6596/1333/7/072034.

[33]. M.J.A. Shobaki, S.S.A. Naser, M.Khair, I.Kassab and S.S.A.Naser. The reality of the application of electronic document management system in governmental institutions-an empirical study on the palestinian pension agency // International Journal of Engineering and Information Systems. Vol.1, №2, 2017. –P. 1-14.

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)
[34]. N.L.Adam, M.A. Mansor, M.F.Pauzi and S.C.Soh. Document and event record management system: a prototype. // 2022 IEEE 12th Symposium on Computer Applications and Industrial Electronics (ISCAIE), 2022. –P. 223–227. doi: 10.1109/ISCAIE54458.2022.9794525.

[35]. Malekani, Andrew and Alphonce, Alex Bahame ABA. Assessing the efficacy of electronic document management system in records management at Sokoine University of Agriculture Library Philosophy and Practice (e-journal). 2022. –P. 7417.