

# ARTIFICIAL INTELLIGENCE FOR COMPUTING AND ITS APPLICATIONS IN THE FIELD OF NATURAL RESOURCES AND ENVIRONMENT

## Lesson 1. Artificial Neural Networks

Dr. Nguyen Du Khang

### Summary:

*Artificial intelligence (TTNT) It is intelligence exhibited by any artificial system. In the current Industry 4.0 revolution, artificial intelligence is considered the core element for development. Artificial intelligence is divided into two schools of thought: traditional AI and computational AI. In this regard, the author presents in several installments some of the main methods of computational AI, such as artificial neural networks. neural network) Fuzzy systems, decision trees, etc. Unlike other related literature, the author will present the concepts in an easily understandable way for readers new to this field. The calculation formulas will be specifically proven, and finally, detailed algorithms will be provided for application to specific cases, especially for certain areas of the environmental resources industry, including surveying, mapping, and remote sensing.*

**Keywords:** *Neural networks, neurons, synapses.*

The development of artificial neural networks is inspired by biological neural networks. The idea is to build network structures and use algorithms based on the operating principles of the brain. However, achieving absolute similarity is impossible; our knowledge of the brain is limited, and we are unlikely to find the precise rules to control it.

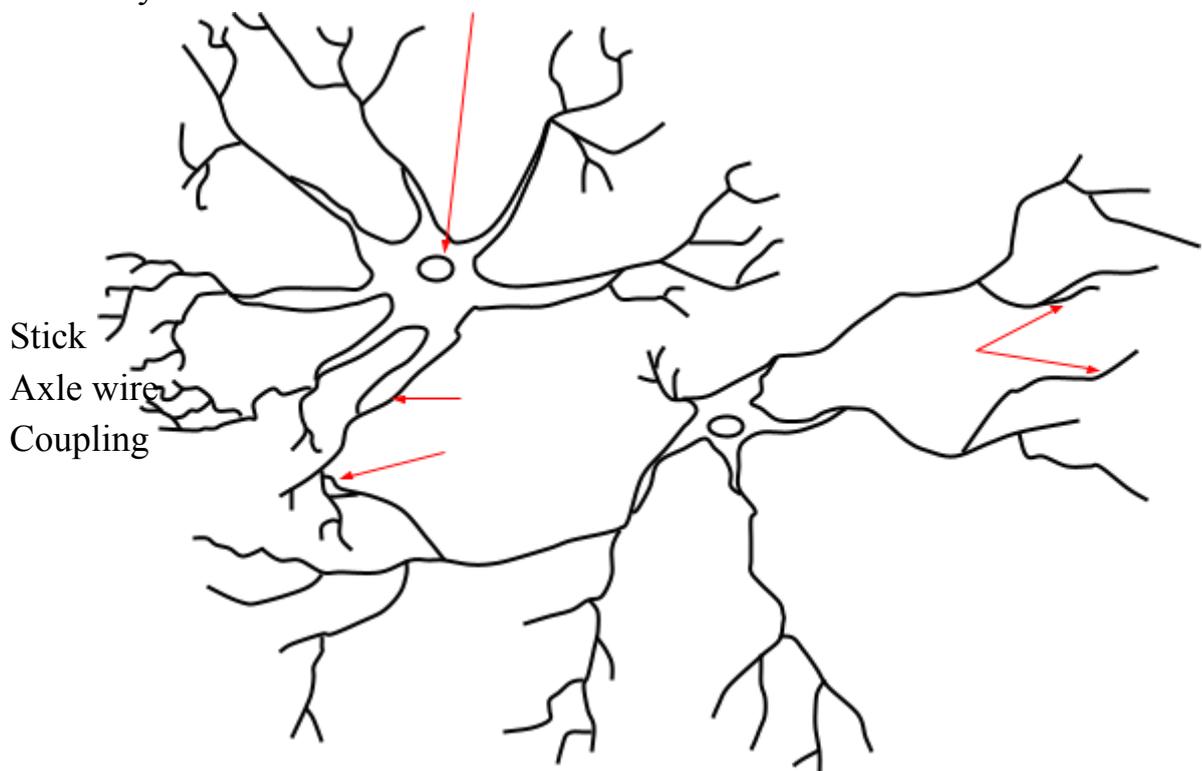
First, let's understand the basic concepts of how biological neural networks work. Next, we will discuss the basic concepts of artificial neural networks.

### 1. **Biological neural networks**

Biological neural networks, the highest level of which is the human nervous system, are composed of elements called nerve cells (neurons) with incredibly complex structures and functions. Approximately  $10^{11}$  Neurons are involved in nearly  $10^{15}$  The transmission connections can be over a meter long. Each neuron shares many characteristics with other parts of the body, but possesses entirely unique capabilities such as receiving, processing, and transmitting electrochemical signals along nerve fibers, forming a communication system within the brain.

Figure 1 shows the structure of a typical pair of biological neurons. Surrounding the cell body are axons, followed by branching nerves. This pair of neurons is connected at a point called a synapse. Input signals are received through the synapse and transmitted to the neuron body. Here, they are synthesized; in addition, some input signals tend to stimulate the neuron, while others inhibit its stimulation.

Cell body



*Figure 1. Structure of a pair of biological neurons.*

When the total stimulation within a neuron exceeds a certain threshold, the neuron is stimulated, sending signals along the axon to other neurons. This basic functional network has many complications and exceptions; therefore, most artificial neural networks only model the simplest properties.

## 2. Artificial neural networks

**Neural networks** A neural network is a series of neurons connected by synapses. The general model is shown in Figure 2, where the circles represent neurons and the lines represent synapses. Neural networks are capable not only of analyzing input information and producing output, but also of reproducing information that has passed through the network from their own memory. In other words, a neural network is a machine-based interpretation tool of the human brain.

**Neuron** (neuron) It is modeled after the properties of a biological neuron at the first approximation level. It is the computing unit, functioning to receive information, perform simple calculations, and transmit information further.

**Coupling** A synapse is the connection between two neurons. A synapse has a parameter called a weight, which changes the input information as it is transmitted from one neuron to another.

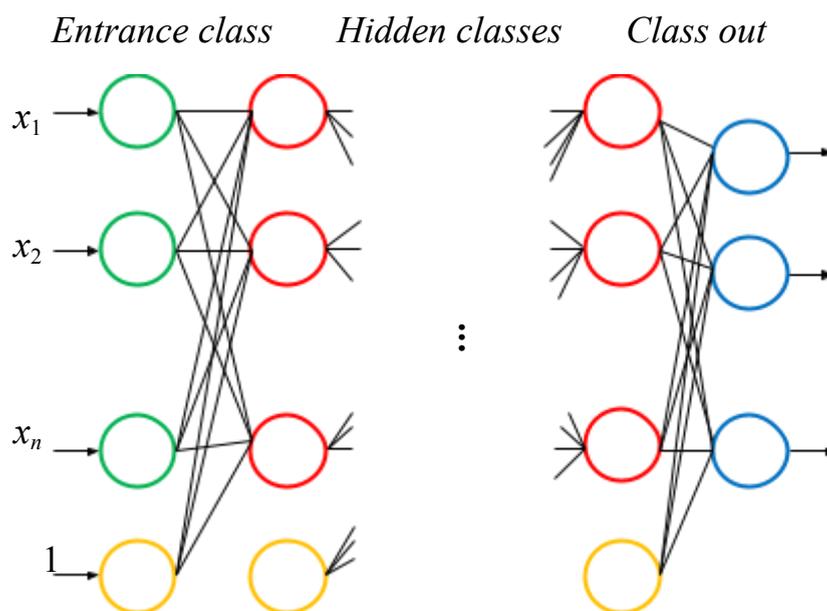


Figure 2. General model of a neural network.

A neural network consists of layers: an input layer that receives information (green), hidden layers that process information intermediately (red), and an output layer that processes the information and produces the final result (blue).

The neurons in the network correspond to layers with input neurons, hidden neurons, and output neurons. In most neural networks, transducer or bias neurons (yellow) are added to improve the signal separation capabilities of the layers.

For simple neural networks, there may be no hidden layers and/or translocated neurons.

Each neuron has input data and output data. For neurons in the input layer, input = output, meaning they do not process information. Transposition neurons, however, always have input and output signals equal to 1, and they do not have input coupling.

**How neurons work** In the network, there are no transduced neurons as described in Figure 3. The set of input signals is denoted by the vector  $X = (x_1, x_2, \dots, x_n)^T$ , in there,  $n$ - is the number of input signals. Each input signal  $x_i$  has a weight that  $In_i$  ( $i = 1 \div n$ ). The symbol for the weight vector is  $W = (In_1, In_2, \dots, In_n)$ .

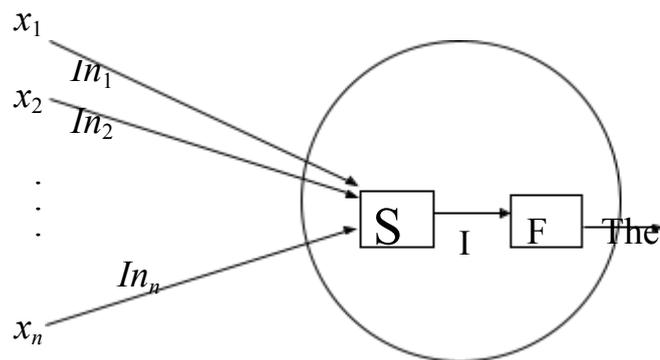


Figure 3. Mechanism of operation of artificial neurons.

The input signals of a neuron (excluding input neurons and translocation neurons) are determined by the following formula:

$$n$$

$$I = W.X = \sum_{i=1}^n x_i I n_i . \quad (1)$$

The output signal is determined by the formula:  $O = F(I)$ , (2) where  $F$  is the activation function, ensuring that the output  $O$  of the neurons has a value in the range  $[0, 1]$  or  $[-1, 1]$ .

There are many types of trigger functions, the most basic being step functions, sign functions, sigmoid functions, and hyperbolic tangent functions.

Step function:

$$F(x) = \text{step}(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \quad (3)$$

Sign function:

$$F(x) = \text{sign}(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ -1 & \text{if } x < 0 \end{cases} \quad (4)$$

The sigmoid function has the form (Figure 4,a):

$$F(x) = \text{sigmoid}(x) = 1/(1 + e^{-x}), \quad (5)$$

This is the most common trigger function, with a value range of  $[0, 1]$ .

The derivative of this function has the form:  $F'(x) = F(x).(1 - F(x))$ .

(6)

The hyperbolic tangent function has the form (Figure 4,b):

$$F(x) = \text{th}(x) = (e^{2x} - 1)/(e^{2x} + 1), \quad (7)$$

This function should only be used when the values in the neural network include both negative (-) and positive (+) numbers. The function has a value in the range  $[-1, 1]$ . If the neural network only contains positive numbers, this activation function should not be used, as it will weaken the network's performance.

The derivative of this function is of the form:  $F'(x) = 1 - F(x)^2$ .

(8)

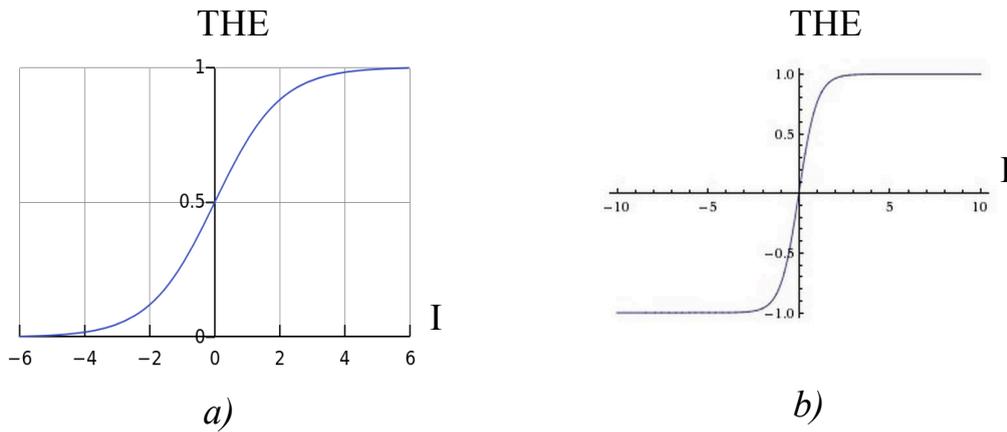


Figure 4. Types of activation functions. a) Sigmoid function; b) Hyperbolic tangent function

When transducer neurons are added to the network, the hidden and output neurons gain additional coupling, meaning additional weighting parameters are added. The expanded input vector takes the form:  $X' = (x_1, x_2, \dots, x_n, 1)^T$  and the expanded weight vector:  $W' = (In_1, In_2, \dots, In_n, In_{n+1})$ .

In that case, the input signals of the neurons (except for input neurons and translocation neurons) would be in the form:

$$I = W' \cdot X' = \sum_{i=1}^n x_i In_i + in_{n+1}. \quad (9)$$

From formula (9), it can be seen that adding transducer neurons aims to shift the graph of the activation function to the left or right. This allows for increased separability of input and output signals.

Thus, neural networks only function when there is a weight matrix of the connections (each layer has an input weight vector and an output weight vector). Determining the weight matrix requires training the network.

**Network training** This is the process of creating a sequence of data used to operate a neural network. The network is trained so that a set of input data, through processing, produces the desired output results. The set of input data, as well as the output results, are represented as vectors. Training is performed by sequentially feeding input vectors and simultaneously adjusting the corresponding weights with each processing step according to a specific

procedure. During training, the weight vectors are gradually adjusted so that each input vector produces the desired output vector.

Network training is divided into supervised training (learning with a teacher) and unsupervised training (learning without a teacher).

**Supervised training** This is done when, for each input vector, there exists an output vector that represents the desired target, and they form a pair. Often, the network is trained for several such pairs. Assuming the input vector is a given vector, calculate the network's output vector and compare it to the desired output vector to determine it. **error** If the error exceeds the allowable limit, the network needs to be retrained by adjusting its weights according to an algorithm that ensures the smallest possible error. This process is repeated until the error is within the allowable limit or the number of weights. **loop** Reached the limit.

An epoch is a count that is incremented each time the neural network is trained.

Error – a quantity reflecting the deviation between the obtained result and the expected result. It is determined after each iteration and should have a decreasing value. If this does not happen, it means there has been confusion in the network training process. Error can be calculated in many ways, of which three main forms are commonly used:

- Mean Squared Error (MSE):

$$\text{MSE} = \sum_{i=1}^m (O_i - T_i)^2 / m, \quad (10)$$

in doing,  $O_i$  and  $T_i$  – is the result of the neuron's calculations and expected outcome.  $m$  – is the number of neurons in the radiating layer.

- Root Mean Square Error (Root MSE):

$$\text{Root MSE} = \sqrt{\text{MSE}}. \quad (11)$$

- Mean squared arctan error:

$$\text{Arctan} = \sum_{i=1}^m \arctan^2(O_i - T_i) / m. \quad (12)$$

In principle, the error calculation is the same in all cases; however, the most commonly used is the MSE error because it maintains balance in the calculation. After each iteration, the error is determined. If the error exceeds the allowable limit, the network training process is repeated until the requirements are met.

***Unsupervised training*** This process occurs when each input vector produces a corresponding output vector, without needing to compare it to a predetermined target. Therefore, the input vector is the only data used to train the network. Training is done by adjusting the network's weights so that the output vector matches; that is, input vectors with sufficient similarity will produce the same output. Thus, based on the training, the neural network can group input vectors into classes.

### 3. **Functions of artificial neural networks**

Neural networks are used to solve complex problems requiring computation and analysis similar to those performed by the human brain. The basic functions of artificial neural networks are widely applied in many fields; in the field of natural resources and environment, these may include:

***Classify***– This involves classifying data according to parameters. For example, classifying land use, forests, areas with potential mineral resources, levels of environmental pollution, drought, saltwater intrusion, etc. This work can be done by an artificial neural network, analyzing information about existing data, current conditions, and phenomena on the surface of the study area.

***Forecast***– This refers to the ability to predict the next step. For example, weather forecasting based on signals such as time of year, temperature, humidity, wind speed, cloud height and thickness, and other information; forecasting landslides (on the coast, rivers, etc.) based on signals such as currents, wave height, soil and rock mechanics, and other information; forecasting landslides, flash floods, etc. inland; forecasting in agriculture (crop growth, pests and diseases, yield, etc.); forecasting water flow of rivers and lakes; etc.

***Identification*** Currently, artificial neural networks are the most widely applied technology. For example, they are used to identify characteristic microecological zones based on shape, size, position relative to the surrounding environment, and other information. Artificial neural network recognition is most commonly used in traditional artificial intelligence fields, such as handwriting, image, and speech recognition.

In subsequent articles, the author will introduce the core functions and parameters of artificial neural networks. Following that, detailed algorithms for training artificial neural networks will be presented, along with specific applications in several fields of the natural resources and environment sector.