Improvement of decisions support subsystems of information-analytical systems on the neural networks ensembles basis

Mikryukov Andrey

PhD, associate professor, Plekhanov University, 89166019804,
Mikrukov.aa@rea.ru

1 Introduction

At present, neural network approaches, in particular, the use of ensembles of neural networks, which are an example of a collective solution of problems, have been widely used for modeling complex sociotechnical systems that are difficult to formalize and poorly structured.

The solution quality of the specific task (data mining, forecasting, pattern recognition, classification, etc.) can be significantly improved using neural network ensembles, in which it is expected to form and train a finite set of neural networks, the results of which are taken into account in the overall solution. At the same time, individual decisions are coordinated in such a way that the overall final decision is the best.

One of the fundamental problems of improving the functioning of an ensemble of neural networks in the subsystems of decision support for information and analytical systems (DSS IAS) in terms of increasing their accuracy and reliability is the generation of diversity ensemble (the differences of individual models) [1]. Aggregation of similar models in the ensemble can not lead to a significant improvement in the quality of the solution of the problem.

To resolve this contradiction, it is proposed to use approaches based on the collective of the neural networks, for the construction of which a new class of artificial neurons, the so-called selective neurons, differing from classical neurons by a more efficient method of processing input information approximating the biological neuron is used [2]. Due to the peculiarities of their construction, training and functioning, selective neurons provide, on the one hand, the construction of more efficient neural networks for solving the tasks posed, and, on the other hand, the greater diversity of the neural network ensemble.

2 Neural network ensembles based on elective neural networks.

The models used in the IAS-based IAS based on neural networks have a number of features and advantages. They are adaptive self-learning systems, which are difficult to dynamically simulate, and often, it is simply impossible, because They often contain a significant array of hidden, uncontrolled, incomplete and noisy parameters and mutual connections between them. Their use makes it possible to solve problems that are difficult or impossible to solve by traditional methods due to the absence of formalized mathematical descriptions of the functioning of the object of investigation. Neural network models have associative memory and in the process of work they accumulate and generalize information, from which their effectiveness increases with time. Their use is based on training the neural network to extract information from experimental data, which ensures the objectivity of the results and increases their reliability and reliability [3].

When constructing an ensemble of neural networks, a finite set of previously trained neural networks is simultaneously used, the output signals of which are combined into a joint estimate, which exceeds the quality of the results obtained with the help of local networks entering the ensembles.

The ensemble $H(\bar{x})$ of the models $h_i(\bar{x})$ (i=1,2, ... N) is the composition of the algorithmic operators h_i: $R^d \to R$ and the corrective operation

F: $R^N \to R$, in which the final estimate of $H(\bar{x})$ [4] is put in correspondence with the set of estimates $h_1(\bar{x})$, $h_2(\bar{x})$,..., $h_N(\bar{x})$:

$$H(\bar{x}) = F(h_1(\bar{x}),\, h_2(\bar{x})\,, \ldots,\, h_N(\bar{x})). \tag{1}$$

As is known, the fundamental task in the construction of ensembles is the generation of a variety of ensemble (or the difference of individual models) [5].

Obviously, the aggregation of similar models in the ensemble can not lead to a significant improvement in the quality of the solution of the problem.

The problem of generating diversity lies in the fact that individual models of classical neuron networks based on a neuron McCullock-Pitts, are trained to solve one task for one training sample and, as a result, usually quite strongly

Correlated, which affects the accuracy of the solution obtained.

Electoral networks are built on the basis of selective neurons, whose model is close to the model of a biological neuron. The training of the elective neural network is carried out not by changing the weight coefficients of the synaptic connections, but by changing the quantity and quality of the inhibitory and exciting dendrites (signal transmitters) on the basis of which selective neurons are constructed [1].

A mathematically artificial neuron is usually represented as a non-linear function of a linear combination of signals with a limited number of inputs. The result of the function is directed to the only output of the neuron, and then to the input of the neuron from the next layer. Scales characterize communication channels, through which signals are received. The task of network training is to select such weights for each connection that would minimize the final value of the output error. However, the classical types of neural networks have a serious drawback – instability to retraining. If the number of weights in the network is large, then the operation of summing up their linear combination will be extremely computationally capacious, and the effect of retraining is likely to arise when the network recognizes images on the training set with extreme accuracy, but it shows a large error percentage on the test set Input data.

Thus, the shortcomings of classical neural networks are a consequence of the use of weighting coefficients of synaptic connections. Indeed, in a biological neuron, there is no weight gradation – this is an artificial device. The solution of the main problems of processing the input signal is achieved by changing the clustering configuration and the number of dendrites at the input of the biological neuron [6].

The analysis showed that, by analogy with biological mechanisms, it is advisable to use a neural network that includes neurons with controlled clustering of the input channels (synapses) in quantity, and then clustering the channels in quality (exciting or inhibiting). As an effective classifier of images, it was suggested in [2] to use a neural network based on selective neurons. The selective neuron does not have weight coefficients of synaptic connections and is close in properties to its biological analogue (Fig. 1).

Cluster formation includes blocking of non-informative communication channels that do not conduct exciting or inhibitory signals [2]. Thus, after training, each neuron has a cluster with an individual transfer characteristic, where some inputs for signals that are obviously different from the training code combination are blocked.

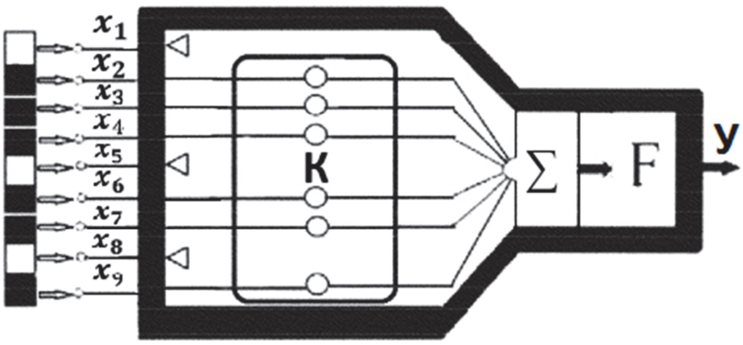

Fig. 1. Scheme of the selective neuron:

$x_1, \ldots, x_9$ – input signals in the form of a binary code;

K – a cluster of communication channels formed in accordance with the binary code at the input;

Σ – adder;

F – a nonlinear threshold function.

3 Mathematical model of the selective neuron

Possible characteristic code combinations of the investigated objects at the input of the selective neuron can be represented in the form of vectors [2]:

$$\mathbf{x_1} = (x_{11}, \ldots, x_{1n}) \ ; \ \ldots \ ; \ \mathbf{x_m} = (x_{m1}, \ldots, x_{mn}) \ , \tag{2}$$

where n is the number of elements of the code combination; M = is the number of objects to be examined. All possible code combinations of input objects form a matrix A, which can be represented as:

$$\mathbf{A} = \begin{pmatrix} x_{11}, x_{12} \ldots, x_{1n} \\ \ldots \quad \ldots \quad .. \\ x_{m1}, x_{m2} \ldots, x_{mn} \end{pmatrix} \tag{3}$$

Let a particular selective neuron contain a cluster of bonds characterized by a code combination $\mathbf{x_i} = (x_{i1}, \ldots, x_{in})$. When entering the input combination of the input neuron at the input of the neuron, we obtain

$$S_{ij} = \sum_{k=1}^{n} x_{ik} x_{kj} = (\boldsymbol{x}_i, \boldsymbol{x}_j) \tag{4}$$

The sum values S_{ij} are equal to the elements of matrix *B*:

$$B = A * A^T \ , \tag{5}$$

Where A^T is the transposed matrix of *A*. We obtain in total *m* x *m* sums. The largest is the sum:

$$S_{ii} = \sum_{j=1}^{n} x_{ij} x_{ji} = (\boldsymbol{x}_i, \boldsymbol{x}_i) = N_i = N \ , \tag{6}$$

where N_i is the number of units in the code combination $\mathbf{x_i} = (x_{i1}, \ldots, x_{in})$. The property of the sums is that $S_{ij} < N$, it is used to decide on the recognition of input signals.

Due to the peculiarities of its construction and functioning, the selective neural network provides:

- selective recognition achieving of input signals without using the weighting of their synaptic connections;
- the possibility achieving of encoding an input signal of a certain type with the channel number or the number of the registering neuron;
- input information compression, due to the preservation of only that information about objects that falls into a given channel or registering a neuron;
- the speed increase of operation;
- the reliability increasing of object recognition when they are large;
- achievement of a much greater adequacy of the selective neuron to a biological single-layer perceptron.

Simulation of the time series forecasting process based on an ensemble of selective neural networks has shown that the average error in forecasting individual models of electoral neural networks has decreased by 8-10%. The quality of the diversity of the models of elective neural networks, calculated as the difference between the quadratic error of the ensemble and the average error of individual models of electoral neural networks has improved (increased) by 10-14%.

4 The conclusion

Studies have shown that the use of a new class of neural networks based on selective neurons for the construction of neural network ensembles in subsystems supporting the decision-making of information and analytical systems makes it possible to significantly improve the accuracy and reliability of decision-making by an ensemble of neural networks.

The key difference between the classical McCulloch-Pitts neuron from the selective one is that the model of the latter is close to the model of a biological neuron. The operation of selective neural networks is based on the controlled clustering of the input channels, individual for each neuron, which in turn ensures an increase in the accuracy and reliability of their operation.

References

1. Kuncheva L.I. Combining Pattern Classifiers: Methods and algorithms. John Wiley & Sons, Hoboken, NJ, 2004.
2. Mazurov M.E. Single-layer perceptron based on selective neurons. Patent for invention No2597497 dated January 13, 2015.
3. A.A. Morozov, V.P. Klimenko, A.L. Lyakhov, S.P. Alyoshin State and prospects of neural network modeling DSS in complex sociotechnical systems // Mathematical machines and systems. – 2010.-No.1. P.127-149.
4. Zhuravlev Yu.I. On the Algebraic Approach to the Solution of the Problems of Recognition and Classification // Problems of Cybernetics. 1978. T.33. – C.5-68.
5. Zhou Z.- H. Ensemble Methods: Foundations and algorithms. Chapman & Hall / Crc Machine Learning & Pattern Recognition. 2012. 236 p.
6. Gutman A. Dendrites of nerve cells: theory, electrophysiology, function. 1984.