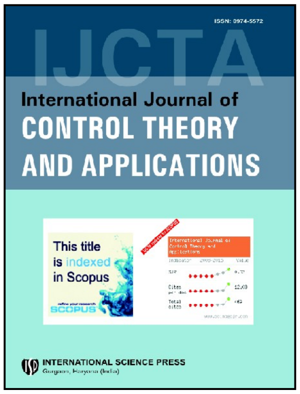

International Journal of Control Theory and Applications

ISSN : 0974-5572

Volume 10 • Number 11 • 2017

Neural Networks as an Opportunity to Create an Artificial Intelligence

Danilina Marina Viktorovna[1], Barannikov Alexandr Lukianovich[2], Ivanova Svetlana Petrovna[3], Sadikova Ksenia Valerievna[3], Fatianova Irina Rudolfovna[3], Sidorova Vera Nikolaevna[3] and Sidorov Nikolai Vladimirovich[3]

[1]*REU G.V. Plekhanov, Finance University*
[2]*Assistant professor, PhD, Moscow Aviation Institute (MAI), Russian Presidential Academy of National Economy and Public Administration (RANEPA)*
[3]*Assistant professor, PhD, REU G.V. Plekhanov*

Abstract: the article is devoted to the analysis of the possibility of the creation of the neural networks as an element of the artificial intelligence, the current state and the perspectives of the development of the future researches in this sphere.

Key words: neuron, neural networks, artificial intelligence, development, mathematical model.

The term “neural network” has appeared in the middle of the XX century. The first work in which the main results were obtained in this direction were made by McCulloch and Pitts. In 1943, their computer model of neural network based on mathematical algorithms and theory of brain activity has been developed. They hypothesized that neurons can be simplistically viewed as devices that operate on binary numbers, and termed this model “threshold logic.”

Like their biological prototype, neurons of McCulloch-Pitts were capable to learn by adjusting the parameters describing the synaptic conductance. Researchers have proposed the construction of a network of electronic neurons and showed that such a network can do almost anything imaginable numeric or logical operations. McCulloch and Pitts have suggested that such a network is also able to learn, recognize patterns, generalize, it means that such network has all the features of the intelligence.

This model laid the foundation of two different neural networks research approaches. One approach has been focused strictly on the study of biological processes in the brain, the other - on the use of neural networks as a method of artificial intelligence for various applications.

Neural networks are adaptive systems for data processing and analysis, which is a mathematical structure that mimics some aspects of the human brain and demonstrate such features as the ability to non-formal learning, the ability to generalize and clustering unclassified information, the ability to build their own forecasts on the basis of the already presented time series. Their main difference from other methods, such as expert systems, is that the neural network, in principle, does not require pre-known model and builds it itself only on the basis of the presented information.

That is why the neural networks and genetic algorithms come into practice everywhere, where you need to solve the problem of prediction, classification, management - in other words, in the area of human activity where there areproblems which need a lot of work to solve them using algorithmic methods and whose solution requires a permanent work of a group of qualified experts, or adaptive systems automation, which are the neural networks.

According to the scheme of the formal neuron of U.S. Mak-Kalloka and U.Pittsa (1943) [1], the neuron is a threshold element (Figure 1). On the inputs of the neuron there are excitatory and inhibitory synapses, the weighted sum of the neuron is determined (based on the weights of synapses) by the input signals, the output signal is produced in the excess of the sum of the neuron threshold.

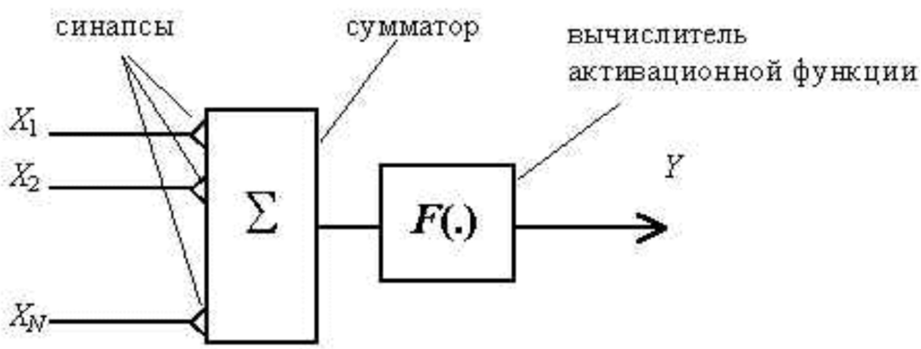

Figure 1: Diagram of the formal neuron. Includes synapses, sum and activation function's calculator, Xi - inputs, Y - the output of the neuron

The functioning of the formal neuron (Figure 1) can be described by the equations:

$$Y_j = F(net_j - K_j), \quad (1)$$

$$net_j = ?_i w_{ji} X_i, \quad (2)$$

where j - number of neurons in the network, Xi - inputs, Yj - the output neuron, wji - weight of synapses, netj - the total effect on the neuron input, Kj - the threshold neuron, F (.) - activation function.

The activation function of the response of the neuron to the input action *netj*, which can be a threshold:

$$F(a) = \begin{cases} 0\ \Pi p И\ a < 0, \\ 1\ \Pi p И\ \mathrm{a} > 0, \end{cases} \quad (3)$$

or some continuous, for example, linear:

$$F(a) = ka \quad (3a)$$

or logistic:

$$F(a) = 1/[1+exp(-a)]. \quad (3b)$$

Depending on the algorithm implemented on the permissible values of the inputs and outputs of the neuron, it imposes the certain restrictions: the values of Xi and Yj can be binary (it means equal to 0 or 1), bipolar binary (+1 or -1), belonging to the interval (0,1), or non-negative real. Similar restrictions are imposed on the weight of neuronal synapses wij.

Let's note that in the paper of McCulloch and Pitts [1] inputs and outputs are assumed to be binary neurons, synapses weight areconsidered binary bipolar and activation function is considered threshold. Research of neural networks [1] carried out in terms of the analysis of the logical calculus, which can be built on the basis of formal neurons. In particular it was shown that "for every logical expression that satisfies certain conditions, you can find a network, having described the behavior of this expression" [1].

The formal neurons to a certain extent reflect the dynamics of the transmission signals in the real biological neurons. Live neurons consist of the cell body, dendrites and axon. In a simple presentation of this problem, the work of the neuron can be described as follows. The dendrites receive signals from other cells through synapses, these signals are supplied to the cell body, where they are summed with other such signals.

If the sum signal for a short period of time is sufficiently large, the cell is excited, generating pulse in the axon, which is transmitted to the next cell. Without going into details, it can be noted that the formal neurons can only very roughly reflect the work of biological living nerve cells.

During the time different models were created. Among them are multilayer unidirectional networks and fully connected Hopfield network. Multilayer unidirectional networks are also called direct distribution networks, or multilayer perceptrons. In the future (where it can not cause different understanding) we will call such a multilayer network. The networks of such type include several types of neurons: neurons of an input layer, neurons of an output layer and a few neurons of "hidden layers". Fig. 2 shows a network in which there are K layers. The neurons of each layer are not connected. The output from each neuron is supplied to the inputs of all neurons in the next layer. The input layer neurons do not carry out the conversion of the input signals, their function is to distribute these signals between the neurons of the first hidden layer. The functioning of the multilayer network is as follows: the input signal is supplied to the network and then is supplied to the input layer neurons which go in its turn through all the layers and then is removed from the output neurons of the output layer [8].

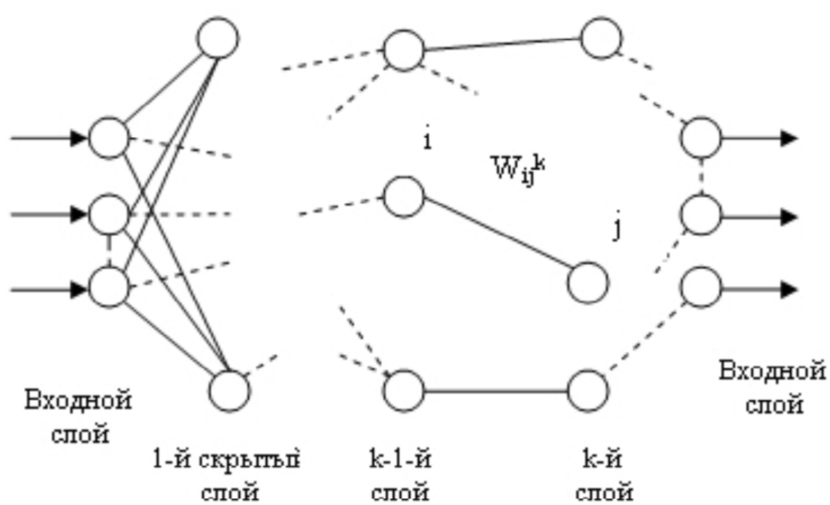

Figure 2: Diagram of the multilayer unidirectional network (includes neurons of an input layer, neurons of an output layer and a few neurons of "hidden layers")

As during the signal propagation over the network it undergoes a series of changes that depend on the initial values of variables and converting functions and weights. Let a chain consists of K layers: one input, and one output and (K> -2) hidden layers, - each layer is composed of n (k) neurons. Also we have a set of output signals of the neurons of the layer k (k = 1: K). Next let's call w^k the set of weights of synaptic connections, linking neurons k - 1-th layer with the neurons of k-th layer; w_{ij}^{k} - is weight connecting i - th neuron of k -1-th layer with j-th neuron of the k-th layer $\left(y_j^{k-1} \in y^{k-1}, y_j^k \in y^k\right)$.

Then the direct operation of the network can be described with the following correlations [8]: $y^{1} = x$

$y_j^k = f\left(\sum_{i-1}^{n(k-1)} w_{ij}^k\, y_i^{k-1}\right.$, $j = 1 : n(k)$, ГДе y_j^k — output of a neural network.

At the heart of multilayer neural networks learning methods most frequently is the so-called Delta rule. Delta rule is implemented as follows: $w_{ij}\,(t = 1) = w_{ij}\,(t) = hx_i\,(d_j - y_j)$,

where h — parameter (step of the study);

d - the reference (desired) value of the output of an element.

Thus, a change of the power of connections occurs in according to the output signal error of 5 = (d - y) and the level of activity of the input element x. The generalization of the Delta rule, called the back propagation (Back - propagation), is applicable to the networks with any number of layers. Education in the network in this case consists of the following steps:

1. Select another tutorial pair (x, d). Post an input vector to the input of the network.
2. Calculate the network output y.
3. Calculate the difference between the network output and the desired output (error).
4. Adjust the network weights so as to minimize the error.
5. Repeat steps 1 to 4 for each training pair, until the error reaches an acceptable level.

The error of the functioning of the network is usually determined as follows:

$$E = \frac{1}{2}\sum_{j-1}^{n(K)} (d_j - y_j)^2,$$

where $y_j =$ y_j^k — is the exit from the network. In order to reduce this error we need to change the weights in the network using the following rule:

$$w^k(t=1) = w^k(t=1) - h.\frac{\partial E}{\partial w^k}.$$

This formula describes the process of the gradient descent in space of scales. Obviously, for the output layer we have the following correlation:

$$(k = K) - \frac{\partial E}{\partial w_{ij}^k} = (d_j - y_j).\, f_j.y_i^{k-1}.$$

As

$$y_l^{k+1} = f_l\left(\sum_{j-1}^{n(K)} w_{jl}^{k+1} y_j^k\right), a\ y_j^k = f_j\left(\Sigma\, w_{ij}^k y_i^{k-1}\right),$$

then for the intermediate or hidden layers we have the following correlation:

$$k = 2:(K-1)) - \frac{\partial E}{\partial w_{ij}^{k}} = \sum_{l-1}^{n(k-1)} \left[(d_l - y_l)\, f_l' w_{jl}^{k+1} \right] f_j' \, y_j^{k-1}$$

If as the non-linear transfer function is used a sigmoid function, it is convenient to use the following recurrent formulas:

$\delta_j^k = (d_j - y_j) y_j (1 - y_j)(j = 1 : n(k))\ \ npu\, k = K$ - for output layer;

$\delta_j^k = \sum_{l=1}^{n(k+1)} \left[\delta_l^{k+1} w_{jl}^{k+1} \right] y_j^k (1 - y_j^k)(k = (K-1){:}2)$ – for hidden layers,

then $-\frac{\partial E}{\partial w_{ij}^{k}} = \delta_j^k y_i^{k-1}\ u\ w^k(t+1) = w^k(t) + h\delta_j^k y_i^{k-1}$.

These correlations are called the back propagation. If the direct operation of the input signal is distributed over the network from the input layer to the output, then the tuning weights network error propagates from the output layer to the input. Application of multilayer neural networks is due to the fact that they approximate the mapping (display) $F{:}D \subset R^{n(1)} \to R^{n(K)}$, using the preliminary training on the training data set... (x_1, d_1), (x_2, d_2), ..., (x_L, d_L), where d_e = F (x_e). Thus, the network can be considered as a model y = φ (x) of a real object y = F (x). Are proved the theorems that using network with reverse error propagation we can approximate any function with any accuracy [8].

Another model is the Fully connected Hopfield network. Hopfield Network (presented in Figure 3.) is a single-layer network. All neurons are connected to each other by bonds wij, where the output of the neuron signal can be given to it as its input and optionally wij = wji.

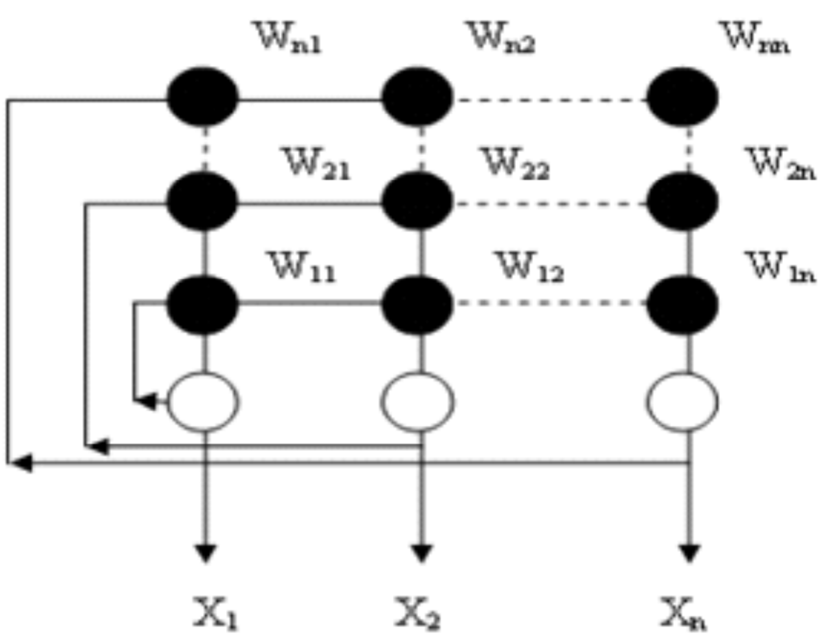

Figure 3: Diagram of the Fully connected Hopfield network

Since the signal outputted from each neuron is fed to the inputs of all the others, the input vector begins to circulate, transformed by the network as long as the network does not come to a stable state (i.e., when all neurons in each subsequent cycle will produce the same signal as in previous). It is obvious, that there may be cases of infinite circulation of the input vector without coming to a steady state. Let's choose a function of elements in the form of [8]:

$$x_j(t+1)=\begin{cases}1, \sum_{i-1}^{n} w_{ij}x_i(t) > \theta_j, \\ 0(-1), \sum_{i-1}^{n} w_{ij}x_i(t) < \theta_j, \\ x_j(t), \sum_{i-1}^{n} w_{ij}x_i(t) = \theta_j, \end{cases}$$

Network Status – is the current set of values of x signals from all neurons. The operation of the network can be represented geometrically as the motion of the vector x, which characterizes the state of the network, on the cube $[0,1]^n$. When is given the new input vector, the network moves from the top to the top, until it stabilizes. Sustainable vertex (top) is determined by the network's weights, the current input and the value of the threshold. If the input vector is partially incorrect or incomplete, then the network will stabilize in the top closest to the desired [8].

In general, all the possible state of the network forms a kind of a hilled surface, and the current states of the network are similar to the way of a heavy ball, which was placed on this surface - it moves down the slope to the nearest local minimum. Each point of the surface corresponds to a combination of active neurons in the network, and the height of the surface rise at this point characterizes the "power" of the state. The energy of this combination of activities is defined as the sum of the weights of connections between pairs of active neurons, taken with a minus sign (for 0 = 0) [8].

Thus, if the correlation between any two neurons has a large positive weight, the combination in which these neurons are active, has a low level of energy - it is to the such combinations that will seek the entire network. In contrast, neurons with a negative correlation with the activation of the energy add to the energy of the network a large amount, so that the network seeks to avoid such conditions. Dynamics of Hopfield network is convenient to describe the so-called energy function, which is in a fairly general terms it can be defined as [8]:

$$E=-\frac{1}{2}\sum_{i-1}^{n}\sum_{j-1}^{n} w_{ij}x_i x_j + \sum_{j-1}^{n}\theta_j x_j .$$

The energy function determines the stability of the network, in other words - it is the Lyapunov function of the Hopfield network, a function that always decreases when the network is changing conditions. In the end, this function must attain a minimum and stop the change, thus ensuring the sustainability of the network.

Changing the status of a network element always causes a decrease in network power. Indeed, suppose that changed state element k (k = 1: n), that is, its status changed from +1 to 0 (or -1), or vice versa, then:

$$\Delta E=-\frac{1}{2}\sum_{i-1}^{n}\sum_{j-1}^{n} w_{ij}x_i(t+1)x_j(t+1)+\sum_{j-1}^{n}\theta_j x_j(t+1)+\frac{1}{2}\sum_{i-1}^{n}\sum_{j-1}^{n} w_{ij}x_i(t)x_j(t)-\sum_{j-1}^{n}\theta_j x_j(t)\Rightarrow$$

$$\Rightarrow \Delta E_k=-x_k(t+1)\sum_{i-1}^{n} w_{ik}x_i(t+1)+\theta_k x_k(t+1)+x_k(t)\sum_{i-1}^{n} w_{ik}x_i(t)-\theta_k x_k(t)\Rightarrow$$

$$\Rightarrow \Delta E_k = (x_k(t) - x_k(t+1))\left(\sum_{i-1}^{n} w_{ik} x_i(t+1) - \theta_k\right)\{x_k(t+1) \equiv x_i(t)\}$$

It is evident that due to changes in k element Δ Ek ≤ 0, Δ Ek = 0, when there is no changes in the network. Due to this continuous striving for reduction of energy, in the end, it should reach a minimum and terminate the changing. By definition, such a network is stable. Hopfield networks are also called networks that minimize their energy. Hopfield networks have numerous applications. Some of them are connected with the ability of these networks to memorize and then recover even incomplete input information in different ways. Other applications relate to the possibility of using Hopfield network for solving optimization problems [8].

Another example of such type of models is Bidirectional associative memory. The main reason for the failure of researchers in the field of artificial intelligence who spent more than 20 years for unsuccessful attempts at intellectual simulation based on conventional digital computer is, apparently, the fact that in modern computers there is a direct dependency on the amount of search time for the stored samples. The computer stores the individual objects in the individual cells, as like if it learns the information by heart, and in the study of sciences - examples for it. Neuro computer that is based on neural networks, has associative memory and has received classifies images at with a speed, which does not depend on the number of the samples already arrived – it immediately connects the new image with the closest available [8]. Human memory is associative - one thing reminds us of another, and that, in turn, on the third, etc. Our thoughts are moving from object to object on a chain of mental associations... Ability of associations can be used to recover lost images ("we met with you somewhere "). Bidirectional associative memory is getero-associative; the input vector is applied to one set of neurons, and the corresponding output vector is generated at a different set of neurons. As Hopfield network, bidirectional associative memory is capable for the generalization, producing the right response, despite the possible distortion of the input [8]. Obviously, the state of the neurons can be considered as the short-term memory, as it can quickly change itself by the appearance of another input vector.

At the same time the values of the weighting coefficients of the matrix form the non-volatile memory (associations) and can vary over a longer period of time, using the appropriate training method. Learning is performed using the training set of pairs of vectors x and y. Let's assume that all the stored samples are binary vectors. Let's consider the solution of the problem with the help of bidirectional associative memory which can be divided into two phases: a training mode and just solution (recognition). Let s consider these two phases in the example. Each neuron a in the first layer A has synapses connecting it with L neurons in the second layer B. We suppose that neurons have the following "sense": a - currency, a2 - USD, a3 – marks, a4 - rubles, the b1 - USA, b2 - Russia, b3 - Canada, b4 - Germany. Let's consider the binary images of the learning mode [8].

We suppose that there are three binary neural connections in the network: (x1, y1), (x2, y2), (x3, y3).

Let

x 1 = (1,1,0,0) → y 1 = (1,1,1,0);

x 2 = (1,0,1,0) → y 2 = (0,1,0,1);

x 3 = (1,0,0,1) → y 3 = (0,1,0,0);

The meaning of learning correlations is obvious: if the excited neurons are a 1 and a 2 (we have dollars), then by the relevant synapses are excited neurons b 1, b 2, b 3 (we can use them in the United States, Russia and Canada), and so on. From the binary correlations we move to the bipolar ones (this is done solely for the simplicity, in order not to have to enter a non-zero threshold neurons):

x 1 = (1,1-1,-1) → y 1 = (1 ,1 -1,-1);

x 2 = (1,-11,-1) → y 2 = (1,-11,-1)

x 3 = (1,-1-1,1) → y 3 = (1,-1-1,1);

We form the matrix of weights:

$$W = \sum_{l-1}^{3} x_l^T y_l = \begin{pmatrix} -1 & 3 & -1 & -1 \\ 3 & -1 & 3 & -1 \\ -1 & -1 & -1 & 3 \\ -1 & -1 & -1 & -1 \end{pmatrix}$$

The recognition mode. We want to evaluate the effectiveness of the memory training links. We must make sure that the matrix W stores the correlations (x 1, y 1), (x 2, y 2), (x 3, y 3). We will provide the input 1 then x 1 = x (2, 2, 2, -2) - this means that in the layer B the first three neurons will be excited (threshold is taken equal to zero). Then, in the binary form y = (1, 1, 1,0), which is the required association. This means that the supply of the input x1, leads to y1, that is, that the computer can really "remember" the correlation (x 1, y 1). Similarly is checked the remembering of the remaining links [8].

The network is bi-directional: y 1 W T = (1, 5, -3, -3) → (1, 1, 0, 0) '! x 1, and so on. Let's determine the energy of correlations in the memory: similarly E (x2, y2) = - 4 and E (x3 y3) = -2. It is expected that the error in the initial data correlation (x 1, y 1) is to attract larger images as a point of stable equilibrium with minimal energy level. Indeed, we will provide the input image x '= (1,1,0,1) - a distorted image on one bit x 1 and x 3 then x' W = (1,1,1, -3) → (1,1,1, 0) → y 1. Similarly, if we take x '= (1, 1, 1) - vector located "between" x2 and x3, we obtain (-3, 1, 3, 1) → (0, 1, 0, 1) → y2 – correlation (x 2, y 2), attracts to itself as its energy is less than energy (x 3, y 3) [8].

Dealing with uncertain data. Let's consider the case where the currency type is undefined: x '- (1,0, 0,0), then x' W = (-1, 3, -1, -1) → (0, 1, 0, 0) '! y3. This means that it can be used only in the country where is used any currency. If the currency can be any, for example, dollars and marks, it can be used anywhere [8]:

x '-(1, 1, 1,0) → x 'M = (1, 1,1,1) → y'.

The research realized shows that the constructed neural network is able to remember the information you need during training and in the operating mode enables one to solve the problem of recognition, that is, implements the functions of associative memory. All information obtained during training is concentrated in the matrix W. Due to the parallel structure, the network solves the problem "instantly" - in one step - the multiplication of the input vector to the memory array [8]. Since the information is integrated into the matrix W, the network is able to solve the problem effectively and by partial distortion in the source data.

Now let's start with the explaining of the relevance of the creation of the artificial intelligence systems. First of all, it must be mentioned about the neuromorphic systems. In this sphere, there were developed several directions.

The first direction is the creation of interfaces "brain - computer" and neuroprosthesis. With the success of the creation of the specifications that are relevant mainly for people with limb injuries, disabled, or for patients paralyzed partially or completely. In the perspective the interface "brain - computer" can also be used to manage complex technical systems - such as robotic systems, aircraft and other large and / or distributed machines. Such an interface will be useful, for example, in the operation of a nuclear power plant, which has to monitor the large number of sensors.

The second trend is the development of the so-called "neuroanimates", which presents the functional connection attempt between the specially trained cultural living nerve cells and the robotic devices or other machine. Cultures were grown from stem cells –the predecessors of the neurons and the already adult neurons in a specific nutrient environment in the laboratory dish.

These neurons are connected to the electrodes, which read the signals from the cells or act on them by electrical pulses. So you can try to teach living neurons how to use the management of technical systems - robots and other machines, including virtual. This is a new trend, with still few number of practical results, but it has shown that the creation of such hybrid systems is possible, in principle, so there are prospects for such researches.

Third, the traditional direction, which is developing the 40-ies of XX century is the creation of artificial intelligence systems. It is a software and hardware implementation of artificial neuromorphic systems, similar in terms of decision-making and operations made on the interaction between a man and the higher animals technique, having developed nervous system, with the external environment.

Analyzing the future development of the neural networks, it can be presented the following possible directions in 2016-2018, presented by the Head of the laboratory of the neutral networks and education of the Center of the alive systems of Moscow physical technical institute Michael Burtzev [2]:

- recognition and classification system of objects in images;
- voice interaction interfaces for the Internet of things;
- quality monitoring services in call-centers;
- identifying the problems of the system (including predictive maintenance time), anomalies, cyber-physical threats;
- IP security and monitoring system;
- replacement of the bots of the operators of call centers;
- analytics system;
- self-learning system, optimizing the management of material flows or the location of objects (warehouses, transport);
- intelligent, self-learning process control systems and devices (including robotics);
- the appearance of "on the fly" universal translation systems for conferences and personal use;
- support the emergence of bots consultants or personal assistants, the functions close to the man.

It means that the neural networks are very popular at the current stage and the problem of the creation of the artificial intelligence remains of the great importance. At the same time neural networks are a form of machine learning, and the machine learning is a form of artificial intelligence. But the category of "artificial intelligence" is so poorly defined that this phrase is not practical. Yes, we can agree that it creates the images of technologically advanced future, but in reality we are still not close to the future, which creates our imagination. Once the optical character recognition was too complex for the machine, but now the phone application can scan documents and convert them to text. And we accepted it and nobody in the modern society calls it a great achievement. The reason that the basic phone features can be considered as artificial intelligence is that in fact there are two types of Artificial Intelligence. Weak or narrow Artificial Intelligence describes any system designed to perform a task from a narrow tasks list. For example, Google Assistant or Siri, being a quite powerful example of Artificial Intelligence, still perform a fairly narrow list of tasks. They receive voice commands and returns a response, or run applications. Research in the field of artificial intelligence nourish these functions, but they are considered "weak" [7].

In contrast, strong Artificial Intelligence - also known as total artificial intelligence, or "full Artificial Intelligence" – is a system capable of performing any human problem. And it does not exist. Therefore, any "smart" application is still an example of a weak artificial intelligence. And although the meaning can be quite vague, practical research in the field of artificial intelligence is so useful, that is probably already included in your daily life. Every time your phone will automatically remembers where you parked, detects faces in your photos, receives search offers or automatically includes all your pictures from the weekend in one group of files, you are somehow related to artificial intelligence. To some extent, “artificial intelligence” really just means that the application will be a little smarter than we are used to [7].

REFERENCES

[1] Мак-Каллок У.С., Питтс У. Логическое исчисление идей, относящихся к нервной активности // Автоматы, под ред. Шеннона К.Э. и Маккарти Дж. М.: ИЛ, 1956. С. 362 - 384.

[2] Дарья ХохловаБум нейросетей: Кто делает нейронные сети, зачем они нужны и сколько денег могут приносить, 12 июля 2016, https://vc.ru/p/neural-networks

[3] Е.В. Котельников, В.Ю. Колевато. Методы искусственного интеллекта в задачах обеспечения безопасности компьютерных сетей // Вятский государственный университет – Статья, 2010 – 32с.

[4] С. Н. Павлов. Системы искусственного интеллекта: Учебное пособие, 2011. — Ч. 1. — 176 с.

[5] Yasnitsky L.N. Diagnosis and Prognosis of Cardiovascular Diseases on the Basis of Neural Networks / L.N. Yasnitsky, K.V. Bogdanov, F.M. Cherepanov, T.V. Makurina, A.A. Dumler, S.V. Chugaynov, A.N. Poleschuk // Biomedical Engineering. - 2013. - Vol. 47, ¹ 3. - P. 160-163.

[6] Ясницкий Л.Н. Технология нейросетевого моделирования и обзор работ Пермской научной школы искусственного интеллекта / Л.Н. Ясницкий, К.В. Богданов, Ф.М. Черепанов // Фундаментальные исследования. - 2013. - № 1/3. - С. 736-740.

[7] И.Хель Нейронные сети, искусственный интеллект, машинное обучение: что это на самом деле? 24.10.2016, https://hi-news.ru/science/nejronnye-seti-iskusstvennyj-intellekt-mashinnoe-obuchenie-chto-eto-na-samom-dele.html

[8] Informational portal, *http://bourabai.ru/einf/chapter136.htm*