А.А. Птухин, Е.М. Божко
Уральский федеральный университет имени первого Президента России Б.Н. Ельцина
Екатеринбург, Россия

Искусственные нейронные сети и их применение

В данной статье рассматриваются принципы работы, возможности и использование искусственных нейронных сетей. Приводится общая характеристика нейронных сетей, описываются принципы их функционирования, достоинства, недостатки и возможности применения в современном мире.

Artificial Neural Networks and Their Application

Nowadays the topic of machine learning, intelligent algorithms, and artificial intelligence is very popular because the artificial neural networks are actively developing and already used in many branches of science and spheres of life [3]. When solving tasks other than number crunching, a human brain possesses many advantages over a digital computer. Our brain is also remarkably reliable because it does not stop working just because a few cells die. But perhaps the most fascinating aspect of the brain is that it can learn without any necessary software [4].

The development of artificial neural networks (ANN) is inspired by biology. The researchers use terms borrowed from the principles of the brain organization while considering network configuration and algorithms [1]. Indeed, our brain is a complicated biological neural network consisting of a large number of neurons, which receives information from the senses and somehow processes it.

A biological neuron is an extremely complicated system. ANN are created by means of computer simulation of models of biological neural networks and represent a total of interacting components [1]. Thus, we can define the ANN as a set of interacting artificial neurons.

The gist of how neurons work lies in the synapses. A synapse is the place where two neurons are joined, and where the signal may be either amplified or reduced. Each synapse is unique. Moreover, synapses can change in the course of time and, therefore, the nature of the signal change will change as well. If the parameters of the synapses are properly selected, the input signal will be converted to the correct output after passing through the neural network. The number of neurons of the ANN and the speed of information processing is far less than in our brain, and even about 1,000 times less than in that of a fly [2]!

Using ANN for solving problems is feasible if there is no specific solution algorithm and for solving highly complicated problems. But the most interesting feature of neural networks is their ability to learn things that people do unconsciously (facial recognition, for example). Beside their ability to solve new problems, neural networks have a number of additional advantages such as *resistance to input data noise*. After being trained, they are able to ignore the noise data. Another advantage of neural networks is their *ability to adapt to changes*. Using data for learning, the ANN always adapt themselves to the environment. ANN are able to function properly even when they are seriously damaged, and this is the evidence of their *fault-tolerance*. One of the major advantages of neural networks is their *supreme processing speed* as each neuron can be represented as a small processor that receives and converts the signal.

Neural networks might seem to be ready to fully replace computers but it is not so. Neural networks have their drawbacks: for example, they are *not able to give precise and unambiguous answers*. The ANN are also *not able to solve the problem step by step* – they can do it only in one go. Thus their last drawback is the *inability to solve computational problems* [2].

To simplify the understanding of the operating principle of neurons and ANN the neurons can be represented in the form of circles and the inputs and outputs as the arrows aimed in the direction of traffic signal. Then the ANN will be represented as a set of circles connected with arrows.

In the biological neural network, the electrical signal is transmitted from inputs to outputs and it changes as it goes through the network. The signal magnitude is also changed in this process, and its value can be represented with a number. Therefore, the numbers served on the network inputs symbolize the electric signal, as if it actually were there. These numbers will change while passing through the network, and we shall obtain some resultant number at the output a response of the network. Remember that the synapses can amplify or reduce the signal passing through them. In the ANN each such relationship is characterized by a certain number called a link weight. The signal transmitted through this link is multiplied by the weight of the corresponding link. Thus, each of the arrows connecting the abovementioned circles corresponds to a number representing a link weight.

Let us consider the internal structure of an artificial neuron and the way it converts the incoming signal. The signals arriving to the input are multiplied by their weights, and then all the products pass on to the adder, which adds together all input signals multiplied by the corresponding weights. The result of it is a number called the weighted sum, which can be represented as an extent of general neuron excitation. The neuron should somehow process this weighted sum and generate the appropriate output signal. The activation function is used for these purposes. Different types of artificial neurons use different activation functions that take a weighted sum as an argument; for example, a unit step function whose value changes from 0 to 1 when passing a certain threshold. A sigmoid is also used quite often [2]. If the horizontal axis contains the values of the weighted sum and the vertical axis contains the values of the output signal, the graph of this function resembles a letter S. It is more flexible than the unit step function, and any number between 0 and 1 can be its resultant.

But how can the ANN neurons be positioned and connected with each other? Most neural networks usually have an input layer which task is only to distribute the input signals to the remaining neurons. The neurons of this layer do not make any calculations. Then in the single-layer neural networks, the signals from the input layer are immediately fed to the output layer, which makes the necessary calculations, and their results are fed directly to the outputs. Meanwhile, the multilayer neural network generally have hidden layers located between the input and output layers. The hidden layers also convert input signals to some intermediate results. Such structure of a neural network copies the multilayer structure of certain human brain divisions. Multilayer neural networks have a much greater capacity than single-layer ones. There are also feedforward ANN, where signal goes only in the direction from the input layer to the output one. In the feedback networks the signal can go in the opposite direction. In this case, the neuron output is determined not only by its input signal and weights but also by the previous outputs.

If you simply send a signal at the input of the ANN, it will make no sense. The network settings have to be changed until the input signal is converted into a desired output. However, the only thing to be changed is the weights of links.

Training a neural network implies searching for such a set of weight coefficients that converts the input signal after passing over the network to the desired output. Training samples i.e. a set of input signals to train a network are used for this purpose. And the test samples – the sets of input signals for evaluating network performance quality – are used to access the quality of network performance after training.

There are two approaches to the training of the network that lead to different outcomes: training with the teacher and without the teacher. When training with a teacher the signal is supplied to the input, and the network response is compared with the correct answer. Then the link weights are changed if needed and the process is repeated until the network answers differ minimally from the correct ones. A computer program usually acts as a teacher because it is often necessary to train the network for hours and days. When training without a teacher the network classifies the input data on its own, the correct output signals not being shown [2].

In fact, the study of neural networks began much earlier than it might seem. Artificial neural networks become available for the first time as early as in the 1940s due to the efforts of Warren McCulloch and Walter Pitts. Seeking to reproduce the functions of the human brain the researchers created a simple hardware (and later software) models of biological neuron system and its links. These ideas were later developed by Frank Rosenblatt, a neuroscientist, who proposed a device simulating the process of human perception and called him a "perceptron" (from the Latin word *perceptio* meaning – perception). In 1960, Rosenblatt introduced the first neurocomputer named "Mark-1", which was able to recognize some of the letters of the English alphabet [2].

The general arrangement of a perceptron is practically the same as it was described above but there are some differences. The input layer consists of the so-called sensors (S-elements), the hidden layer consists of the associative ones (A-elements), and an output layer is composed of the reacting (R-elements) elements. The S-elements may be in the state either 0 or 1. The S-A-links may have a weight of -1, 0 and 1. If the signal on the A-elements passes through a certain threshold, its condition changes to 1, otherwise it remains equal to 0. The A-R-links may be of any weight. If the signal on the R-element passes a certain threshold, its condition changes to 1, otherwise it remains equal to -1. The R-element determines the total perceptron output. There are also multilayer perceptrons, which have more than one A-elements layer.

The perceptrons are good at solving the problems of classification, and even a maximally simplified perceptron, where A-R-links may take only integer values, each A-element has only one S-element and there is only one R-element, can solve the simplest task on recognizing digits.

ANN are used in various fields of science: starting from the speech recognition systems to recognizing the protein secondary structure, classifying different types of cancer, and genetic engineering [3].

For example, in September 2016 the adventurous Japanese from Omron Corporation showed the world the robot named FORPHEUS. This robot was awarded the Guinness Book of Records certificate as the first artificial intelligence machine that not only can compete with a human in table tennis but also train its own students. Moreover, FORPHEUS is a self-learning machine that can learn from people and adapt to them, and such abilities would not be possible without the use of ANN!

At the same time, Google Inc. has updated its text translation service – Google Translate. Now neural networks are used for translation. The advantages of this method are that the network is able to learn on its own and the quality of translation will not remain the same and can be improved as people use the service.

The neural networks are now a part of our lives and you can find goals for them in every area.

Список литературы:

1. Основы теории нейронных сетей // НОУ ИНТУИТ URL: http://www.intuit.ru/studies/courses/88/88/info (дата обращения: 17.10.2016).

2. Учебник по нейронным сетям // Нейронные сети. Изучаем технологии будущего. URL: http://neuralnet.info/ (дата обращения: 17.10.2016).

3. Что такое искусственные нейронные сети? // Хабрахабр URL: https://habrahabr.ru/post/134998/ (дата обращения: 19.10.2016).

4. Krogh, Anders. What are artificial neural networks? // Nature Biotechnology. – 2008. – No. 26(2). – PP. 195-197 (дата обращения: 19.10.2016).