# Use of mathematical model of artificial neural network for current forecasting of tax base of region

*Maksadkhan* Yakubov[1], *Adilbek* Turgunov[2], and *Gulchekhra* Jamalova[3*]

[1]Tashkent University of Information Technologies named after Muhammad al-Khorezmi, Tashkent, Uzbekistan
[2]Karshi branch of the Tashkent University of Information Technologies named after Muhammad al-Khorezmi, Karshi, Uzbekistan
[3]Karshi Engineering and Economics Institute, Karshi, Uzbekistan

**Abstract.** In the article, the method allowing the possibility of a neural(neuronic) network for forecasting the tax base of a region is offered. The given method is suitable for working out of the positive short-term forecast.

## 1 Introduction

Forecasting has become an integral tool for studying and using in socio-economic relations. One of the most important application areas for the development and use of forecasts is the taxation system.

Tax forecasting is an assessment of the tax potential of the regions and the receipt of taxes and fees in the budget system (consolidated, regional, and territorial budgets). It is based on the results of the socio-economic development of the subjects of the Republic.

The forecasting process includes determining the tax bases for each tax, monitoring the dynamics of tax receipts over several periods, and calculating the collection levels of taxes and fees. Forecasting has three main stages: retrospective, positive, and normative analysis.

In the context of strengthening and expanding the tax base of the republican budget, as the main stabilizing factor in the state economy, the most important task of tax forecasting at the stage of retrospective analysis is a comprehensive study of the effectiveness of the tax process and the changes taking place in the tax system in conjunction with the development of the macroeconomic situation in the regions based on forecasting development trends tax base [1-3].

### 2 Main part

The simplest method for determining trends in the tax base development is the extrapolation trend method, the continuation in the future of those trends that have developed in the past [1]. This method is based on information on the volume of taxable

*Corresponding author: guli.jamalova@mail.ru

bases, on the receipt of specific taxes and fees for determining the periods of previous years, on the volume of shortfalls in income, on the status of tax debts, on an analysis of trends in the development of the taxable base, the structure of taxpayers [2]. For current and operational forecasting of the tax base of the regions, mathematical models of artificial neural networks can be used since forecasting economic parameters in the short term using these models gives good results.

## 3 Object and methods of research

Consider the methodology for the current forecasting of the tax base of the regions using an artificial neural network.

Artificial neural networks (ANN) - mathematical models, as well as their software or hardware implementations, built on the principle of organization and functioning of biological neural networks - networks of nerve cells of a living organism.

From the point of view of machine learning, a neural network is a special case of pattern recognition methods, discriminant analysis, clustering methods, etc. From a mathematical point of view, neural network training is a multi-parameter nonlinear optimization problem. From the point of view of cybernetics, the neural network is used in adaptive control tasks and as an algorithm for robotics. From the point of view of the development of computer technology and programming, a neural network is a way to solve the problem of efficient parallelism. And from the point of view of artificial intelligence, ANN is the basis of the philosophical trend of connectivism and the main direction in the structural approach to studying the possibility of a building (simulating) natural intelligence using computer algorithms.

Neural networks are not programmed in the usual sense of the word; they are trained. The ability to learn is one of the main advantages of neural networks over traditional algorithms. Technically, learning is about finding the coefficients of connections between neurons. In the learning process, the neural network can identify complex relationships between input and output data and perform generalization. This means that in case of successful training, the network will be able to return the correct result based on the data missing in the training sample, as well as incomplete and/or "noisy" partially distorted data.

The ability of a neural network to predict directly follows from its ability to generalize and highlight hidden dependencies between input and output data. After training, the network can predict the future value of a certain sequence based on several previous values and/or some currently existing factors. It should be noted that forecasting is possible only when the previous changes predetermine the future ones to some extent.

In our case, the neural network can identify patterns of changes in the tax base for a short period. We will assume that the period of up to a year will be short-term since the factors affecting the receipt of taxes in the regional budget are quite inert and, even with a great change in their parameters, will significantly determine the future values of the tax base indicators in a given year.

There are many artificial neural networks; however, when performing prediction functions, a multilayer perceptron has proven itself well - a special case of the Rosenblatt perceptron, in which one backpropagation algorithm trains all layers. The neural network training takes place with the teacher [3]. This multilayer perceptron is a system of connected and interacting simple processors (artificial neurons) (Fig.1). The inputs of the processors receive signals that are amplified or weakened depending on the weight coefficients w. The values of all amplified or attenuated signals from the input of each neuron are fed to the adder. The sum of these values comes as an argument to the activation function from the adder. The activation function can be any nonlinear function; however, the system's learning ability and the forecast's accuracy largely depend on its choice.

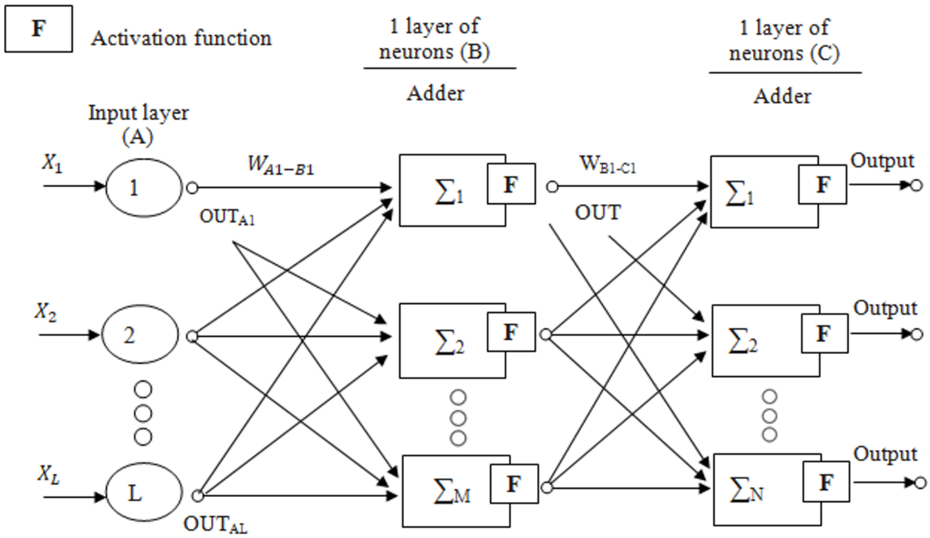


**Fig. 1.** Multilayer perceptron

It is recommended to use the logistic function - the sigmoid, given by an expression of the form:

$$OUT = \frac{1}{1+exp(-\alpha Y)} \qquad (1)$$

In this function, the argument $Y$ is totalizer output value, factor $\alpha$ determines the steepness of the function's graph, and accordingly OUT neuron state – excited (close to 1) or inhibited (close to 0). Outputs OUT are transmitted further through the layers of the network in a similar way until the exit.

As we indicated earlier, this multilayer perceptron is trained using the backpropagation method.

Let us describe the algorithm of this method

According to the least squares method [4], the objective function of the TS error to be minimized is the value:

$$E(w) = \frac{1}{2}\sum_{j,p}\left(y_{j,p}^{(N)} - d_{j,p}\right)^2 \qquad (2)$$

where$y_{j,p}^{(N)}$ is real output state of the neuron$j$output layer$N$neural network when applying an image to its inputs; $d_{j,p}$ is the ideal (desired) output state of this neuron.

The summation is carried out over all neurons of the output layer and all images processed by the network. Minimization is carried out using the gradient descent method, which means adjusting the weight coefficients as follows:

$$\Delta w_{ij}^{(n)} = -\eta \cdot \frac{\partial E}{\partial w_{ij}} \qquad (3)$$

Where $w_{ij}$ is synaptic connection weight, connecting $i$-th layer neuron $n-1$ with j-th layer neuron$n$; $\eta$ is learning rate factor, $0 < \eta < 1$.

 https://doi.org/10.1051/e3sconf/202340102062

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial y_j} \cdot \frac{dy}{ds_j} \cdot \frac{\partial s_j}{\partial w_{ij}} \quad (4)$$

Here, *y_j,* as before, means the output of the neuron$j$, and under $s_j$ – the weighted sum of its inputs, that is, the argument of the activation function. Since the multiplier $dy_j/ds_j$ is the derivative of this function concerning its argument; it follows from this that the derivative of the activation function must be defined on the entire x-axis. In this regard, the unit jump function and other activation functions with inhomogeneities are unsuitable for the considered TS. They use smooth functions such as the hyperbolic tangent or the classical sigmoid with exponent. In the case of hyperbolic tangent

$$\frac{dy}{ds} = 1 - s^2 \quad (5)$$

Third multiplier $\partial s_j / \partial w_{ij}$, obviously, equal to the output of the neuron of the previous layer $y_i(n/1)$.

As for the first factor in (4), it is easily decomposed as follows:

$$\frac{\partial E}{\partial y_j} = \sum_k \frac{\partial E}{\partial y_k} \cdot \frac{dy_k}{ds_k} \cdot \frac{\partial s_k}{dy_j} = \frac{\partial E}{\partial y_k} \cdot \frac{\partial E}{\partial y_k} \cdot w_{jk}^{(n+1)} \quad (6)$$

Here summation over $k$is performed among the neurons of the layer $n + 1$.
By introducing a new variable

$$\delta_j^{(n)} = \frac{\partial E}{\partial y_j} \cdot \frac{dy_j}{ds_j}, \quad (7)$$

we get a recursive formula for calculating the quantities $\delta_j(n)$ layer n from quantities $\delta_k(n+1)$older layer $n + 1$:

$$\delta_j^{(n)} = \left[\sum_k \delta_k^{(n+1)} \cdot w_{jk}^{(n+1)}\right] \cdot \frac{dy_j}{ds_j} \quad (8)$$

For output layer

$$\delta_l^{(N)} = \left(y_l^{(N)} - d_l\right) \cdot \frac{dy_l}{ds_l} \quad (9)$$

Let's find the weight increments

$$\Delta w_{ij}^{(n)} = -\eta \cdot \left[\sum_k \delta_k^{(n+1)} \cdot w_{jk}^{(n+1)}\right] \cdot \frac{dy_j}{ds_j} \cdot y_i^{(n-1)}. \quad (10)$$

Now we can write (3) as:

$$\Delta w_{ij}^{(n)} = -\eta \cdot \delta_j^{(n-1)} \cdot y_i^{(n-1)} \quad (11)$$

Sometimes, to give the process of weight correction some inertia, which smooths out sharp jumps when moving over the surface of the objective function, (11) is supplemented with the value of the weight change at the previous iteration

$$\Delta w_{ij}^{(n)}(t) = -\eta \cdot \left(\mu \cdot \Delta w_{ij}^{(n)}(t-1) + (1-\mu) \cdot \delta_j^{(n)} \cdot y_i^{(n-1)}\right) \quad (12)$$

where $\mu$ is inertia coefficient; $t$ is number of the current iteration.

1. Thus, the complete TS training algorithm using the backpropagation procedure is constructed as follows:
2. Apply one of the possible images to the network inputs, and in the normal operation mode TS, when signals propagate from inputs to outputs, calculate the latter values. Recall that

$$s_j^{(n)} = \sum_{i=0}^{M} y_i^{(n-1)} \cdot w_{ij}^{(n)} \tag{13}$$

where$M$ – number of neurons in a layer$n - 1$taking into account the neuron with a constant output state +1, which sets the bias; $y_i(n-1) = x_{ij}(n) - i$-th neuron input $j$layer$n$.

$$y_o(n) = f\big(s_o(n)\big) \tag{14}$$

where$f(s)$ – sigmoid.

$$y_q(0) = I_q \tag{15}$$

where$I_q q$-th component of the input image vector.

3. Calculate $\delta(N)$ for the output layer by the formula (9).

Calculate by formula (11) or (12) weight changes $\Delta w(N)$ layer $N$.

4. Calculate by formulas (8) and (11) (or (8) and (12)) respectively $d(n)$ and $\Delta w(n)$ for all other layers, $n = N - 1$, ..., 1.

Adjust all weights in TS

$$w_{ij}^{(n)}(t) = w_{ij}^{(n)}(t-1) + \Delta w_{ij}^{(n)}(t) \tag{16}$$

5. If the network error is significant, go to step 1. Otherwise, the neural network's learning process is completed, and you can start testing it.

## 4 Results and discussion

Having chosen a training method, preparing training and testing samples is necessary. When preparing the training and testing samples, it should be borne in mind that the neural network does not learn well on large values. Therefore, to form training and testing samples, it is recommended that only changes in the tax base values be fed to the network input, normalized in such a way that, if possible, the values belong to the segment [0;1]. This is also because the output values of the neural network will be from the same segment.

The training sample will be data for the previous period, objectively reflecting the real state of the tax base. It is advisable that the frequency of these data be equal to a month. The testing sample is used to check the adequacy of the network training and allows concluding the sufficiency or insufficiency of training.

The structure of the neural network is determined by the ultimate goal of the study - to give the most accurate forecast possible. The predicted value is the value of the tax base for a certain period following the period whose value is known. Therefore, the output of the neural network will be one. The inputs are the values of the tax base for previous periods in order, and the output is the next known value. To effectively train the network, you need to have enough data. The required values in the training sample and the prediction accuracy depend on the number of inputs. As the number of inputs increases, more data is required for training, while the accuracy increases significantly. Most often, the number of inputs is found from a trade-off between the required accuracy and the amount of information available. During training, n values of the tax base are fed in parallel to the inputs, and the output $(n+1)$-th value. Then there is a shift to the left of the values at the inputs and output. That is, the inputs are values from 2 to$(n+1)$-th, and on the way out – $(n+2)$-th meaning. In this way, the shift occurs k times, taking into account the constraint$(n+k) = N-1$, where$N$– the number of values in the training set.

To achieve the required accuracy with the backpropagation method, you will probably have to spend several iterations of training - epochs. The number of epochs will depend on the required accuracy and neural network settings - learning rate, number of hidden layers, and number of neurons in the hidden layer. After the required accuracy is achieved, the training adequacy is checked based on the testing sample. This sample also contains a pair of values - inputs and outputs; however, unlike the training sample, the output value in this pair serves only to evaluate learning. Only the first values of the pair in the testing sample are fed into the input, and the adequacy of training is checked using the second values. After the end of testing, a conclusion is made about the adequacy of training. If the training goes well, then they start predicting; otherwise, they train the network again with the same data or look for an opportunity to increase the number of elements in the training sample, tune the network, and select the optimal number of inputs. The values of the network settings are selected empirically, so it can be quite a laborious process.

After the conclusion is drawn that the network is trained adequately, it is possible to start forecasting. When applying $(n+k+1)$output value, the network will produce a new predicted value for the subsequent period. To predict the value of the second period, the forecast value for the first period is fed to the network input. Thus, recursively, the network can issue forecasts for any period. However, it should be noted that as you move away from the known forecasting periods, the accuracy inexorably decreases. This can be explained by the accumulation of prediction error over all$M$predicted values and its impact on $(M+1)$-th the value of the forecast of the tax base of the subject of the Republic [5-9].

Thus, the described methodology makes it possible to predict and track the trend of the region's tax base.

## 5 Conclusion

Short-term forecasting of tax revenues and fees is of particular importance since the unreality of annual forecast estimates causes difficulties in mobilizing the planned revenues by the tax authorities into the budgetary system of the Republic or an unreasonably high fulfillment of the established tasks for the subjects of the state, which entails a violation of the budget balance, and, as a result, destabilization processes in the socio-economic system. Trend analysis can be useful for making decisions on changing the parameters of the tax system when assessing decisions already made by state legislative and executive authorities at all levels and developing measures aimed at stabilizing the functioning of the socio-economic system of the region.

## References

1. Black J., Hashimzade N., and Myles G. A dictionary of economics. Oxford university press. (2012).
2. Downes J., and Goodman J. E. Dictionary of finance and investment terms. Simon and Schuster. (2014).
3. Kruglov V. V., Borisov V. V. Artificial neural networks. Theory and practice. Hotline - Telecom, p. 382. (2001).

4. Yakubov M. S., Jamalova G. B. Intellectual models and methods of supporting managerial decision-making in the tax service. Modern concepts of scientific research. 72nd International Scientific Conference. Eurasian Scientific Association. No. 2 (72), (2021).
5. Jamalova G. B. Tax policy at macro and microeconomic levels, its essence and principles of development. International scientific and technical journal. Innovation technical and technology Vol.1(4).
6. Jamalova G.B. Information modeling using artificial neural networks. Academic research in educational sciences, Vol. 1(3), (2020).
7. Yakubov M. S, Jamalova G. B. Adaptation of C-Average and Gustafson-Kessel algorithms for intellectual information processing in the tax unit services - tax inspection. ACADEMICIA An International Multidisciplinary Research Journal (Double Blind Refereed and Peer Reviewed Journal), Vol. 11(2), (2021).
8. Turgunov A. M., Jamalova G. B. Modeling and implementation of intellectual support of management activities in the tax inspectorate. Economy and Society, № 6(85) (2021).
9. Turgunov A.M., Jamalova G.B. Management decision-making based on forecasting tax revenues. Academic research in educational sciences. Vol. 2(4), (2021).