Information and Neural Educational System for Training Standard and Selective Neural Network Technologies in Universities

Mikhail Mazurov [1[0000-0001-9993-4687]], Andrey Mikryukov [1[0000-0002-8206-677X]], M. Gorbatih [2], T. Bergaliev [1], Natalia Shchukina [1[0000-0002-5825-1051]]

[1] Plekhanov Russian University of Economics, Stremyanny lane 36, 117997 Moscow, Russia
mazurov37@mail.ru mikrukov.aa@rea.ru
[2] Lomonosov Moscow State University, Leninskie gory 1, 119991 Moscow, GSP-1, Russia

Abstract. The theoretical and mathematical substantiation of standard and selective neural network technologies is given. Models have been developed for visual modeling of processes in standard neural networks based on McCulloch-Pitts neurons and selective based on selective neurons. An assessment of the accuracy of recognition in selective neural networks is given, based on an assessment of the fulfillment of the basic conditions for recognition. The neuroeducational system allows effective teaching of neurotechnologies to schoolchildren, students, and specialists in related professions.

Keywords: McCallock-Pitts neuron, selective neuron, single-layer Rosenblut perceptron, selective perceptron, Monte-Carlo selective learning method

1 Introduction

The theoretical and mathematical substantiation of standard and selective neural network technologies is given. The neuro-educational system for teaching neural network technologies and their applications based on standard and selective neural networks is designed for pupils, students, graduate students, people of the "silver age", specialists in related professions. The possibility of solving more complex problems using the developed software, accessible to a wide audience of schoolchildren, students, specialists in related professions, has been realized. This is the calculation of the weight coefficients of single-layer and multi-layer neural networks with McCallock-Pitts neurons using the selective Monte-Carlo method. Familiarization with the methods of "deep learning" in neural networks.

The neuro-educational system is made in two versions. The first option is a material instrumental implementation in the form of training electronic working models.

The second option is computer implementation in the form of separate programs for the learning process and the use of more complex training tasks.

Material instrumental implementation in the form of training electronic operating mock-ups intended for training standard neurotechnologies based on McCallock-Pitts neurons and selective neurotechnologies based on selective

Proceedings of the of the XXIII International Conference "Enterprise Engineering and Knowledge Management" (EEKM 2020), Moscow, Russia, December 8-9, 2020.

neurons. For training standard neural networks, a visual technology based on the Monte-Carlo selective method has been developed, the programs are presented in the form of executable exe files.

In the second computer version, more complex training examples for recognizing objects on a large-screen monitor are considered. The developed software allows you to go to the solution of practically useful tasks of object recognition and control tasks using selective neurotechnologies. Tasks are developed that illustrate the training methods used in neural networks of "deep" learning.

In order for the neural network to be able to complete the task, it needs to be trained. Training the most common neural networks using McCallock-Pitts neurons comes down to calculating its weighting coefficients [1-5, 12-20]. The process of teaching with a teacher is the presentation of a network of sample training examples. Each example is fed to the inputs of the network, then it is processed inside the structure of the neural network, the output signal of the network is calculated, which is compared with the corresponding value of the target vector representing the desired output of the network. Then, according to a certain rule, an error is calculated, and the weighting coefficients of the connections within the network change, depending on the selected algorithm. The vectors of the training set are presented sequentially until the error over the entire training array reaches an acceptably low level. The structural diagram of the neural network learning process is shown in Fig. 1.

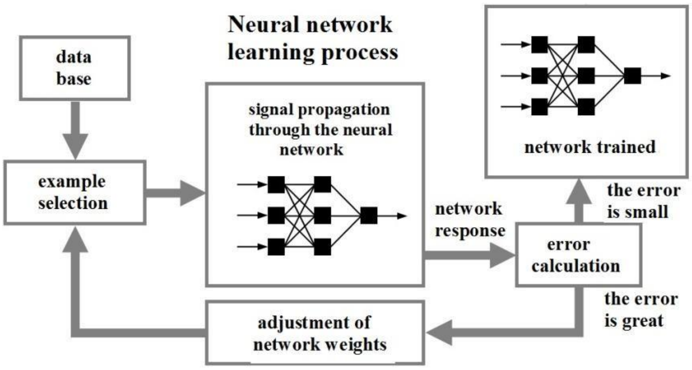

Fig. 1. The structural diagram of the learning process of a neural network

Various iterative methods are used to calculate weights. As is known, these methods have a number of fundamental irremovable drawbacks, we list some of them: 1. The technical complexity of performing iterative procedures for finding training weights associated with a large amount of computation [13]; 2. The need to recalculate weights when adding new input features; 3. Instability, the possibility of ambiguous solutions to recognition and control problems for some sets of weighting coefficients [13]; 4. The main disadvantage is the ambiguity of sets of weighting coefficients satisfying the inequalities from which they are determined [13].

For effective training in neuro-educational technologies, a series of software tools have been developed, as well as the material implementation of the neuro-educational system.

2 Instrumental implementation of a neural educational system based on McCallock-Pitts neurons

For educational purposes, a single-layer perceptron neuro-educational system based on three McCallock-Pitts neurons was developed. The perceptron was supposed to recognize the letters L, T, X at the weight coefficients generated during the training. The proposed perceptron was implemented in the form of an electrical circuit shown in Fig. 2.

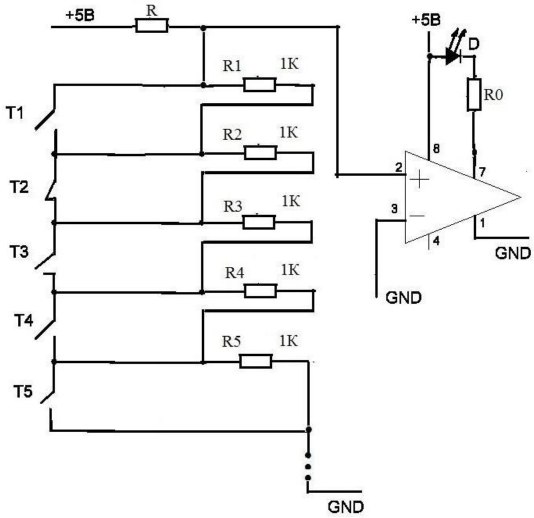

Fig. 2. The prototype of the proposed McCallock-Pitts neuron, based on the electrical circuit and the use of the comparator as a nonlinear threshold element

Provided $R \gg \sum_{i=1}^{9} R_i$ the output voltage is

$$u = \frac{R_{экв}}{R} E = \frac{R_1}{R} x_1 + \frac{R_2}{R} x_2 + \ldots + \frac{R_9}{R} x_9 = \sum_{i=1}^{9} \alpha_i x_i ,$$

where $\alpha_i \, (i = 1,\ldots, 9)$ are the weights; $x_i \, (i = 1, \ldots, 9)$ - input signals. Thus, we find that the sum linearly depends on the magnitude of the weighting factors.

To increase the recognition stability, a circuit is used that contains comparators - voltage limiters above and below.

The magnitude of the weight coefficients is regulated using slide variable resistances. Exceeding the threshold level is registered by the LED indicator. The perceptron contains three McCallock-Pitts neurons, which is enough for training purposes.

3 Selective neurons and neural networks

From the many shortcomings of neural networks based on training by selecting weights, selective neural networks using selective neurons described in [6-11]. In this paper, we consider the training of selective neural networks with selective neurons described in these works. The structures of the McCallock-Pitts neuron and the selective neuron in a black and white image are shown in Fig. 3, respectively, left and right.

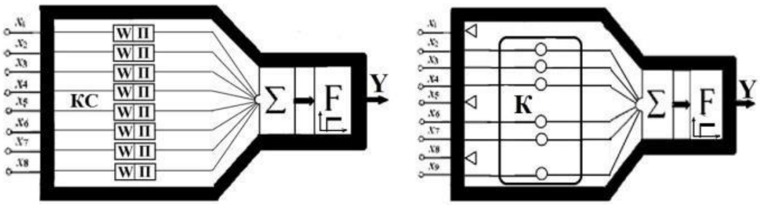

Fig. 3. Structures of the selective neuron (right) and the McCallock-Pitts neuron (left)

An effective way to obtain selectivity in a perceptron, a system of neurons, was implemented in a selective perceptron composed of selective neurons. In each neuron, clusters of specialized communication channels are created, tuned to the corresponding characteristic code combinations of the input signals. The structure of a single-layer Rosenblut perceptron using McCallock-Pitts neurons and a selective perceptron are shown in Fig. 4 left and right.

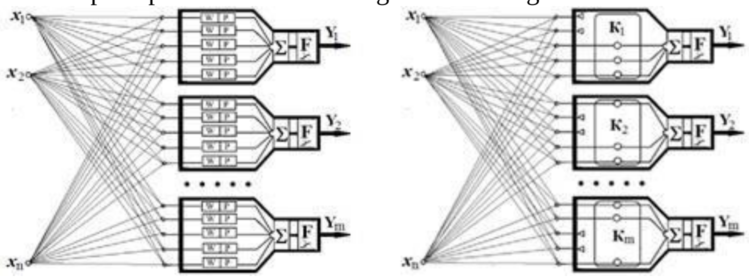

Fig. 4. The structure of a single-layer Rosenbluth perceptron using McCallock-Pitts neurons on the left and a selective perceptron on the right

In fig. 4: K - formed clusters of communication channels; - adder, F - threshold nonlinear elements. Triangles indicate blocked communication channels from among the input ones that are not essential for objects at the input of the perceptron. The effectiveness of the selective perceptron for the recognition of contour objects follows from the theorem given in [9, 10]. The fundamental difference between standard neural networks and selective neural networks is that standard neural networks are trained by selecting weights, and selective neural networks are trained through selective clustering of communication channels.

3.1. Theoretical information about neural networks on selective neurons

The structure of a material device was developed that implements the instrumental method of training a selective neural network using the example of 3 selective neurons and 3 input signals characterized by binary sequences of 9 cells, shown in Fig. 5

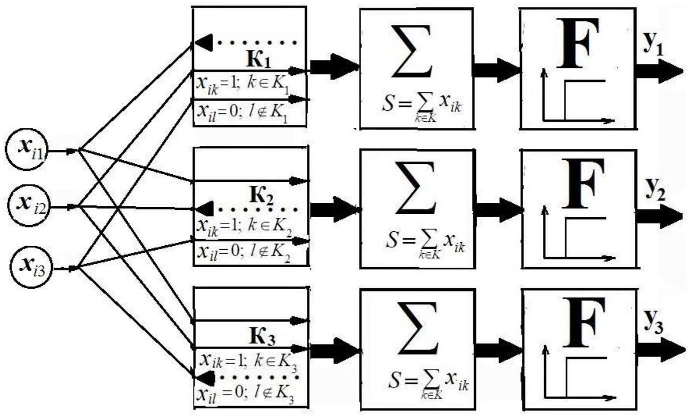

Fig. 5. The structure of the material device that implements the instrumental method of training the selective neural network

In fig. 5 for clarity, the recognition of 3 vector objects is modeled. Selective clustering is illustrated by the first three vertical blocks in the structural diagram. Blocking channels and the formation of clusters is done using black triangles. Summation of the elements of vectors $\boldsymbol{x}_i = (x_{i1},..., x_{i9})$ $(i = 1, 2, 3)$ is performed in the summation blocks. Then a nonlinear threshold transformation is performed in the blocks indicated $\boldsymbol{F}$.

To justify the uniqueness of recognition of contour images by a selective neural network, a pixel scan of normalized contour objects was implemented, matching each object with a binary string of 0 and 1 (0 - no contour, 1 - contour point). Then for mismatching contour objects the relation holds $(\bar{x_i}, \bar{x_j}) < N$, where $(\bar{x_i}, \bar{x_i}) = N \quad (i = 1,..., l)$. We believe that it is impossible to combine contour objects with a linear transformation $O_1,..., O_l$, which allows us to realize the linear independence of the binary codes of these objects. A uniqueness recognition theorem for incompatible contour objects has been proved [9, 10].

Theorem: Suppose that in a two-dimensional region divided into pixels by a rectangular lattice, *m* contour objects are given that are incompatible with movements — shift, rectangular transfer, and rotation. Let the objects be scanned using a horizontal scan into binary sequences - vectors of 0 and 1 in length n , that is $\mathbf{x_i} = (x_{i1},..., x_{in})$ $i = (1,..., m)$. Let all possible code combinations of input signals be collected in matrix A

$$\mathbf{A} = \begin{pmatrix} x_{11}, x_{12}..., x_{1n} \\ x_{21}, x_{22} ..., x_{2n} \\ ... \quad ... \quad ... \\ x_{m1}, x_{m2} ..., x_{mn} \end{pmatrix}$$

Let a particular neuron contain a cluster of connections characterized by a code combination in the form $\mathbf{x_i} = (x_{i1},..., x_{in})$. Consider the amount

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

We represent the sums S_{ij} in the form $\boldsymbol{A}\boldsymbol{x}_j^T$, where $\boldsymbol{x}_j^T$ is the transposed column vector from the row vector of the vector $\boldsymbol{x}_k$. We represent all possible sums S_{ij} $(i = 1,...,m;\ j = 1,...,m)$, the number of which is equal m^2, in the form of the matrix B

$$\boldsymbol{B} = \boldsymbol{A}\boldsymbol{A}^T ,$$

where A^T is the transposed matrix A. Consider the amount S_{ii}

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

where N_i is the number of units in the code combination $\mathbf{x_i} = (x_{i1},..., x_{in})$. We use the property to recognize input objects

$$S_{ij} < S_{ii} = N .$$

Then the recognition of each of m input objects by the considered single-layer perceptron will be unique.

The theorem states that the solution to the problem of the selective perceptron can be found in one iteration, that is, for a time interval much shorter than during the iterative process for the Rosenblut perceptron. In this case, the solution itself will be the only one.

3.2. Advantages achieved by implementing a selective perceptron

The selective perceptron based on selective neurons has several advantages over the Rosenbluth perceptron based on McCallock-Pitts neurons. We list the useful properties of a selective perceptron that can be achieved: 1. Solving recognition, control, and other problems when using the instrumental method in selective neural networks is unambiguous under certain conditions, while recognition in neural networks with McCallock-Pitts neurons is ambiguous; 2. The training procedure for selective neural networks is simplified, it becomes available to a layman; 3. Reduces training time; 4. Training can be achieved by tools without the use of computing tools.

3.3. Implementation of the selective perceptron

The selective perceptron on selective neurons has been implemented in practice. The electrical circuit of the selective neuron provides a light indication of all involved communication channels of the neuron with the control panel, which allows you to visualize the clustering of working communication channels. The control panel is common to all three selective neurons that make up the selective perceptron. The selective adjustment of the communication channels of a selective neuron to informative cells of input signals is performed with the corresponding internal communication channels located in the selective neuron itself. In this case, the active communication channels included in the selective cluster are identified by ignition of the LEDs located on the neuron. Behind the indicators of communication channels is a neuron excitation recorder.

4 The selective Monte-Carlo method for training and testing standard neural networks based on McCallock-Pitts neurons

The elimination of the shortcomings of iterative algorithms can be achieved using an innovative calculation procedure based on the selective Monte-Carlo method. Inequalities for determining weights have innumerable solutions. Various iterative procedures are used to find these solutions. However, what solution we will find after the iteration procedure is unknown. Some clarification of the situation is given by the well-known theorem of the American mathematician Novikov. Its essence is that it claims, under a number of restrictions (rather stringent), that learning - the iterative process will end in a finite time. And what set of weights and what quality of the system we get is unknown. The developed program allows you to find training weights directly using the selective Monte-Carlo method. The method allows you to find a large number of sets of weighting coefficients and selects those that satisfy the given optimality conditions. You can find sets lying in a given numerical range, sets that provide a stable mode of operation of a neural network, and sets that exclude ambiguous recognition.

4.1. Monte Carlo selective method software

The software is made in the programming language Matlab7. We present the results of testing the developed program. Consider an example of calculating weights for one of the training options. The calculation results of one possible set of weights for the signals, each of which is a binary sequence of 9 cells from 0 and 1, are presented below. The matrix of weights for binary input signals is equal to

0.9113	0.1775	0.3599	1.4603	0.4626	0.2949	1.2842	1.4467	0.3703
0.9878	1.0770	0.7659	0.2474	0.8311	0.3239	0.5889	1.1790	0.3721
0.4015	0.1678	1.1979	0.7318	1.3805	0.8073	0.8069	0.3543	1.0721

The initial matrix of input signals is

1	0	0	1	0	0	1	1	1
1	1	1	0	1	0	0	1	0
1	0	1	0	1	0	1	0	1

The matrix that controls the quality of the found set of weights is

5.4729	3.3580	3.3884
3.3752	4.8407	3.5458
3.3667	3.5020	4.8589

The quality of the found weighting coefficients can be judged by the homogeneity of the diagonal members of the matrix; in accordance with the calculations, the diagonal terms must be in the interval $4.8 < C_{ii} < 5.9$. In addition, the condition of "high quality recognition" must be met $C_{ij} \ll C_{ii}$.

This condition is also fulfilled. The number of calculation cycles is 375533. The counting time is less than 1 second.

A certificate of state registration of computer programs has been received for the software “The program for calculating perceptron weight coefficients using the selective Monte-Carlo method”.

5 Experimental study of the McCullock-Pitts perceptron

5.1. Perceptron testing based on McCallock-Pitts neurons

On the control panel that simulates the receptors of the input signals, sequentially activate the receptors corresponding to the letters L, T, X. The recognition indicators for the letters L, T, X should light up. Identification indicators of unidentified neurons should not light up. The entire learning process is illustrated by the Rosenbluth perceptron, which includes three McCallock-Pitts neurons. The single-layer Rosenblut perceptron in the recognition mode of the letters of the English alphabet L, T, X is shown in Fig.6.

Fig. 6. Single-layer Rosenbluth perceptron in the mode of recognition of letters of the English alphabet L, T, X

In Fig. 6 top left, the perceptron recognizes the letter L, the letter X right, in the center of the perceptron to recognize the letter T.

5.2. An experimental study of a perceptron based on selective neurons

Consider the recognition of the letters of the English alphabet L, T, X using a selective perceptron.

The appearance of a selective perceptron based on 3 selective neurons, designed to recognize three input objects in the form of letters, numbers and other characters that can be set on the control panel screen, is shown in the photograph in Fig. 7.

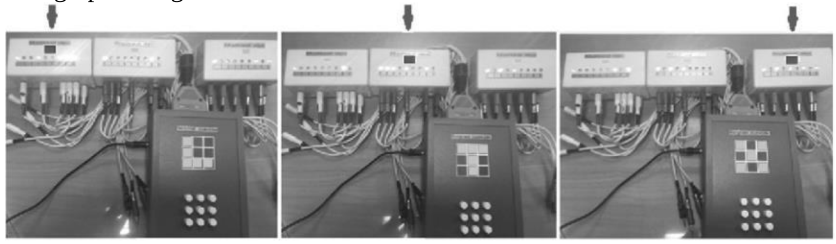

Fig. 7. Appearance of a perceptron based on selective neurons configured to recognize the letters of the English alphabet L, T, X. An experiment on the recognition of letters L, T, X is shown

In fig. 7, the perceptron recognizes the letter L from the top left, the perceptron recognizes the letter T from the top right, the perceptron recognizes the letter X from the bottom center.

5.3. An example of more complex recognition without mathematics in computer implementation

Consider the binary coding of 10 digits 0, 1, ..., 9 on a 4x6 monitor screen. The numbers on the monitor screen and their decomposition into binary sequences are shown in Fig. 8.

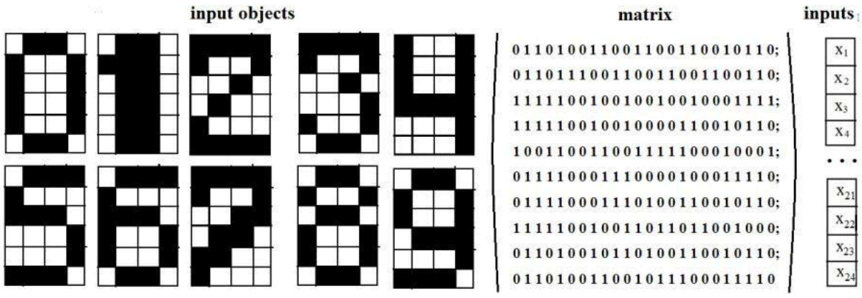

Fig. 8. The location of the numbers on the monitor screen and their decomposition into binary sequences

Decomposition of 10 digits from 0-9 in the form of a binary string. The significance of each expansion in the form of normalization according to the norm is taken into account. Each digit corresponds to a binary string of 24 cells of 0 and 1. The totality of all lines is shown in Fig. on the right, in the form of a 24x10 matrix. The binary input signals of the objects are shown in rows of a 24x10 matrix.

5.4. Smart ECG Recognition

For training, other more complex recognition options were implemented using an instrumental method based on the use of selective neural networks. This is a "smart" recognition of electrocardiograms. The well-known method of ECG recognition is based on the determination of individual characteristics of an ECG type: maximums, minimums, inflection points, etc. Smart ECG recognition based on selective neural networks is based on a comparison of a real ECG with a reference from a database. Software has been developed that allows you to enter a real ECG and recognize its belonging to one of the ECGs from the database [8].

5.5. Evaluation of recognition accuracy in selective neural networks

The estimation of the recognition accuracy in neural networks is estimated using the relation

$$\delta = \frac{M - M_0}{M} 100\% ,$$

where M - the total number of recognitions, M_0 is the number of erroneous recognitions in practice. For selective neural networks, it is expedient to evaluate the recognition accuracy by the accuracy of the fulfillment of the basic recognition conditions

$$S_{ij} < U_p \wedge S_{ii} \geq U_p ,$$

where S_{ij} ($i, j = 1,..., m$) – integral sums of input signals $\mathbf{x}_i = (x_{i1},..., x_{in})$, n - number of vector components $\mathbf{x}_i$. Integral sums S_{ij} are determined using the matrix $\boldsymbol{B}$

$$\boldsymbol{B} = \boldsymbol{A} \cdot \boldsymbol{A}^T ,$$

где $\boldsymbol{A}^T$ - matrix transposed to $\boldsymbol{A}$, $\boldsymbol{A}$ - matrix composed of vectors of input signals $\mathbf{x}_i$. The threshold value U_p is determined from the ratio $U_p = \min_i S_{ii}$.

Let us define a function M_0 - the number of erroneous recognitions as the number of condition failures $S_{ij} < U_p$. Then the recognition accuracy is determined from the relation

$$\delta = \frac{m^2 - M_0}{m^2} 100\% ,$$

where m^2 number of matrix elements $\boldsymbol{B}$.

You can give a visual geometric interpretation of the number of misidentifications. To do this, you need to build a graph of the function $Z = B(i, j)$ in three-dimensional space (i, j, Z). Incorrect recognition is characterized by off-diagonal matrix elements that will be larger than diagonal elements.

Let's give an example of evaluating the accuracy of a function $Z = B(i, j)$ when recognizing 10 digits (0, ..., 9) in a 4x6 field. Each digit corresponds to a binary string of 24 cells from 0 and 1. The totality of all strings is shown in Fig. on the right, in the form of a 10x24 matrix. Binary input signals characterizing objects are shown by rows of a 10x24 matrix. A geometric illustration of the selectivity of a single-layer perceptron based on selective neurons is shown in Fig. 9.

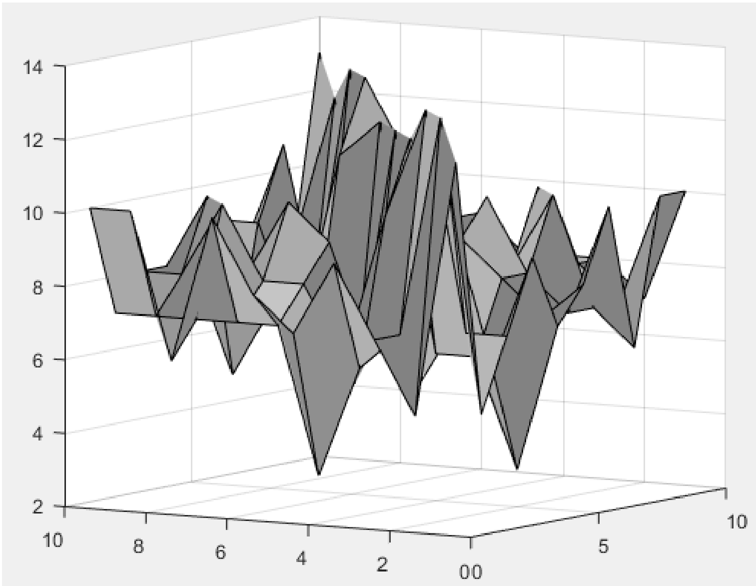

Fig. 9. Geometric illustration of single-layer perceptron selectivity based on selective neurons

Threshold value $U_p = 12$. As you can see from the graph of the function $B(i, j)$, all values $S_{ij} < U_p = 12$. Therefore, $M_0 = 0$ and recognition

6 Conclusion

The theoretical and mathematical substantiation of standard and selective neural network technologies is given. Mock-ups have been developed for the visual simulation of processes in standard neural networks based on McCallock-Pitts neurons and selective ones based on selective neurons. The neuro-educational system allows for effective training in the neurotechnology of senior schoolchildren, students, and specialists in related professions. Examples for teaching more complex recognition without mathematics in computer implementation are considered.

Acknowledgments

The article was prepared with the support of the Russian Foundation for Basic Research, grants No. 18-07-00918, 19-07-01137 and 20-07-00926.

References

1. Alexandrov Yu.I., Anokhin K.V., Sokolov E.N., Grechenko T.N. et al. Neuron. Signal processing. Plastic. Modeling. Fundamental leadership. Publishing House of the Tyumen State University. 2008. 548 p.
2. Alyautdinov M. A., Galushkin A. I., Kazantsev P. A., Ostapenko G. P. Neurocomputers: from software to hardware implementation. M.: Hot line - Telecom. 2008 . 152 p.
3. Galushkin A.I. Neural networks. Fundamentals of the theory. M., Hotline - Telecom, 2010.
4. Grechenko T.N. Psychophysiology. M.: Gardariki, 2009. 572 p.
5. Yezhov A.A., Shumsky S.A. Neurocomputing and its applications in economics and business. M., MEPhI, 1998.
6. Mazurov M.E. Neuron, modeling the properties of a real neuron. Patent for invention No. 2597495. 11/07/2014.
7. Mazurov M.E. A single-layer perceptron based on selective neurons. Patent for invention No. 2597497 01/13/2015.
8. Mazurov, ME.: Intelligent Recognition of Electrocardiograms Using Selective Neuron Networks and Deep Learning. International Conference of Artificial Intelligence, Medical Engineering, Education. Moscow, Russia. 2017. P.182-198
9. Mazurov M.E. Synchronization of relaxation self-oscillating systems, synchronization in neural networks. Izv. RAS. Ser. physical 2018.V. 82. No. 1. 83-87 P.
10. Mazurov M.E. Non-linear dynamics, almost periodic summation, self-oscillating processes, information coding in selective pulsed neural networks. Izv. RAS. Physical Series, 2018, V. 82, No. 11. 1564-1570 P.
11. Mazurov, ME .: Modeling of Intellect with the use of Complex Conditional Reflexes and Selective Neural Network Technologies. International Conference of Artificial Intelligence, Medical Engineering, Education. Moscow. 2018. 365-376 P.
12. Mayer R.V. Practical electronics: from a transistor to a cybernetic system. Glazov: GSPI, 2010.
13. Khaikin S. Neural networks: a full course. 2nd ed. M., "Williams", 2006.
14. Abimbola R. Iyanda, Olufemi D. Ninan, Anuoluwapo O. Ajayi, Ogochukwu G. Anyabolu. Predicting Student Academic Performance in Computer Science Courses: A Comparison of Neural Network Models. IJMECS 2018. Vol. 10, No. 6, Jun. PP.1-9, DOI: 10.5815/ijmecs.2018.06.01
15. Afnan A. Alkhathlan, Ahmad A. Al-Daraiseh. An Analytical Study of the Use of Social Networks for Collaborative Learning in Higher Education. IJMECS 2017. Vol. 9, No. 2, Feb. PP.1-13. DOI: 10.5815/ijmecs.2017.02.01
16. M. Moshref, R. Al-Sayyad. Developing Ontology Approach Using Software Tool to Improve Data Visualization (Case Study: Computer Network). IJMECS, 2019,Vol. 11, No. 4, Apr. PP.32-39. DOI: 10.5815/ijmecs.2019.04.04

17. C. S. Cheah, L. Leong. Investigating the Redundancy Effect in the Learning of C++ Computer Programming Using Screencasting. IJMECS, 2019, Vol. 11, No. 6, PP. 19-25. DOI: 10.5815/ijmecs.2019.06.03
18. Nurafifah Mohammad Suhaimi, Shuzlina Abdul-Rahman, Sofianita Mutalib, Nurzeatul Hamimah Abdul Hamid, Ariff Md Ab Malik. Review on Predicting Students' Graduation Time Using Machine Learning Algorithms. IJMECS. 2019. Vol. 11, No. 7, Jul. PP.1-13. DOI: 10.5815/ijmecs.2019.07.01.
19. L. E. Alvarez-Dionisi, M. M. R. Balza. Teaching Artificial Intelligence and Robotics to Undergraduate Systems Engineering Students. IJMECS 2019. Vol. 11, No. 7, Jul. PP. 54-63. DOI: 10.5815/ijmecs.2019.07.06
20. Salako E. Adekunle, Adewale Olumide S., Boyinbode Olutayo K. Appraisal on Perceived Multimedia Technologies as Modern Pedagogical Tools for Strategic Improvement on Teaching and Learning. *IJMECS* 2019. Vol. 11, No. 8, Aug. PP.15-26. DOI: 10.5815/ijmecs.2019.08.02