UDC 004.89 DOI: 10.63978/3083-6476.2025.1.1.08

Yurii Lysetsyi
Doctor in Technical Sciences,
associate professor
professor of the department
Yevhenii Bereznyak Military Academy
Kyiv, Ukraine
e-mail: Yurii.Lysetskyi@snt.ua
ORCID: 0000-0002-5080-1856

THE FUNDAMENTALS AND COMPONENTS OF ARTIFICIAL INTELLIGENCE

Abstract*. The article is devoted to highlighting the basics and components of artificial intelligence. The science of artificial intelligence began to develop in the 40s of the last century, when a group of scientists from various fields came to the conclusion that it was possible to create an artificial brain. At that time, its theoretical and mathematical foundations and components began to take shape. Research in the field of neuroscience has shown that the brain is a network of neurons that exchange electrical signals on an all-or-nothing, 0 or 1 basis. The cybernetics of the American mathematician Norbert Wiener described the basics of control and stability in electrical networks. The information theory of the American electrical engineer and mathematician Claude Shannon made it possible to learn about digital signals. The English mathematician and cryptographer Alan Turing theorized the possibility of any computation using digital operations. The American neuropsychologist and neurophysiologist Warren Ma-Culloch and the neurolinguist and mathematician Walter Pitts analyzed networks consisting of idealized artificial neurons and demonstrated how they could perform the simplest logical functions. Canadian physiologist and neuropsychologist Donald Gebb proposed a theory of learning based on the conjuncture of neural networks and synapses that can strengthen or weaken over time. These scientists were the first to describe what researchers would later call a neural network. In 1957, Frank Rosenblatt proposed the Perceptron, a mathematical or computer model of the brain's perception of information (a cybernetic model of the brain), first implemented as the Mark-1 electronic machine in 1960. The perceptron was one of the first models of neural networks, which is essentially a mathematical model that emulates the human nervous system. Neural networks are considered: Feedforward, Recurrent, Convolutional, Long Short-Term Memory, Convolutional Recurrent Neural Networks, Generative Adversarial Networks, Hopfield Networks, Boltzmann Machines, Memory Networks, Generative Adversarial Networks. It is shown that neural networks are a modern powerful tool of artificial intelligence that can model and simulate the work of the human brain. And this is a practical mechanism that ensures the implementation of Machine Learning and Deep Learning technologies. Thus, artificial intelligence models based on neural networks are used to solve complex problems that cannot always be solved using mathematical formulas or rules. They are capable of learning from large amounts of data and recognizing complex dependencies in this data. They can be used in various technological products and industries, including cybersecurity.*

Keywords: *artificial intelligence, neural network, model, cybersecurity.*

Introduction

Statement of the problem. Artificial Intelligence (AI) as a field of study has been founded on seminar, held in Dartmouth College in the USA in 1956 [1]. This seminar was of great importance for science: it brought together scientists interested in the problem of modeling the human mind, confirmed the emergence of a new field of science and gave it the name – artificial intelligence. The organizers of the event were John McCarthy, Marvin Minsky, Claude Shannon and Nathaniel Rochester. They invited all the famous American researchers whose work was somehow connected with the issues of control theory, automata theory, neural networks, game theory and the study of intelligence.

Since the early 2000s, with the rapid development of information technologies, AI has begun to be used very actively and widely, in particular in information security, and eventually in cybersecurity. Therefore, in this context, highlighting the basics and components of AI is quite appropriate.

Analysis of recent research and publications. The issues of individual theoretical and mathematical components of AI are the subject of scientific works by both foreign and domestic researchers [2–9], however, with the emergence of opportunities for using AI technologies in the field of cybersecurity, their effective application requires a more systematic study of the foundations and components of AI.

Purpose of the article is to illuminate fundamentals and components of AI.

Presentation of the main material

The science of artificial intelligence (AI) began to develop in the 1940s, when a group of scientists from various fields came to the idea of the possibility of creating an artificial brain [1]. Then its theoretical, mathematical foundations and components began to form [2]. Research in the field of neuroscience has shown that the brain is a network of neurons that exchange electrical signals on the principle of "all or nothing", 0 or 1. The cybernetics of the American mathematician Norbert Wiener described the basics of control and stability in electrical networks [3]. From the information theory of the American electrician and mathematician Claude Shannon became aware of digital signals [4]. English mathematician and cryptographer Alan Turing theoretically substantiated the possibility of any computation using digital operations [5]. American neuropsychologist and neurophysiologist Warren McCulloch and neurolinguist and mathematician Walter Pitts analyzed networks consisting of idealized artificial neurons and demonstrated how they could perform the simplest logical functions [6]. Canadian physiologist and neuropsychologist Donald Gebb proposed a theory of learning based on the conjuncture of neural networks and synapses that can strengthen or weaken over time [7]. These scientists were the first to describe what researchers would later call a neural network.

In 1957, Frank Rosenblatt proposed the Perceptron - a mathematical or computer model of information perception by the brain (cybernetic model of the brain), first embodied in the form of the electronic machine "Mark-1" in 1960. The Perceptron became one of the first models of neural networks, and the "Mark-1" was the world's first neurocomputer [8].

So, a neural network is essentially a mathematical model that emulates the functioning of the human nervous system [9]. In turn, a mathematical model is a system of mathematical relationships that describe the object under study process or phenomenon. Thus, a neural network is a mathematical model that describes the interdependence between various factors using mathematical formulas and allows you to identify patterns in the interdependence of input data and the final result in the learning process. In the mathematical model of a neural network, there are nodes – artificial neurons. The nodes are interconnected by connections and transmit signals to each other (Fig. 1).

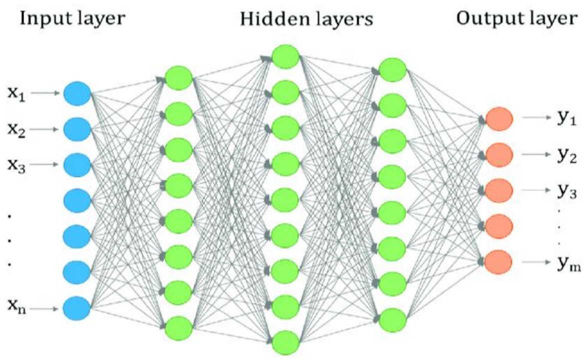

Fig. 1. Scheme of a multilayer neural network

The simplest mathematical model of a neuron is called Perceptron [10]. The perceptron may be visualized as shown in Fig. 2:

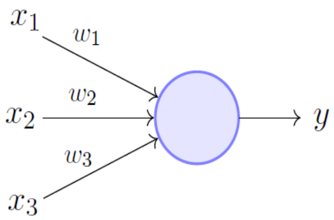

Fig. 2. Schematic representation of a perceptron

A perceptron is a mathematical function that is given several parameters “x”, and the result of the calculation is a number “y”, for example y = 3*x. Each input value “x” is given a certain weight “w”, which can be adjusted, depending on how important this value is and how it affects the result of the calculation. In this case, this function will look like this:

$$y = \begin{cases} 0 \text{ if } \Sigma iwjxj \leq \text{threshold value} \\ 1, \text{ if } \Sigma iwjxj > \text{threshold value} \end{cases}$$

In fact, the function looks more complicated: in essence, it means that the result of the neuron's operation is 0 if the sum of the products of all input parameters and their weights is less than a certain threshold value or is equal to 1 if the sum of all input parameters and their weights is greater than a certain threshold value. 0 – the neuron is passive, 1 – the neuron is active. The result of the calculation is passed as input parameters to other neurons of the next layer of the neural network. The larger “y” – the more strongly the neuron is “activated” in the next layer, and so on.

So, the work of a perceptron can be imagined as a device that makes decisions based on weighted input data. If you imagine a mathematical model with a thousand perceptrons, that is, a thousand neurons, each of which will have thousands or tens of thousands of “x” at the input. And the more nodes, the greater the variability that the network can model. By controlling the weights of the connections between neurons, we can adjust the neural network, and thus train it to produce the expected results.

The perceptron is the simplest but least efficient model of a neural network node. In addition to the perceptron, there are also linear neurons, sigmoid neurons, ReLU neurons, tanh neurons, and SoftMax neurons [11]. Each has specific functions for different neural network tasks.

The architecture of a neural network can be divided into two components: the architecture of the mathematical model of the neural network and the architecture of the software that implements the functions of the neural network. To begin with, let's list the most commonly used architectures of mathematical models:

Feedforward neural networks. Signals are transmitted in one direction from one layer to another. This is the simplest structure designed for classification or prediction.

Recurrent neural networks. In this structure, neurons can have connections not only forward but also backward, creating loops in the network. Used, for example, in machine translation, where context and previous context are important for correct translation.

Convolutional neural networks. Enable local data processing and feature recognition in images. Convolutional neural networks are effective in object recognition and pattern detection in images.

In addition to the above, there are recurrent neural networks with long-term memory (Long Short-Term Memory, LSTM), convolutional recurrent neural networks (Convolutional Recurrent Neural Networks), generative adversarial neural networks (Generative Adversarial Networks), associative memory networks (Hopfield Networks, Boltzmann Machines), memory networks (Memory Networks), generative adversarial networks (Generative Adversarial Networks, GAN) and others [12]. Neural network architectures are constantly being developed and modified in order to increase the efficiency of learning, reduce errors in work, and perform complex complex missions.

From the software application point of view, when implementing neural network functions, three levels can be identified:

- *Basic level* – this is the basic part of the neural network, which is formed in the process of basic deep learning;
- *Level of adaptation* – this is adaptation to a specific organization, user, legislation, and this level contains the unique experience of each practical implementation of a neural network;
- *Contextual level* – the result of training a neural network during the interaction of specific subjects with it, that is, the context of a specific user or organization.

Therefore, neural networks are a modern powerful AI tool that can model and simulate the work of the human brain. They are a practical mechanism that ensures the implementation of Machine Learning and Deep Learning technologies.

Machine Learning is a subfield of AI that studies algorithms and models that allow computers to learn and improve their performance based on data, without explicit programming [13].

Deep Learning is a subfield of machine learning that uses neural networks with a large number of layers to automatically learn data representations. Deep learning allows computers to understand and analyze complex relationships, discover hidden patterns, and make predictions based on large amounts of data [14].

The difference between Deep Learning and Machine Learning is in the complexity of the tasks [15]. Deep Learning detects more complex dependencies and uses much more data to build patterns. Companies such as Google, Microsoft, and Facebook are implementing these technologies in their products. The real sensation was the ChatGPT language model (the latest available version of GPT-4), which generates text based on user queries and taking into account the context of previous queries [16]. This is a deep neural network model based on the GPT (Generative Pre-trained Transformer) architecture. GPT uses a transformer architecture and multilayer recurrent neural networks (RNN) to generate text and answers to questions.

Today, neural networks are used in various technological products and industries, including cybersecurity.

Conclusions

The foundations and components of AI are cybernetics, information theory, theory of computation using digital operations, logic, neuropsychological theory, theory of neural networks, learning theory based on the conjuncture of neural networks and synapses, mathematical (computer) model of the brain. The real breakthrough in the use of AI began in the 2010s with the development of machine and deep learning. Digitalization, Big Data and the ability to process large data sets provided the necessary resource for this.

Therefore, AI models based on neural networks are used to solve complex problems that cannot always be solved using mathematical formulas or rules. They are capable of learning from large amounts of data and recognizing complex dependencies in this data. They can be used in various technological products and industries, including cybersecurity.

References

1. The History of Artificial Intelligence from the 1950s to Today. URL:https://www.freecodecamp.org/news/the-history-of-ai/(date of application 10/24/2024).
2. Lysetsky Yu.M., Sokur V.V. Theoretical and mathematical foundations of artificial intelligence. *Military Intelligence Bulletin.* 2024.Vol.85. P. 51-56.
3. Norbert Wiener. Cybernetics or Control and Communication in the Animal and the Machine, Paris - Cambridge, Mass., 1948. 194 p.
4. Shannon CEA Mathematical Theory of Communication. *Bell System Technical Journal*. 1948. No. 27. P. 379-423.
5. Turing AM Computing Machinery and Intelligence. *Mind.* 1950. No. 49. P. 433-460.
6. Warren S. McCulloch and Walter Pitts, A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics.* 1943. Vol. 5. P. 115-133.
7. Donald Olding Hebb The Organization of Behavior: A Neuropsychological Theory. DO Hebb: The Organization of Behavior, Wiley: New York, 1949. 337 p.
8. Rosenblatt F. The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. *Psychological Review.* 1958. Vol. 65. No. 6. P. 386-408.
9. Kalbazov D.Y., Danchenko O.I., Lysetsky Yu.M. Neural networks. Development and prospects. *Mathematical machines and systems.* 2024. No. 2. P. 24–32.
10. What is Perceptron? A Beginners Guide for 2023. URL: https://www.simplilearn.com/tutorials/deep-learning-tutorial/perceptron (date of application 15.11.2024).
11. 7 architectures of neural networks for solving NLP problems.URL: https://neurohive.io/ru/osnovy-data-science/7-arhitektur-nejronnyh-setej-nlp/(date of application 09/23/2024).
12. What are the types of neural networks? URL:https://devzone.org.ua/qna/iaki-isnuiut-tipi-neironnix-merez.
13. What is machine learning? URL: https://www.ibm.com/topics/machine-learning(date of application 17.01.2024).
14. What is deep learning and how does it work? URL: https://www.techtarget.com/searchenterpriseai/definition/deep-learning-deep-neural-network (accessed 12/12/2024).
15. What Is Deep Learning? How It Works, Techniques. URL: https://www.mathworks.com/discovery/deep-learning.html (accessed 12/12/2024).
16. What is ChatGPT from OpenAI capable of and how can it be used? URL: https://www.imena.ua/blog/chatgpt/?gclid=EAIaIQobChMIp-mGo9f8gQMVvJiDBx22rAZYEAAYASAAEgJcnPD_BwE (date of application 12/15/2024).