Hierarchical approach to artificial intelligence

K Zhigalov$^{1,2}$, K Avetisyan$^3$, V Y Konyukhov$^4$ and A D Osmaev$^{2,5}$

$^1$V.A. Trapeznikov Institute of Control Sciences of the Russian Academy of Sciences, Moscow, Russia
$^2$Kh. Ibragimov Complex Institute of the Russian Academy of Sciences (CI RAS), Russia
$^3$Moscow University of the Ministry of Internal Affairs of the Russian Federation n.a. V. Ya. Kikotya, Moscow, Russia
$^4$Irkutsk National Research Technical University, 83, Lermontov str., Irkutsk, 664074, Russia
$^5$Chechen State University, 32a, Sheripov str., Grozny, Russia

E-mail: kshakalov@mail.ru

Abstract. There are three promising directions of artificial intelligence in the 21st century: focused solutions, i.e. solutions used only in a specific field; artificial intelligence of the general direction – a neural computer capable of performing any intellectual action typical for a man or a society and development of artificial superintelligence with scientific creativity, wisdom and communication skills. This paper is aimed at considering a hierarchical approach to the application of artificial intelligence. It is assumed that the hierarchical approach will allow more accurately choosing the required type of solution when dealing with any automation task.

1. Artificial intelligence. Introduction

The National Strategy for Artificial Intelligence Development until 2030 gives the following definition of AI: AI is a set of technological solutions that allow simulating cognitive functions of a person (including self-learning and search for solutions without a predetermined algorithm) and obtaining results comparable, at least, to the results of intellectual activity of a person when performing specific tasks.

Here, AI is considered in conjunction with its sub-areas. At the same time, there are many articles in the Internet devoted to AI, and their definition is slightly different, but this is because it is considered without machine learning and deep learning. So what is initially understood as the AI?

AI is some computer program implemented in some software, designed to automate any process, simplify the execution of monotonous routine tasks, in other words, it is a machine code created to facilitate the life and work of a person, or, perhaps, to replace it at all [1, 2]. In particular, AI can be used in the field of the earth remote sensing and decryption of remote sensing materials [3]. For this purpose, AI shall mimic our intelligent activities. AI can be implemented in a software code using such branches as if – else (if “it” – we do one action, otherwise – another). As a result, AI looks as an algorithm that describes the behavior of a person in solving a problem.

Before moving on to the definition of a neural network, from the point of view of a computer program, without touching on the historical issue of its development and from where the idea was
taken, it is necessary to understand the meaning of a neuron.

2. Neural network (NN) and its structure
In order to further understand what a neural network is, it is necessary to get acquainted with some basic elements of a neural network, and its basic (simplest) structure.

![Neural Networks Diagram](image)

**Figure 1.** Basic structure of neural networks

It may be said that the input cell represents the input data (Figure 2).

The output cell can show, for example, the probability of belonging to a relevant class when solving the classification problem, for example, dog – cat (Figure 3).

The hidden cell (a neuron itself) is something between the input and output of the neural network, what causes “all magic” (Figure 4).

![Input Cell](image)

**Figure 2.** Input cell of the neural network

![Output Cell](image)

**Figure 3.** Output cell of the neural network

![Hidden Cell](image)

**Figure 4.** Hidden cell of the neural network

The mathematical model of information perception by brain. The example of a single-layer perceptron. There is also a multi-layer perceptron. The very concept of a multi-layer perceptron is a neural network with direct propagation (Figure 5). However, the multi-layer neural network is not a multi-layer perceptron, since the neural network can be recurring and have feedback. In the case of perceptron, everything goes from the entrance to the exit in one direction.

The simplest example of a neural network (a basic neural network that includes 2 inputs, 2 hidden neurons (these 2 hidden neurons give 1 hidden layer) and 1 output), as well as an example of direct propagation, is called the Feed Forward. With this neural network, for example, a nonlinear XOR problem can be solved (Figure 6).

Figure 7 shows the Deep Feed Forward. As can be seen, this structure is similar to the basic neural network, but there are already 3 inputs, 8 hidden neurons, which form 2 hidden layers of 4 neurons and 2 outputs.
Neuron as a class that has input data, certain weights (they are also called synapses), activation function (there are quite a few of them, but the most often used is sigmoid and ReLU) and output.

Next, let us consider the development steps of each neuron element:

1. Input data – despite all the simplicity of this concept, they are not obvious. Before working with neural network, it is necessary to understand what to supply to it at the entrance, what will it classify, forecast since it will carry out a task and to clearly, visually display an exit for the end user. Given the fact that at the lowest level the computer works with the machine code, which is 0 or 1, and the neural network is a computer program, the input data will be 0 and 1.
When working with an image, the neural network does not work with the image that we see, but the image in grayscale, where 0 – black, 1 – white, and the grayscale ranges from 0 to 1, hence, the closer to the black the gray shade, the closer it is to 0 and vice versa. In addition to the range 0 and 1, the range from −1 to 1 may also occur. Depending on the input and on how they will be presented in a particular neural network, a different activation function is selected, for example, if the range is from 0 to 1, a conventional sigmoid or RELU can be chosen, if the range is from −1 to 1, a hyperbolic tangent or bipolar sigmoid is chosen.

2. Weight is the importance of a specific input given to a neuron when making a final decision, for example, when we want to go to a movie, some will choose a film only based on their interest in a given film, and they will not be concerned with the price since they are well off, and for others the price will be important.

3. The activation function normalizes the value of the neuron weight multiplication by the input data and makes it between 0 and 1. There are many activation functions (Figure 8, Figure 9, Figure 10 show the main functions used).

![Figure 9. Sigmoid activation function](image1)

![Figure 10. Hyperbolic tangent activation function](image2)

The sum of input value products per neuron weight (f(input*weight)) passed through the activation function. To obtain the output in the basic neural network, with the scale matrix already configured, direct propagation is used, in English it is called feed forward. Neural network is a set of neurons connected to each other by means of a matrix of weights and having a number of outputs, where each output means a class, for example, in the recognition problem it can be: a cat, a dog, a bird, a person.

Neural networks have existed for a long time (they were invented in the first half of the 20th century), but they reached their popularity in the first half of the 21st century.
The matter is that after the neural networks was discovered there was a problem of choosing a scale matrix between each neural. After the method of machine learning was invented, the neural networks were found to be in a new perspective, it became possible to teach them, i.e. to select this matrix of weights, for a certain problem, and accordingly to apply it. Besides, large data sets have not yet been collected for complex classification problems.

From a mathematical point of view, a neural network is a function with many parameters (weights), which play an important role in solving the problem for which this function was created. In addition to parameters, the neural network also has hyperparameters. Unlike parameters that are tuned using the chosen ML mathematical algorithm, hyperparameters are chosen manually, often by trial and error. Hyperparameters include the following:

1. Learning rate.
2. Neural network architecture.
3. Presence of displacement neurons.

The learning rate affects what minimum number of iterations, the parameters of the neural network will be adjusted so that the neural network can perform the task for which it was created. It is possible to set the static learning rate throughout the learning process. But to optimize, it can be reduced as the iteration increases.

The neural network architecture defines the general principles of its construction (plane layered, fully-connected structure, weakly connected, feedforward, recurrent, etc.).

The configuration specifies the network structure within the given architecture: the number of neurons, the number of inputs and outputs of the network, the activation functions.

There are the following basic architectures:

- feedforward networks – all connections are directed strictly from input to output neurons. Such networks include, for example, the Rosenblatt perceptron and multi-layered perceptron;
- recurrent neural networks – the signal from the output neurons or neurons of the hidden layer is partially transmitted back to the neurons of the input layer of the network;
- radial basis function networks – networks containing a single hidden layer of neurons, which use a radial-symmetrical activation function, are used to solve classification and forecast problems;
- Kohonen networks – a class of networks that use unsupervised learning to solve clustering
problems. They contain only two layers: input (distribution) and output (clustering);
  • Kohonen maps or self-organizing feature maps – a form of Kohonen networks in which the
    number of output neurons is selected to be much larger than the number of clusters formed. They are
    used to visualize multi-dimensional data clustering results;
  • fully connected networks – neural networks in which each neuron is connected to all other
    neurons. Such networks have the highest bond density;
  • weakly connected networks – neurons are connected only to their nearest neighbors;
  • plane-layered neural networks – neurons form cascades called layers, and the neurons of each
    layer are connected to all neurons of the next and previous layers, and inside the layer there are no
    connections. The plane-layered networks can be single-layer (containing one hidden layer) and multi-
    layer (containing several hidden layers).

Each architecture of a network is used to solve a certain class of data analysis problems (regression, classification, clustering, forecast), and utilizes special learning algorithms.

3. Machine Learning (ML) and Deep Learning (DL)

Machine learning models:
  • Linear regression;
  • Logistic regression;
  • Linear discriminant analysis;
  • Decision trees;
  • K-nearest neighbors;
  • Vector quantization networks;
  • Support vector method;
  • Bagging and random forest.

![Figure 12. Linear regression model](image1)

![Figure 13. Logistic regression model](image2)
Figure 14. Model of linear discriminant analysis

Figure 15. Vector quantization network model

Figure 16. Support vector model
Which machine learning is related to neural networks? Artificial neural networks (ANN) – a mathematical model of biological neural networks (BNN).

After the invention of neural networks, their problem was to form the correct scale matrix to solve a certain problem, for example, to classify cats and dogs. Accordingly, it should be noted that for each problem the weights are different. Therefore, there is such subarea of machine learning as Deep Learning.

4. Deep Learning (DL)
Deep learning is an advanced field of machine learning (ML). It represents a few hidden layers of artificial neural networks. The deep learning methodology applies nonlinear transformations and high-level model abstractions on large databases. It is the shift to to robotic execution in critical areas where the human factor can have negative consequences. It forms the maximum automated production, able to operate and rebuild without manual control.

5. Conclusion
Artificial intelligence (AI) allows solving simple problems through operators if – else, turn left if something, turn right – if something different. Neural networks (NN) solve nonlinear, more complex problems. The NN represent a complex function with many variable parameters called weights. There is Machine Learning (ML) and Deep Learning (DL) as weight matrix setup methods.

The aggregate of methods that allow simulating the behavior of a person in a narrow area together with the functions of adjustable variable (the number of variables depends on the number of input data, and the number of configurable parameters is usually selected depending on the problem) allows modeling different spheres of activity of the system of social relations.

Machine learning is any method that allows learning without being explicitly programmed to do so. Deep learning is a subset of machine learning that uses suitable multilayer neural networks.
There are three promising directions of artificial intelligence in the 21st century: focused (weak) artificial intelligence (FAI) – used only in a specific field; general (strong) artificial intelligence (GAI) – a neural computer capable of performing any intellectual action typical for a man or a society. Artificial superintelligence (ISI) with abilities for scientific creativity, wisdom and communication skills.

Besides, it is necessary to study the system by means of the system analysis, mathematical modeling and programming [4].

Various sciences, such as law, study the issues of artificial intelligence and consider the problems associated with the introduction of new trends and intelligent solutions today [5].

References
[1] Makhovikov A B, Katuntsov E V, Kosarev O V and Tsvetkov P S 2019 Digital transformation in oil and gas extraction In Innovation-Based Development of the Mineral Resources Sector: Challenges and Prospects 11th conf. of the Russian-German Raw Materials 2018 pp. 531–538
[2] Shutaleva A V, Kerimov A A and Tsiplakova Yu V 2019 Humanization of education in digital era Perspect. of Sci. and Ed. 42(6) 31–43 DOI: 10.32744/pse.2019.6.3
[3] Strizhenok AV, Ivanov AV and Suprun I K 2019 Methods of decoding of the geoeological conditions of natural-anthropogenic complexes based on the data of Earth remote sensing J. of Phys.: Conf. Ser. 1399(4) 044077
[4] Youn R B, Klyuev R V, Bosikov I I and Dzeranov B V 2017 The petroleum potential estimation of the north Caucasus and Kazakhstan territories with the help of the structural geodynamic prerequisites Sustainable Development of Mountain Territories 9(2) 172–178
[5] Vasilyev A S, Murzin D V 2020 Trends of application of subjective good faith in russian law
Approaches to Intellectual Propesitos y Representaciones 8(3) doi: http://dx.doi.org/10.20511/pyr2020.v8n3.516