Adaptive Learning Using Artificial Intelligence in Distance Education[*]

Petr V. Chetyrbok[1][0000-0002-0115-9158], Marina A. Shostak[1] [0000-0002-3834-0559],
Lennara U. Alimova [2] [0000-0001-7815-5262]

[1] V.I. Vernadsky Crimean Federal University, Simferopol, Russia
[2] State Budgetary Educational Institution "Crimean Engineering and Pedagogical University"
named after F. Yakubov, Simferopol, Russia

petrchetyrbok@gmail.com

Abstract. Reviewed and analyzed is the use of artificial intelligence (AI) in distance education. The authors highlight several areas of artificial intelligence's application in education. Adaptive learning is perhaps one of the most promising lines of artificial intelligence's application in education.

Keywords: artificial intelligence, distance education, neural networks, platform, course, adaptive learning

1 Introduction

The distant method allows organizing the learning process online in the real-time mode. Students and teachers communicate via the Internet. Teachers transmit knowledge to students who receive new information and assignments and pass tests. At the same time, teachers can be removed from the student at any distance, they can live in different countries and on different continents.

Under the current conditions, the quality of education depends on information transmission and the methods of its presentation, including the use of artificial intelligence elements.

We can distinguish the following features and properties of artificial intelligence systems distinguishing them from conventional automation systems:

1) a purpose or group of purposes;
2) planning their actions and search for solutions to problems;
3) learning and adapting their behavior patterns within the operating process;
4) working in a poorly formalized environment, under the conditions of uncertainty and following fuzzy instructions;
5) self-organization and self-development;
6) understanding natural language texts;

---

7) data generalization and abstraction.

The purpose of the article is to review and analyze how the directions of artificial intelligence application in adaptive learning could be used in the educational system.

2   Related Works

Adaptive learning is the learning mode taking into account learners' abilities, knowledge, skills, and moods. In the 1950-60s, programmed learning algorithms started to develop actively. The programs created on their basis controlled students' current knowledge, which facilitated the increase in the effectiveness of educational materials' gradual mastering. Adaptive testing appeared in this area. In his writings, Gordon Pask subsequently outlined the main provisions of adaptive learning. In his theory, he substantiated the idea that a high-quality curriculum should be individualized and based on learners' knowledge, taking into account their activity and attitude to the subject. The need for algorithms taking into account learners' characteristics has become urgent nowadays. It is essential to develop services implementing these algorithms and using artificial intelligence elements. Jose Ferreira, the founder of Knewton, notes: "In April 2013, I wrote a post in which I predicted that over the next several years, all educational materials would become digital and adaptive. Knewton played an important role in this revolution. We want to create a world in which each learning application is a priori adaptive. But it is not so simple. Nowadays, there are numerous applications in the market having rich educational experience, beautiful interfaces, interesting content, and well-thought-out methodology. But a significant drawback of many of these applications is the inability to adapt to learners' knowledge".

Knewton is known to be one of the first companies to actively apply data analysis technologies in the field of education. As a result, an adaptive educational platform was created that can be connected to any modern educational process management system (LMS). The Knewton platform allows creating universal algorithms and developing an extensive infrastructure for collecting, analyzing, and using information about learners' performance, such as:

1. Collecting data on learners' knowledge and the degree of certain concepts' mastering.
2. Conclusions based on the collected data on learners' characteristics and their reactions to changes in learning.
3. The choice of the optimal learning strategy for every student based on personalization. The Knewton algorithms based on the personalization of learning success forecasts offer the optimal structure and complexity level of adaptive learning courses. Knewton provides an infrastructure platform offering software developers the algorithms to adapt educational resources and helps them create unique, personalized, and flexible educational courses. "Analytics" is linked to the Knewton platform using the API and allows answering the following questions:

– what do learners know?

- why has a student made a mistake?
- what topics and disciplines are best studied by a certain student at a given time?
- what are the chances that the student will solve the problem correctly?

Being of particular relevance, personalized learning is a wide range of educational programs in which the method and pace of learning depend on learners' needs, special interests, and preferences.

Artificial intelligence adapts the learning process to the individual learning rate of every student and offers tasks of increased complexity.

Adaptive learning technologies are applied in commercial projects in the sphere of Human Resources (HR). Competentum, Ispring, E-mba are available in the Russian market. Artificial intelligence is introduced in language teaching ("Skyeng", "Lingualeo", "Websoft"), as well as programming and design ("Geekbrains", "Netology").

The Moodle learning management system is a distance learning system distributed in the open-source code, which makes it possible to customize it to the characteristics of a particular educational project using certain elements of artificial intelligence.

Moodle is a platform used in distance education providing both teacher-student and student-student types of communication. It offers the following services:

- exchanging files of any format;
- newsletters allowing to quickly inform all the course participants or individual groups about current events;
- creating and storing each student's portfolio containing all the tasks submitted by him or her, all the assessments and teacher's comments on the tasks, all the messages in the forum.

The teacher creates his or her assessment system within the framework of the course, which allows issuing summary statements of performance, controlling students' attendance, activity, and the time they spend in the educational network. The modular structure of the system makes it easy-to-use by students and teachers.

Educational institutions pay attention to students' attitudes towards teachers and conduct surveys. Even though digital ones have replaced paper questionnaires, the feedback process has not changed much. However, the time has come to revise it because students' feedback is an important source of information.

Artificial intelligence offers several interesting opportunities to optimize this process:

- Chatbots can collect information using a dialogue interface simulating real interviews. This process does not require much effort from students.
- Conversations can be adapted to students' individual features of character and changed depending on their answers.
- Chatbots can filter out rude comments and personal insults that sometimes occur in feedback forms.

The "Smart Campus" answers any student inquiries related to studying and living on campus: how to find a lecture hall, register for a chosen course, get assignments, find a free parking space, or contact a professor.

The University of Western Australia (UWA) already has such a "Smart Campus". It is operated by the Watson supercomputer system created by IBM.

3   Results

The Optimal Model. In this article, the models will include:

1. Measuring students' cognitive and emotional state;
2. Creating an open student model to stimulate students' thinking and reasoning;
3. Providing dynamic assistance (effective feedback) to increase students' motivation and involvement;
4. Using models of social modeling, e.g. so that language learners could more successfully interact with native speakers of the language they learn, as well as understand cultural and social norms.

Century Tech Platform. Teachers try to pass the same knowledge to students at different levels, but often fail to do it. Artificial intelligence helps to adapt lessons to the individual characteristics of every student. Within the following five years, Century Tech will be deployed in all 700 Flemish schools. It is the first time that the authorities have decided to introduce artificial intelligence in schools to such an extent. The goal is to replace the standard model, in which teachers try to pass knowledge to students at different levels, but often fail to do it, with a fundamentally new one, in which artificial intelligence helps to adapt lessons to the individual characteristics of every student. This includes identifying areas of knowledge they are familiar or unfamiliar with, where they feel confident or lack confidence, and an assessment of how much they like certain activities. The artificial intelligence system constantly analyzes every student's performance, identifying trends in knowledge mastering, and adapting lessons by them. In other words, if your third-grade student cannot understand the equation, he will not have to get ahead of the rest of the class and listen to you explaining "what is X" in the evenings. Artificial intelligence will determine his problem and let him master the material in the form of micro tricks.

The introduction of artificial intelligence systems in schools will become possible when students have constant access to computers at school and the educational technologies collected in them. When every student starts using his or her digital device, it will be possible to say that the introduction of a full-fledged digital school has begun. This is the only way to get a digital educational footprint and an initial technological foundation for forming individual educational trajectories using artificial intelligence. And then teachers will begin to acquire and adapt their knowledge about students thanks to the data obtained during each interaction with the system. What have our schools managed to implement so far? Moscow electronic school is actively used in Moscow. It is important to understand that it is not an artificial intelligence system yet, but rather the first attempt to automate the learning process.

A national project called "Education" has been developed in Russia. Its overarching goal is to create a unified educational platform allowing every student to receive a high-

quality quality education, including the use of adaptive learning and individual educational trajectories. The model of a unified digital educational environment was developed in 2019. Artificial intelligence as technology can become part of this ecosystem.

On the other hand, in our country, there are already examples of the use of artificial intelligence in education. For example, the Russian startup "Parla" created an application for learning English. The application is based on a program that is taught alongside students and adapted to their tasks and progress. Already at the registration stage in the application, the program can analyze data from social networks and offer an individual learning program based on a particular person's interests. It is a commercial project, but most technologies appear primarily as a commercial product.

According to British developers, their platform is used 20 minutes a week in elementary schools and 40 minutes a week in high schools (on average). Applicants claim that the system maximizes six hours a week. This time can be spent on creative activity and physical education, which are often "neglected" in comparison with academic disciplines due to a serious examination load on students.

The platform also applies the basic principles of neurophysiology to determine students' abilities, levels of knowledge and endurance, individual pace and appropriate time for learning, as well as the speed of information transfer from the short-term memory to the long-term memory. These analytical data are later incorporated into automated learning algorithms.

The use of artificial intelligence can solve the problem of accompanying students along their educational path (IEP) at a new qualitative level. The fact is that the real IEP is dynamically rebuilt as the student develops. To do this, you need to conduct regular monitoring, as well as monitor any educational activities (reading and watching educational materials and problem-solving). This is what we call "an experienced teacher's intuition". But in fact, we imply the processing of the resulting large data array, which can only be done by artificial intelligence. The role of human teachers must be transformed: from the carrier of knowledge to the carrier of the philosophy of the studied subject. Conceptual things can be comprehended and mastered by people, not by artificial intelligence systems. Teachers will pass them to children in small groups due to enhanced free time. "It seems that the Belgian initiative may well become a Russian reality, but not very soon. On the other hand, individual educational trajectories are already being introduced to higher education institutions. Universities such as the Higher School of Economics provide their students with online courses provided by other universities according to students' specialties. It means that students can already change the curriculum to some extent and independently choose were to master the material and up to what extent.

Big data are just an archive in which the accumulated information is stored. This archive is very large containing millions of records or data points/sets of data, such as students' names, their annual grades, absence due to illness, and the number of lost textbooks or goals scored in physical education. Previously, no one collected such data, because there would not be enough resources for accounting, storage, and, especially, for analysis. At the moment, the easiest way of collecting data is used by developers of electronic textbooks, online courses, or mixed-type learning systems. They are used immediately, helping to set up the system for on-the-fly learning: the more data, the

more accurate it is. Big data analysis can speed up the solution of scientific, research, and educational problems. Studying statistics, one can work with both individual trajectories and global educational systems. The methods used in big data analysis are based on computer-aided learning, image recognition, psychometry, and statistics. The most popular PSLC Data shop repository stores information collected over 250,000 hours spent by students on educational programs – about 30 million actions, answers, and results.

Trial and Error Method. Using neural networks based on this method, it has been possible to solve complex problems in education. For example, a system can improve students' essays and indicate answers to questions [8].

Neural Network Duelists. The structure is as follows: one system consists of a network generating new data after training, and another system divides the data of the previously specified network into "true" and "false". Such a set can ultimately produce very realistic synthetic data, and so far such "duels" of neural systems are excellent for creating landscapes in computer games, improving pixel videos, or for applying stylistic changes to a design that was created using a computer. Duelists can also be useful for improving the quality of educational materials. For example, in studying foreign languages there is a huge segment for filling dictionaries with current word meanings, where the filter is most often used by other users or native speakers [9], [10].

Data Collection and Personalization. Using geolocation data and our previous searches, artificial intelligence can already offer us the perfect cafe nearby or, for example, give directions to the nearest store of our favorite comic books. Nowadays, smartphones can hold a charge longer, because they keep a record of our interaction with applications and prioritize, quitting those that we hardly use. Most commonly used applications start up faster and issue notifications.

Now let us imagine how we learn a grammar rule based on the examples only from the area we are interested in, and all educational materials are adapted to meet our needs. This is the type of content we will continue to search for.

The Main Models of Big Data. The most interesting model of working with big data is forecasting, where a combination of the known data allows predicting something unknown that a person is looking for. The known data are collected from records of school systems, online services, surveys, and observations during experiments. It is very important to collect such data. One needs to know what to pay attention to and be able to determine the necessary information. The model is used to predict the future (we calculate whether the student can solve the following problem and what result he or she is likely to achieve) or the present (judging from the statistics over the last hour, we find out whether a student is interested in viewing online materials now). The estimated value may be a number: for example, the time spent on solving a problem, the number of sentences used, the percentage of video views, or the test result in points. There may be a category: quit / not quit, ask, try to solve/ask for a hint, which will be A / B / C / G. In such cases, the classification method and various algorithms, such as a decision tree or clustering, are used (Fig. 1) [11], [12].
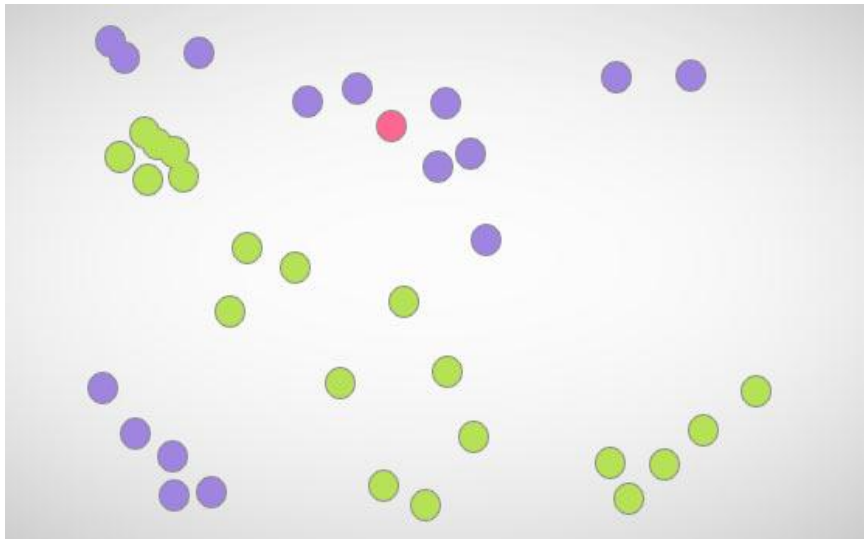
Fig. 1. Clustering

Advanced algorithms take into account error costs and the effectiveness of proper system intervention. For instance, if a student covers 0.05% of the course per minute, the wrong forecast "costs" him 1 extra minute of learning, and the correct one – 0.03%. Forecasts should be checked. Are there dependencies in the data, or just random coincidences? To find it out, one can separate the data and see whether the dependency relevant for one group is repeated in all the others. The main question is as follows: is the determined dependency applicable to new data? There are cases when the results obtained in a sterile laboratory do not coincide with the results shown by students in real life. Unlike the forecast, where it is known what needs to be determined, the structure determination method is used to identify unknown patterns and subsequent data clustering. Another method, called "network analysis", considers all the participants in the educational process as "nodes" interconnected by ties, which can be stronger or weaker depending on the intensity and frequency of communication.

The system includes various types of interaction: teamwork with one resource, leadership, help, criticism, or even insult. The interaction data are determined by the following important parameters:

- Density: how many possible ties are established among students? This parameter shows up to what extent a group of students may be called a real class. Sometimes everyone communicates with everyone, and sometimes only a few students in the entire class make up the active core, and the rest prefer to listen passively.
- Accessibility: are there students who are not contacted? Some students do not communicate with anyone, they exist in the system "on their own."
- Distance: how many "nodes" go along the route from one student to another? The principle of five handshakes works here. The shorter the chain, the more active the class.

- Flow: how many possible routes (through different "nodes") from one student to another exist? Numerous routes stipulate more diverse ties.
- Centrality: how important is each student? Who is the most influential student in the class? This parameter is determined using the following three components. The first one is the number of ties leading to the learner. Incoming and outgoing ties are taken into account separately: a student who wants to be friends with everyone does not always look like the one with whom everyone wants to be friends. The second one is proximity, which is the sum of students' distances from each other. Strong ties are considered closer here. And the third one is the number of routes passing through the student among other "nodes".
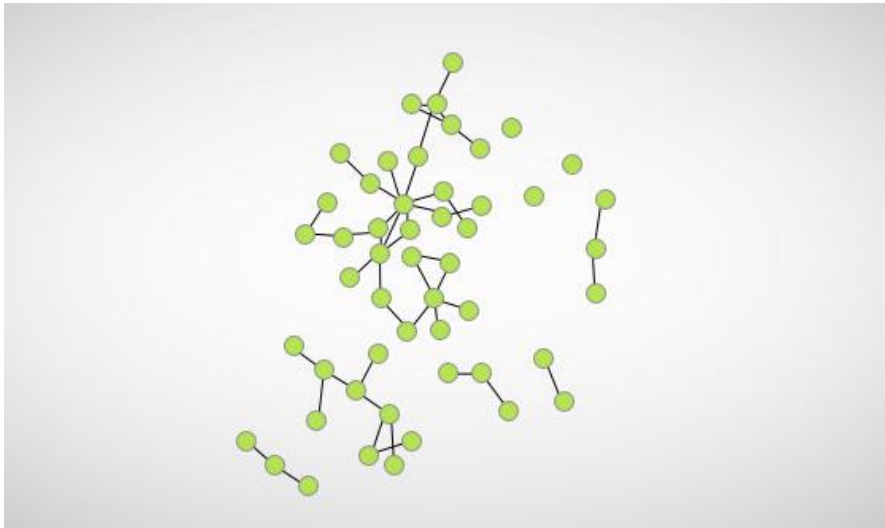

Fig. 2. Network analysis

Reciprocity: the number of bidirectional ties in all pairs of students (Fig. 2).

Proper vector: A parameter calculated mathematically and based on the number and strength of links. This algorithm is also used by Google to generate PageRank and determine the order in which results are displayed on the search page.

Assess students' knowledge to expand it competently. Also, these data will help evaluate teachers' work. Finally, following on from such data, the system can independently make pedagogical decisions. Academic performance is constantly being recorded. These data help to form templates by which one can estimate every student's knowledge. After all, students may not know something, but correctly answer a question (there is always a chance to guess) or, conversely, accidentally make a mistake.

Big data allow understanding of how students behave when they are bored. The system can determine one of the models of such behavior.

- Students try to "outsmart" the system, i.e., to succeed without any preparation. For example, they go through all possible answers until the correct answer appears.

- Students are distracted by other tasks, such as communicating with the teacher or commenting on Facebook.
- Students answer thoughtlessly and make random choices, without even trying to think.
- Students behave unpredictably.

There are other metacognitive problems. For example, students often do not use hints and help, even if it takes a long time to solve a problem. In another case, having made a mistake, they linger on the correct answer for a long time, trying to understand its origin, instead of moving on. And sometimes they quickly look through the answer choices and answer the question correctly, but then stop and slowly analyze it, thus wasting time. The same data on students' behavior can be viewed in different ways. After analyzing students' behavior patterns for a week, one can tell which student tried to "outsmart" the system (this student will have additional classes), which day of the week was the least effective, and which lesson was the most boring. To determine a student's behavior pattern, it is important to collect the correct data.

It is not appropriate to ask students direct questions in such a situation. One cannot ask students: "Are you lying to me now? Yes / No". Instead, one needs to observe students' behavior, record their actions, analyze the sequence of their responses, or even use the PrintScreen function. There are also turnkey solutions offered by Neuromation using synthetic data. The main idea of synthetic data is to model the missing data to create a model of students' behavior. In other words, one needs to model a dataset for a training set of deep neural networks. Such a neural network will be able to predict students' behavior patterns depending on their specific actions during the test procedures.

Big data allow:

- creating methods adapted for a large number of students;
- personalizing educational materials;
- choosing a teacher.

4   Conclusion

Nowadays, distance education is not only popular but also quite a promising form of education. However, to use it with maximum efficiency, its technical and theoretical basis must be up to the proper level. Moreover, the process of teacher-student interaction should be characterized by mutual interest and the use of artificial intelligence elements.

This article discusses the possibilities of using artificial intelligence in education. Of particular importance are such features as adaptive learning using artificial intelligence elements, personalized learning, automatic assessment, interactive learning, students assessing their teachers, smart campuses, and smart agents.

Reference list:

1. Murphy R. F. *Artificial Intelligence Applications to Support K–12 Teachers and Teaching. A Review of Promising Applications, Opportunities, and Challenges*. Perspective. RAND Corporation, 2018.
2. Chetyrbok P.V. Artificial Intelligence in Distance Education. *Distancionnye obrazovatel'nye tekhnologii Materialy III Vserossijskoj nauchno-prakticheskoj konferencii.* [Distance Education Technologies. Proceedings of the 3rd All-Russian Scientific and Practical Conference], 2018, pp. 91-95.
3. Akinshina G. V. *Design Methodology Development of a Secure Web-based Data System as Exemplified by Distance Education System. Infokommunikacionnye tekhnologii* [Information and Communications Technologies], 2008, v. 6, no. 1, pp. 56-71.
4. Haykin S. *Neural Networks*: the Complete Book. 2nd ed. Tr. from English. Moscow, VilLajams Publ., 2006. 1104 p.
5. Robert I. V. *Theory and Methodology of Informational Support of Education*. Psychological, Pedagogical and Technological Aspects. Moscow, BINOM Publ., 2014. 400 p.
6. The Market for *Artificial Intelligence in the US Education Sector in 2018-2022*. Retrieved from: https://www.technavio.com/report/artificial-intelligence-market-in-the-us-education-sector-analysis-share-2018?utm_source = usa1 & utm_medium = bw_wk41 & utm_campaign = businesswire).
7. Oztyurk A., Ajdyn S. *Segmenting Students in an Online Learning Environment. Online Conference of Open and Flexible Higher Education*. 2015.
8. Dorogov A.Ju., Alekseev A.A. *Fast Neural Networks* // Proceedings of Seventh International Conference on Advanced Computer Systems (ACS-2000) Poland, Szczecin, October 2000.- P.267-270.
9. Dorogov A.Yu. *Parametrical and Topological Plasticity of Multilayer Neural Networks* // Processing of 4-th International Conference "New Information Technologies" (NITe'2000) Minsk, Belarus 5-7 December, 2000. - Minsk, 2000. - Vol. 1. - P. 15-19.
10. Dorogov A.Yu. *Estimation of Multilayer Neural Network Plasticity* // Eleventh IF AC International Workshop Control application of optimization CAO'2000 Perga-mon An Imprint of Elsevier Science Oxford, UK. - 2000.- VI.- P. 81-85.
11. Dorogov AJu. *Plasticity of Multilayer Neural Network* // First international conference on mechatronics and robotics: Proceeedings (M&R'2000) St-Petersburg: NPO Omega BF Omega, 2000. May 29-June 2.- 2000.-V1.- P.33-38.
12. Hopfield J.J. *Neural Networks and Physical systems with emergent collective computational abilities* // Proc. Nat. Sci. USA. 1982. - V.79. - P. 2554-2558.