TEACHING METHODS FOR ARTIFICIAL NEURAL NETWORKS

https://doi.org/10.5281/zenodo.7818684

Abdusalomov Sherzodjon Inomjonovich
Ferghana State University

Abstract: *This article provides information about the research conducted in the study of artificial neural networks. The training methods and algorithms for artificial neural networks are also discussed.*

Keywords: *neural network, artificial neural networks, artificial intelligence, perceptron, genetic algorithms, initialization, madaline, algorithm, neurophysiological, deterministic method, biological neurons, formal neurons.*

Annotation: *This article provides information about the research conducted in the study of artificial neural networks and also, artificial neural network training methods and training algorithms are covered.*

Keywords: *neural network, artificial neural networks, artificial intelligence, perceptron, genetic algorithms, initialization, madaline, algorithm, neurophysiological, deterministic method, biological neurons, formal neurons.*

Introduction.

In the age where modern technologies are developing rapidly, artificial intelligence and its components are being applied in almost every aspect of life. This includes modern appliances used in everyday life such as televisions, refrigerators, washing machines, ovens, security systems, and smart home devices. Therefore, since ancient times, people have been trying to understand how their thinking works. In this regard, many neurologists, neuroanatomists, and scientists have conducted extensive research on how the brain works and have achieved significant results. They have gained a lot of knowledge about the structure and functions of the nervous system, but very little is known about how it actually works.

Studies have shown that the brain is incredibly complex, and each of the billions of neurons is connected to hundreds or thousands of other neurons. Despite the fact that supercomputers created by humans have achieved great results, they still fall short of the human brain's natural intelligence and complexity. This demonstrates the high level of complexity and perfection of human intelligence. Therefore, studying artificial neural networks is an important issue in the field of research.

In the field of studying neural networks and artificial intelligence, many scientists, researchers, and engineers have conducted extensive work. For example, the first step in studying neural networks was taken in 1943 when neurophysiologist Warren McCulloch and mathematician Walter Pitts published a paper on the model of neural networks using artificial neurons and electrical circuits.

In 1949, D. Xebb reported on the connection properties of neurons and their interactions, as well as proposed rules for teaching neural networks.

In 1957, F. Rozenblatt developed the formation and operation methods of perceptrons and proposed the first technical implementation of a neural computer.

In 1958, Jon Fon Neyman created a vacuum tube system that mimicked the simple functions of neurons.

In 1959, Bernard Widrow and Marcian Hoff developed the ADALINE (Multiple Adaptive Linear Elements) and MADALINE (Multiple Adaptive Linear Elements) models. The MADALINE was used as a filter to remove noise from telephone lines, and the model is still used today.

Also in that year, neurologist Frank Rosenblatt began work on the perceptron model. His single-layer perceptron is now considered a classic neural network model. Unfortunately, a single-layer perceptron can only perform a limited set of tasks. In 1969, M. Minski and S. Papert published "Perceptrons," in which the limitations of perceptrons were proven.

Several Uzbek scientists have also worked in this field, including academics V.Q. Qobulov and S.S. G'ulomov, professors A.T. Shermuhamedov and D.A. Khalilov, and researchers Q. Rahimov and I. Tojimamatov, who have published scientific articles on the subject.

Research methodology

In describing this research, the scientific works, educational literature, and creations of scholars, researchers, and practitioners on the subject were systematically studied. Their conclusions and ideas were analyzed comparatively, and the information was reworked.

Imagination and results.

“The concept of "Artificial Neural Network" was introduced in the 1940s for the first time. Artificial Neural Network models the activity of the nervous system of humans and animals in an arithmetic logical manner. The official model of the neuron was developed in 1943. Such models were able to solve problems of considerable complexity. By combining official neurons into a network, it is possible to eliminate these difficulties. These systems have much wider capabilities: artificial neural networks can solve problems related to the "human activity" field in a traditional way. For example, making decisions based on recognizing patterns and even incomplete data.

In particular, neural networks are interesting for learning and memorizing information that reflects human thinking processes. Therefore, in the early works on learning neural networks, the term "artificial intellect" was often used. Recently, interest in artificial neural networks has grown rapidly. They have been adopted by specialists such as biologists. Artificial neural networks are actually models of natural nervous systems, so creating and studying such networks gives us the opportunity to learn a lot about how natural systems work. The theory of artificial neural networks emerged in the 1940s due to the latest achievements in biology, as artificial neurons are made up of elements that model the elementary functions of biological neurons. These elements are formed to correspond to the anatomy of the environment or may not correspond to it at all. Despite these similarities, artificial neural networks demonstrate remarkable features that are not found in natural neural networks. For example, an artificial neural network can change its behavior based on external factors. It can read the input signals provided to it and learn in a way that provides the desired output. After learning, the network does not respond to small changes in input signals. The ability to recognize images through vibration and distortion is very useful in solving image recognition problems. It should be noted that neural networks implement generalizations automatically due to their structure, rather than through specially written programs.

One interesting feature of neuron networks is that they are reliable: even if several elements are not working correctly or fail, the network can still produce the correct results, but with less accuracy. Certain types of neuron networks have the ability to create a vague image based on several input signals. For example, you can teach the network to recognize the letter "A" by presenting a sequence of broken images of the letter. After training, the network can create the letter "A" without any breaks, meaning it can create things that were never presented to it. However, it should be emphasized that artificial neuron networks are not perfect. They are not suitable for tasks that require precise mathematical calculations. Researchers have not yet come to a consensus on the definition of a neuron network. There are many variants in literature.

A neuron network is a system composed of many simple computational elements that work in parallel. The result of network operation depends on its structure, connection strength, and the type of computation performed by each element. A neuron network is a parallel distributed processor capable of independent processing of information from input signals. This works similarly to the way humans acquire knowledge during the learning process: the knowledge obtained is not stored in a particular element, but rather distributed throughout the entire network.

A neuron network is a system composed of many simple computational elements. The result of each element depends only on its internal state. All elements work independently of each other, without synchronization with other elements.

Artificial neuron networks are systems that can accept, store, and use knowledge. However, most researchers consider them to be systems composed of many simple processors, each with its own local memory. This memory is usually referred to as the processor state. Processors can exchange numerical information with each other. The result of each processor's operation depends only on its state and the input data it receives.

Before using a neuron network, a training procedure must be performed, during which the network's behavior is based on input data and each element's state is set to calculate the correct response.

Differences between neural network architecture and classical von Neumann architecture.

We can illustrate the following analogy: let's say there is a function. How is it obtained? Very simple: two numbers are multiplied, then one is added and the result is divided by two. This sequence of operations is the simplest program. However, there is another method to solve this problem. It is possible to create a graph of this function and then find a solution from the graph. For example, the image of a line can be given. It is clear that finding a function that describes this image is very difficult.

If we continue this analogy, then the process of learning a neural network has its own unique graph. That is, we inform about the collection of coordinates. Points are constructed from these coordinates, and then the closest points are connected with straight lines. Thus, a graph is obtained, which can be used to find any value for a given input. In this case, no calculations are required, the result is obtained from the graph.

However, there is one difficulty here. Through given points, an infinite number of curved lines can be drawn. Therefore, when we try to determine the value of an input later on, we may get countless answers. But this problem can be solved: firstly, the values of the given points are close, and secondly, there is a method to minimize the error.

This is the main advantage of the neural network architecture. Any task needs to be formalized in a traditional computer (the image of a letter needs to be transformed into a function). Along with this, if there is a small error in the initial data or even one of the statements is broken, the final result will also be incorrect.

Today, there are many neural models that differ in complexity of computation and similarity to living neurons. Let's take a look at the classic model called "formal neuron" (1-illustration).

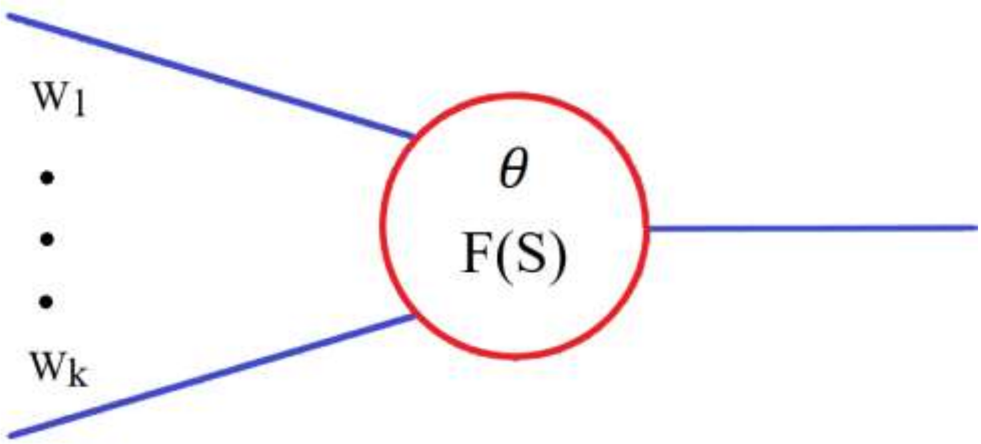

Figure 1. Formal neuron.

In a neuron, there are several input channels and only one output channel. Through the input channels, the neuron receives task information and the result is produced at the output. The neuron $W_1,\ldots,W_k$ calculates the weighted sum of input signals and then modifies the sum $F(S)$ using the given non-linear function.

Let's enter the following parameters:

X_i - the value of the input signal,

θ - the threshold level of the neuron,

W_i - the weight coefficient of the neuron (this value is usually referred to as weight, connection strength, or connection weight),

F - the activation function that performs the transformation,

y - the output value of the neuron. The threshold level and all weights make up the set of neuron parameters. Similarly, network parameters are the set of parameters of all its constituent neurons.

The output of the neuron is given by the following formula:

$$y = F\left(\sum (x, w_1) = \theta\right)$$

There is a modification of the formal neuron without a threshold. In this case, another input channel is added to the neuron (let its number be equal to) and for any input signal. Without a doubt, these models are equivalent.

$$\sum_{i\neq k}(X_i W_i + X_k W_k) = \sum X_i W_i = \theta$$

Formal neural modeling disadvantages.

It is estimated that a neuron calculates its output in one cycle, so it is not possible to model complex systems with such neurons that have internal states.

• Formal neurons, unlike biological neurons, cannot process information in a synchronized manner.

• There are no precise algorithms for selecting an activation function.

• It is not possible to order the operation of the entire network.

• For real neurons, the activity of the threshold neuron changes dynamically depending on the general state of the network, and the weight coefficients change according to the signals of the increase.

A single neuron can perform simple calculations, but the main functions of a neuron network are not performed by individual neurons, but are provided by their connections. A single-layer perceptron is a simple network consisting of a group of neurons that make up a layer. The input data is ranked with a vector of values, with each element of the vector given to the corresponding input of each neuron in the layer. The neurons compute the output independently from each other. The size of the input (i.e., the number of elements) should be the same as the number of synapses for each neuron and should be proportional to the size of the input signal. Despite its apparent simplicity, a single-layer perceptron can perform useful tasks, such as classifying images or calculating the values of logical functions.

A multi-layer perceptron is capable of calculating the output value for a given input value. In other words, the network calculates the value of some vector function: $Y = F(X)$.

Thus, the problem posed to the learner is to represent the problem in the form of a collection of vectors $(x1,...,x^S)$. The solution to the problem is in vector form, where . The thing that $\{y1,...,y^S\}$ a Perceptron can do is to generate a representation for $\forall s$ $y^S = F(x^S)$.

We cannot completely extract this mapping from the Perceptron, but we can only count the images of a certain number of points $F : X \to Y$ to $\forall x \in X$. Formalization, that is, determining the meaning that the compositional parts of input and output vectors have, is done based on practical experience by the person. Unfortunately, there are no strict formalization rules for neural networks.

To build a multilayer Perceptron, we need to select its parameters according to the following algorithm:

Determine how the components of the X input vector are related to the meaning. For example, the formalized condition for the input vector is that it should contain all the necessary information to get the answer.

Choose the Y output vector so that its components provide a complete solution to the problem.

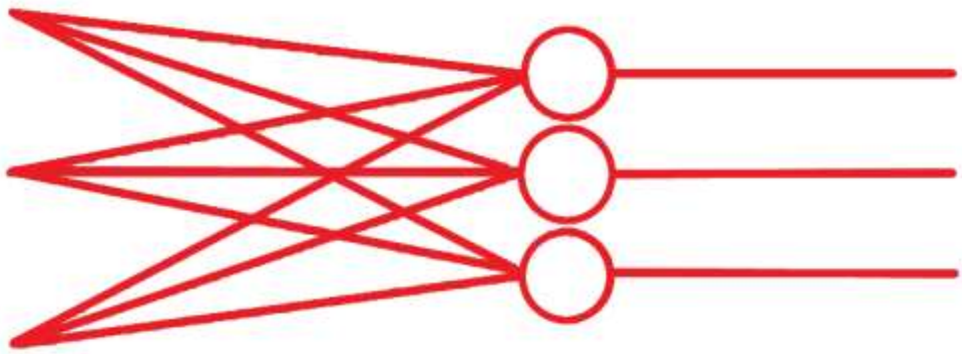

Figure 2. One-story Perceptron.

To create a multilayer perceptron, the following steps should be taken according to the algorithm:

Determine the type of activation function to be used for the neuron. In this case, it is necessary to take into account the specific characteristics of the problem, since good selection improves the speed of learning.

Choose the number of layers and neurons for each layer.

Define the range of input, output, weight, and threshold values based on the chosen activation function.

Set initial values for weights and thresholds. To prevent neurons from being saturated, initial values should not be too large. For most neurons to output nonzero values, initial threshold values should not be too small.

Conduct exercises to adjust network parameters for optimal problem-solving. At the end of training, the network can solve problems of this type.

Present problem conditions to the network in vector form. Calculate the output vector that gives the formalized solution to the problem.

Learning ability is the main characteristic of meaning. It is possible to see the network architecture, as well as adjustments of connections, for effective execution of the learning process in the context of artificial neural networks. Typically, adjusting the weights requires matching the heavy examples presented by the neural network. The learning ability of the network makes it more attractive compared to systems that work based on predefined patterns..

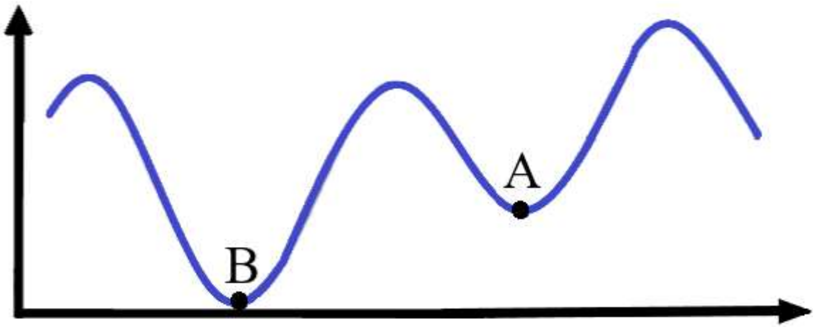

Figure 3. Local minimum problems.

It is possible to divide all teaching methods into two types: deterministic and stochastic. Deterministic methods determine the parameters of the system iteratively based on the current parameters, input values, and actual and required outputs. The clear description of this method is the backward propagation method. Stochastic learning methods randomly change the system parameters. In this case, only improvements are retained. The following algorithm can be given as an example of stochastic learning:

1.Choose system settings randomly. Provide input data and calculate the results.

2.Compare these outputs with the required outputs and calculate the difference between them. This difference is called error. The goal of training is to minimize this error.

.If the error is reduced, save the changes, otherwise discard the changes and choose new random settings. Repeat steps 2 and 3 until the system is fully trained. It should be emphasized that the stochastic learning method may converge to a local minimum (Figure 3).

For example, if the steps of random setting of the actual value are small, any deviations from the point will increase the error and be rejected. Thus, the smallest error value at the point will never be found. If the randomization of the system parameters is large, the error will change so sharply that it will never converge to any minimum. To solve these problems, the average size of random settings can be gradually reduced during the stages of random setting. If the step size is large, the error value will be accepted with equal probability for all values. If the step size is gradually reduced, the error value will eventually converge to a state where it remains stationary for a short period of time. If the step size is continuously reduced, it will eventually converge to a volume that is sufficient to reach the local minimum.

If during the learning process, the algorithm has correct answers (network output) for each input example, it is called a supervised learning algorithm under control. That is, a set of paired vectors $\{(x^S, d^S)\}$ is given in advance, $x \in X$ where the vector indicates the condition of the problem, $d^S \in \mathrm{Y}$ is the known solution to the problem for vecrtor r.

To give the network $X \to Y$ the necessary parameters for mapping during the learning process, it is necessary for the size $\{(x^S, d^S)\}$ of the dataset to be sufficient to form the learning algorithm.

Although the teaching method in practical problem solving is successful under supervision, many researchers criticize its teaching method for the artificial neural networks' biological basis. In reality, it is difficult to imagine a mechanism that compares actual results with the necessary results. The algorithm for learning without supervision can only be used when the input signals are known. Based on these, the network learns to provide the best output values. The concept of "best value" is determined by the learning algorithm. Usually, the algorithm adjusts the parameters to provide the same results for the same input values.

Habb's method is the oldest teaching principle of education. Based on physiological and psychological research, Habb proposed the assumption about how biological neurons learn. If both neurons are active at the same time and regularly, then the strength of the connection between them will increase. The important feature of this rule is that the weight change depends only on the activity of the neurons that are interconnected by the synapse. The algorithm itself looks like this:

1. At the initialization stage, small random values are given to all weight coefficients.
2. The input signal is applied to the network, and the cost of money and production is calculated.
3. The weight coefficients are changed based on the output values obtained by the neuron.
4. Starting from step 2, the input set is repeated with a new form until the network output becomes stable at a certain accuracy.

The error correction rule is a principle to correct errors. In 1957, Rosenblatt created a model that attracted great interest among researchers. The model used the learning algorithm under supervision, i.e., the learning set consisted of input vectors. An output vector was given for each of them. It became the basis for most complex learning algorithms under supervision today.

The essence of the algorithm is that the required output is specified for each input example. If the actual network output does not match the required one, then the network parameters are adjusted. To calculate the correction value, the difference between the actual and the desired network output is used. In addition, the weights are only adjusted if an incorrect answer is given.

Competitive learning. Unlike the Hobbian method, where many output neurons can work simultaneously, in competitive learning, output neurons compete with each other to be activated. This leads to only one output neuron from the entire set of output neurons to be active. This learning method has become the basis for most clustering algorithms today.

Summary and suggestions:

Artificial neural networks and their development issues have always been studied and criticized by various experts. However, they have been evolving rapidly with unique speed and enthusiasm. Some neural networks are developing much faster than representatives of certain fields, according to some people's opinions. Despite this, neural networks are already successfully used in control systems, identification of patterns (images), and daily tasks. The healthcare industry has also achieved great success in disease diagnosis and prevention, but optimal solutions have not been found yet for some issues. Comparative studies of various approaches (including neural networks) do not always lead to clear conclusions.

Therefore, it is necessary to fully understand the capabilities, necessary conditions, and scale of all available approaches, including neural networks, and maximize their advantages for the further development of intelligent systems. Such actions require the creation of entirely new algorithms that combine artificial neural networks with other technologies.

LITERATURE:

1. Gulamov, S.S., Shermukhamedov, A.T., & Khayitmatov, U.T. EXPERIENCE OF ARTIFICIAL INTELLIGENCE DEVELOPMENT IN CHINA.
2. Khalilov, D. (2022). MATHEMATICAL FOUNDATIONS OF ARTIFICIAL INTELLIGENCE AND RADIAL NEURAL NETWORKS. Science and innovation, 1(A6), 664-671.
3. Rakhimov, Q., & O'g'li Sotvoldiev, A.D. (2022). TRENDS IN THE USE OF MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE IN PRACTICAL FIELDS. YOUTH, SCIENCE, EDUCATION: TOPICAL ISSUES, ACHIEVEMENTS AND INNOVATIONS, 1(5), 85-91.
4. Vasenkov, D.V. (2007). METHODS OF TRAINING ARTIFICIAL NEURAL NETWORKS. Computer tools in education, (1), 20-29.

5. Tojimamatov, I.N. (2022). SOCIAL PROBLEMS OF THE SOCIAL NETWORK, 4(1), 702-705.

6. Tojimamatov, I. (2021). BIG DATA TECHNOLOGY IN DIGITAL ECONOMY, 420-430.

7. Mamasidiqova, I., Husanova, O., Madaminova, A., & Tojimamatov, I. (2023). DATA MINING TEXNALOGIYALARI METODLARI VA BOSQICHLARI HAMDA DATA SCIENCE JARAYONLAR. *Центральноазиатский журнал образования и инноваций*, *2*(3 Part 2), 18-21.

8. Tojimamatov, I. N., Mamalatipov, O. M., & Karimova, N. A. (2022). SUN'IY NEYRON TARMOQLARINI O 'QITISH USULLARI. *Oriental renaissance: Innovative, educational, natural and social sciences*, *2*(12), 191-203.

9. Nurmamatovich, T. I. (2021). RAQAMLI IQTISODIYOTNING GLOBALLASHUV JARAYONIDA IQTISOD TARMOQLARIDA QO'LLANILISHINING ASOSIY YO'NALISHLARI. *Н34 Наука и инновации в XXI веке: Материалы Международной*, 291.

10. Tuychievich, B. M., & Nurmamatovich, T. I. (2021). ЖАМИЯТДА РАҚАМЛИ ИҚТИСОДИЁТ. *Н34 Наука и инновации в XXI веке: Материалы Международной*, 189.

11. Kizi, A. Z. I., & Nurmamatovich, T. I. (2021). ZAMONAVIY DASTURLASH FANINI O'QITISHDA PYTHON DASTURLASH VOSITALARI YORDAMIDA AMALIY DASTURLAR YARATISHNING AHAMIYATI. *Н34 Наука и инновации в XXI веке: Материалы Международной*, 264.

12. Tojimamatov, I. N., Mamalatipov, O. M., & Karimova, N. A. (2022). SUN'IY NEYRON TARMOQLARINI O 'QITISH USULLARI.

13. Usmonov, B., Rakhimov, Q., & Akhmedov, A. (2023, March). The problem of takeoff and landing of a hereditarily deformable aircraft in a turbulent atmosphere. In AIP Conference Proceedings (Vol. 2612, No. 1, p. 060015). AIP Publishing LLC.

14. Усмонов, Б. Ш., & Рахимов, К. О. (2020). Построение математической модели в прямой и вариационной постановке задачи изгибно-крутильного колебания наследственно-деформируемого крыла самолета. Проблемы вычислительной и прикладной математики, (5), 108-119.

15. УСМОНОВ, Б., & РАХИМОВ, К. ПРОБЛЕМЫ ВЫЧИСЛИТЕЛЬНОЙ И ПРИКЛАДНОЙ МАТЕМАТИКИ. ПРОБЛЕМЫ ВЫЧИСЛИТЕЛЬНОЙ И ПРИКЛАДНОЙ МАТЕМАТИКИ Учредители: Научно-инновационный центр информационно-коммуникационных технологий, (4), 50-59.

16. Usmonov, B., & Rakhimov, Q. (2019). Vibration analysis of airfoil on hereditary deformable suspensions. In E3S Web of Conferences (Vol. 97, p. 06006). EDP Sciences.