

## APPLICATION OF ARTIFICIAL INTELLIGENCE (NEURAL NETWORKS) IN EDUCATION

### Abstract

*This article focuses on the use of artificial intelligence in the educational environment. We consider the latest technology, which already plays a huge role for both teachers and students. Currently, there are many systems for the development of training systems, among which artificial neural networks occupy a substantial place. The article presents an example of the use of artificial neural networks, which can play a significant role in developing educational systems.*

**Keywords:** *artificial intelligence, neural networks, learning process, system, e-learning system development*

### Introduction

For many years, modern education has remained unchanged against the background of rapid changes in other areas of contemporary society. We can say that the evolution of education lags behind the evolution of society. One of the ways to solve the problems that have arisen is the involvement in the educational process of various technical means, training systems, the use of the Internet.

The development of learning systems is currently a prevalent and rapidly developing type of scientific activity, due to the renewed interest in the use of artificial intelligence technologies in practice, as well as the rapid development of Internet technologies, which provided developers of training systems with powerful new development tools that did not exist before. Such popularity of this area of scientific research has led to the fact that at present, there are a large number of scientific papers on this topic, dozens of training systems have been developed, most of which are, in fact, hypertext documents and cannot claim to be called complete.

### Literature review

The birth of AI goes back to the 1950s when John McCarthy organised a two-month workshop at Dartmouth College in the USA. In the workshop proposal, McCarthy used the term artificial intelligence for the first time in 1956 (Russel & Norvig, 2010, p. 17):

The study [of artificial intelligence] is to proceed based on the conjecture that every

aspect of learning or any other feature of intelligence can, in principle, be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make devices use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves.

Baker and Smith (2019) provide a broad definition of AI: “Computers which perform cognitive tasks, usually associated with human minds, particularly learning and problem-solving” (p. 10). They explain that AI does not describe a single technology. It is an umbrella term to describe a range of technologies and methods, such as machine learning, natural language processing, data mining, neural networks or an algorithm.

AI and machine learning are often mentioned in the same breath. Machine learning is a method of AI for supervised and unsupervised classification and profiling, for example, to predict the likelihood of a student to drop out from a course or being admitted to a program or to identify topics in written assignments. Popenici and Kerr (2017) define machine learning “as a subfield of artificial intelligence that includes software able to recognise patterns, make predictions, and apply newly discovered patterns to situations that were not included or covered by their initial design” (p. 2).

The concept of rational agents is central to AI: “An agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators” (Russel & Norvig, 2010, p. 34). A vacuum cleaner robot is a straightforward form of an intelligent agent, but things become very complex and open-ended when we think about an automated taxi.

Experts in the field distinguish between weak and strong AI (see Russel & Norvig, 2010, p. 1020) or narrow and general AI (see Baker & Smith, 2019, p. 10). A philosophical question remains whether machines will actually think or even develop consciousness in the future, rather than just simulating thinking and showing rational behaviour. It is unlikely that such strong or general AI will exist shortly. Therefore, we are dealing here with GOFAI (“good old-fashioned AI”, a term coined by the philosopher John Haugeland, 1985) in higher education – in the sense of agents and information systems that act as if they were intelligent.

Given this understanding of AI, what are potential areas of AI applications in education, and higher education in particular? Luckin, Holmes, Griffiths, and Forcier (2016) describe three categories of AI software applications in education that are available today: a) personal tutors, b) intelligent support for collaborative learning, and c) intelligent virtual reality.

Intelligent tutoring systems (ITS) can be used to simulate one-to-one personal

tutoring. Based on learner models, algorithms, and neural networks, they can make decisions about an individual student's learning path and the content to select, provide cognitive scaffolding and help, and engage the student in dialogue. ITS have enormous potential, especially in large-scale distance teaching institutions, which run modules with thousands of students, where one-to-one human tutoring is impossible. A vast array of research shows that learning is a social exercise; interaction and collaboration are at the heart of the learning process (see, for example, Jonassen, Davidson, Collins, Campbell, & Haag, 1995). However, online collaboration has to be facilitated and moderated (Salmon, 2000). AIED can contribute to collaborative learning by supporting adaptive group formation based on learner models, by promoting online group interaction or by summarising discussions that can be used by a human tutor to guide students towards the aims and objectives of a course. Finally, also drawing on ITS, intelligent virtual reality (IVR) is used to engage and guide students in authentic virtual reality and game-based learning environments. Virtual agents can act as teachers, facilitators or students' peers, for example, in virtual or remote labs (Perez et al., 2017).

With the advancement of AIED and the availability of (big) student data and learning analytics, Luckin et al. (2016) claim a “[r]enaissance in assessment” (p. 35). AI can provide just-in-time feedback and assessment. Rather than stop-and-test, AIED can be built into learning activities for an ongoing analysis of student achievement. Algorithms have been used to predict the probability of a student failing an assignment or dropping out of a course with high levels of accuracy (e.g. Bahadir, 2016).

In their recent report, Baker and Smith (2019) approach educational AI tools from three different perspectives; a) learner-facing, b) teacher-facing, and c) system-facing AIED. Learner-facing AI tools are software that students use to learn a subject matter, i.e. adaptive or optimization learning management systems or ITS. Teacher-facing systems are used to support the teacher and reduce his or her workload by automating tasks such as administration, assessment, feedback and plagiarism detection. AIED tools also provide insight into the learning progress of students so that the teacher can proactively offer support and guidance where needed. System-facing AIED is a set of tools that provide administrators and managers information on the institutional level, for example, to monitor attrition patterns across faculties or colleges.

In the context of higher education, we use the concept of the student life-cycle (see Reid, 1995) as a framework to describe the various AI-based services on the broader institutional and administrative level and support the academic teaching and learning process in the narrower sense.

## Research question

This paper will attempt to answer two basic queries: What is “neural networks”? – What can offer neural networks used in artificial intelligence in education?

Currently, there are many systems for the development of educational systems, among which the use of artificial neural networks occupies a vital place. Neural networks are computational structures that model simple biological processes commonly associated with processes in the human brain. Adaptable and trainable, they represent parallelized systems capable of learning by analyzing positive and negative impacts. An elementary converter in these networks is the artificial neuron, formal neuron or a neuron, named by analogy with the biological prototype.

The term “neural networks” was formed in the 40s of the XX century among researchers who studied the principles of organization and functioning of biological neural networks. The main results obtained in this area are associated with American researchers W. McCulloch, D. Hebb, F. Rosenblatt, M. Minsky, J. Hopfield, and others.

Interest in neural networks has grown and then faded. Currently, there is an explosion of interest in trained neural networks. The neural network consists of formal neurons (Fig. 1)

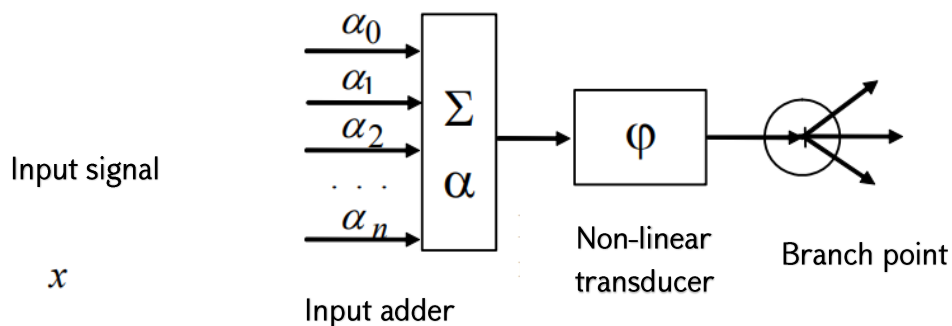


Fig. 1. Formal neuron

The adaptive adder calculates the dot product of the input signal vector  $x$  by the parameter vector  $\alpha$ . We call it adaptive because of the presence of a vector of adjustable parameters  $\alpha$  (vector of synaptic weights of a neuron).

Nonlinear Signal Converter – Receives a scalar input signal  $x$  and translates it to  $\phi(x)$ .

A branch point is used to send one signal to several addresses. Linear connection – the synapse – does not occur separately from adders. It multiplies the input signal  $x$  by

the “synapse weight”  $\alpha$ .

Artificial neural networks are already being used to solve many problems: Image classification: The task is to indicate the belonging of an input image (for example, a speech signal or a handwritten symbol), represented by a feature vector, to one or more predefined classes. Known applications include letter recognition, speech recognition, electrocardiogram signal classification, blood cell classification.

Clustering / categorization (no teacher). In solving the clustering problem, which is also known as the classification of images of “unsupervised” learning, is not fetch with class labels. The clustering algorithm is based on the similarity of images and places close images into one cluster. There are known cases of using clustering for knowledge extraction, data compression, and data properties exploration.

Function approximation. Suppose we have a training set  $((x_1, y_1), (x_2, y_2), \dots, (x_N, y_N))$  (input-output data pairs), which is generated by an unknown function  $F(x)$ , distorted by noise. The approximation problem is to find an estimate for the unknown function  $F(x)$ . Function approximation is essential for solving numerous engineering and scientific modelling problems.

Prediction / forecast. Let  $k$  discrete samples  $\{y(t_1), y(t_2), \dots, y(t_k)\}$  be given at successive times  $t_1, t_2, \dots, t_k$ :- The problem is to predict the value of  $y(t_k + 1)$  at some future time  $t_k + 1$ . The prediction/forecast has a significant influence on decision making in business, science and technology. Stock price prediction and weather forecasting are typical applications of the prediction/forecasting technique.

Optimization. Numerous mathematics, statistics, engineering, science, medicine, and economics can be considered optimization problems. The optimization algorithm's task is to find a solution that satisfies the system of constraints and maximizes or minimizes the objective function. The famous travelling salesman problem is a classic example of an optimization problem.

**Content-addressable memory.** In the Von Neumann computation model, memory access is available only through an address that does not depend on the memory contents. Moreover, if an error is made in calculating the address, then utterly different information can be found. Associative memory, or content-addressable memory, is accessed by specifying a given content. Memory contents can be recalled even on partial input or garbled contents. Associative memory is highly desirable when creating multimedia information databases.

**Management.** Consider a dynamical system defined by the set  $\{u(t), y(t)\}$ , where  $u(t)$  is the input control action, and  $y(t)$  is the output of the system at time  $t$ . In control

systems with a reference model, the purpose of control is to calculate such an input action  $u(t)$  in which the system follows the desired trajectory dictated by the reference model. An example is optimum motor control.

Tasks related to pattern recognition can be applied when building a reference model of a student that meets qualification requirements, assessing the current state (portrait) of a student in the space of features that maximally reflect all aspects of his activity, operational monitoring of the dynamics of a student's portrait, making the necessary management decisions to optimization of the educational process.

The classification problem with guided learning of the network can be interpreted as the problem of recognizing the current state of the student, based on the comparison of the trainee's image presented to the network with the control target sample and mapping the input image to the target output sample, which gives the network information about which class the input sample should be learned to refer to. Clustering images with network learning without control (self-learning), based on grouping data using competition, allows you to select groups (clusters) of learners and analyze the quality of the educational process at various levels.

Today, there is no universal teaching methodology, so the teacher chooses the most acceptable teaching ways based on his own experience, which is not always optimal. As a criterion of effectiveness, we consider the depth of mastery of the subject by the learners, the completeness and strength of the knowledge they have acquired, the level of study of theoretical material and the acquisition of practical skills.

**The novelty.** To solve this problem, we propose to create a system that would divide a group of trainees into subgroups, depending on the results of the entrance testing. After that, the system offered some effective teaching methodology for each of the subgroups, directing it to the corresponding training layer. At the end of the course, the group is invited to pass the absolute control, "analyzing" the system's results, adjusting the distribution weights for the input testing. The next group of learners was divided into subgroups already under the new weights.

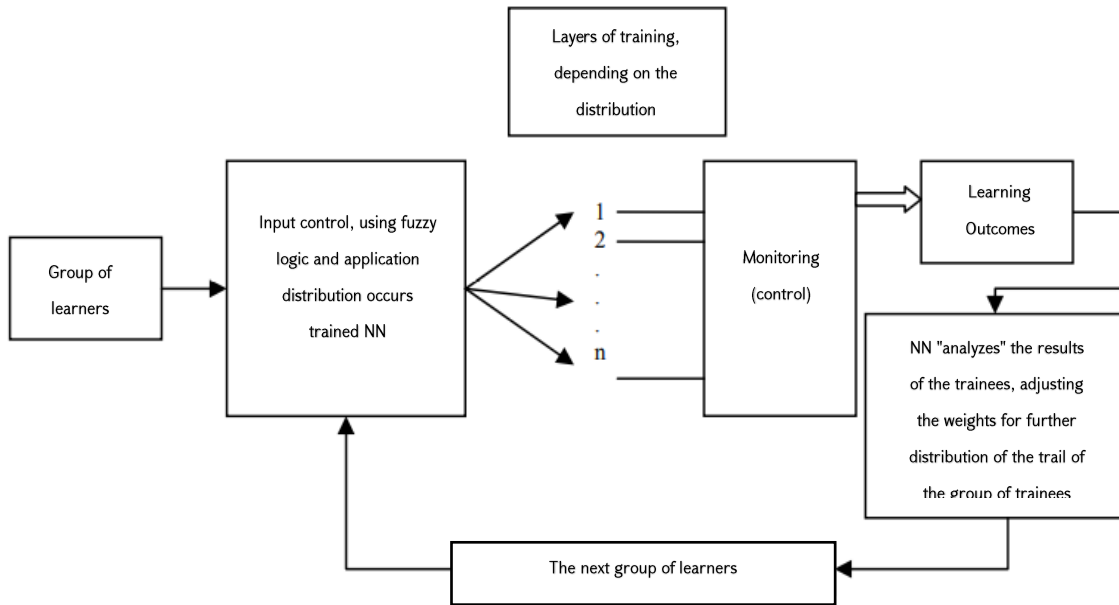


Fig. 2

In this system's design, it is necessary to develop an input testing system, select the architecture and topology of the network, the learning algorithm for the network, the function of activating neurons (for example, "sigmoid"). Let us more detail on the architecture and learning algorithm.

Learning ability is a fundamental property of the brain. In the context of ANN, the learning process can be considered setting up the network architecture and linking weights for the efficient performance of a particular task. Typically, a neural network must adjust the link weights for the available training sample. Network performance improves as the weights are iteratively adjusted.

A network's property to learn from examples makes them more attractive than systems that follow a particular system of rules of operation formulated by experts.

There are three paradigms of learning: "with a teacher", "without a teacher" (self-study) and mixed. In the first case, the neural network has the correct answers (network outputs) for each input example. The weights are adjusted so that the network produces responses as close as possible to the known correct answers. Unsupervised learning does not require knowing the correct answers for each example of a training set. In this case, the internal structure of the data or the correlations between the samples in the data system is revealed, which allows you to categorize the samples. In blended learning, some weights are determined through supervised learning, while the rest is obtained through self-learning. There are four main types of learning rules: error correction, Boltzmann machine, Hebb rule, and competition learning.

**Rule of error correction.** In supervised learning, each input example is given the desired output  $d$ . The actual output of the network may not match the desired one. The principle of error correction during training is to use signal  $(d-y)$  to modify the weights to reduce the error gradually. Learning takes place only when the perceptron is wrong.

**Boltzmann's training.** The goal of Boltzmann training is to adjust the weights so that the visible neurons' states satisfy the desired probability distribution. Boltzmann training can be considered a special case of error correction, in which the error is understood as the discrepancy between the state correlations in two modes.

**Hebb's rule.** The oldest teaching rule is Hebb's teaching postulate. Hebb relied on the following neurophysiological observations: if neurons on both sides of the synapse are fired simultaneously and regularly, then the strength of the synaptic connection increases. An essential feature of this rule is that the change in synaptic weight depends only on the activity of neurons that are connected by a given synapse.

**Competition training.** Unlike Hebb learning, in which multiple output neurons can be fired simultaneously, in competitive learning, the output neurons compete with each other for firing. This phenomenon is generally known as "winner-take-all." Similar learning takes place in biological neural networks (Tab 1.).

Paradigm	Teaching rule	Architecture
With teacher	Error correction	Single and multilayer perceptron
	Boltzmann	Recurrent
	Hebb	Multilayer feedforward
	Competition	ART network
Mixed	Error correction and competition	RBF network

**Table 1. A possible architecture and associated type of training rule for the classification problem**

The architecture of a multilayer feedforward network can be applied in our case because it is most suitable for classification problems. A single-layer network is not suitable. The most superficial single-layer neural networks can solve only linearly separable issues, and this limitation is overcome when using multilayer neural networks.

**Conclusion.** Artificial neural networks are used to classify information in the case of limited, incomplete and non-linear data sources. Neural network technologies are



versatile, and the same program provides an opportunity to work in different areas of knowledge. Such systems do not need to be reprogrammed when changing the composition of the training base. The importance of this feature of neural networks can hardly be overestimated in light of the ever-increasing volume of information. All of the above allows us to say that the introduction of neural network technologies in processing and interpreting information is an important and promising direction in the development of training systems.

## References

- Bahadir, E. (2016). Using neural network and logistic regression analysis to predict prospective mathematics teachers' academic success upon entering graduate education. *Kuram ve Uygulamada Egitim Bilimleri*, 16(3), 943–964. <https://doi.org/10.12738/estp.2016.3.0214>.
- Baker, T., & Smith, L. (2019). Educ-AI-tion rebooted? Exploring the future of artificial intelligence in schools and colleges. Retrieved from Nesta Foundation website: [https://media.nesta.org.uk/documents/Future\\_of\\_AI\\_and\\_education\\_v5\\_WEB.pdf](https://media.nesta.org.uk/documents/Future_of_AI_and_education_v5_WEB.pdf) (accessed Monday, April 19, 2021)
- Baker, T., & Smith, L. (2019). Education rebooted? Exploring the future of artificial intelligence in schools and colleges. Retrieved from Nesta Foundation website: [https://media.nesta.org.uk/documents/Future\\_of\\_AI\\_and\\_education\\_v5\\_WEB.pdf](https://media.nesta.org.uk/documents/Future_of_AI_and_education_v5_WEB.pdf) (accessed Monday, April 19, 2021)
- Haugeland, J. (1985). *Artificial intelligence: The very idea*. Cambridge, Mass.: MIT Press
- Jonassen, D., Davidson, M., Collins, M., Campbell, J., & Haag, B. B. (1995). Constructivism and computer-mediated communication in distance education. *American Journal of Distance Education*, 9(2), 7–25. <https://doi.org/10.1080/08923649509526885>.
- Kol'cov YU. V. Dobrovol'skaya N. YU. Nejrosetevye modeli v adaptivnom komp'yuternom obuchenii (Neural network models in adaptive computer learning) - // *Educational Technology & Society*. - 2002. - 5(2). - 213-216 str.
- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). Intelligence unleashed - an argument for AI in education. <http://discovery.ucl.ac.uk/1475756/> (accessed Monday, April 19, 2021)
- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). Intelligence unleashed - an argument for AI in education. <http://discovery.ucl.ac.uk/1475756/>

(accessed Monday, April 19, 2021)

- Minsky M., and Papert S., (1971). Perceptrons. Cambridge, MA: MIT Press. (Русский перевод: Минский М. Л., Пейперт С. Перцептроны. –М. Perez, S., Massey-Allard, J., Butler, D., Ives, J., Bonn, D., Yee, N., & Roll, I. (2017). Identifying productive inquiry in virtual labs using sequence mining. In E. André, R. Baker, X. Hu, M. M. T. Rodrigo, & B. du Boulay (Eds.), Artificial intelligence in education, (vol. 10,331, pp. 287–298). [https://doi.org/10.1007/978-3-319-61425-0\\_24](https://doi.org/10.1007/978-3-319-61425-0_24). (accessed Monday, April 19, 2021)
- Popenici, S., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. Research and Practice in Technology Enhanced Learning. <https://doi.org/10.1186/s41039-017-0062-8>.
- Reid, J. (1995). Managing learner support. In F. Lockwood (Ed.), Open and distance learning today, (pp. 265–275). London: Routledge.
- Russel, S., & Norvig, P. (2010). Artificial intelligence - a modern approach. New Jersey: Pearson Education.
- Russel, S., & Norvig, P. (2010). Artificial intelligence - a modern approach. New Jersey: Pearson Education.
- Salmon, G. (2000). E-moderating - the key to teaching and learning online, (1st ed.,). London: Routledge.