Modeling the construction of an individual learning path

D M Filippov¹, A N Kazak¹, D V Samokhvalov² and N P Shamaeva³

¹ V.I. Vernadsky Crimean Federal University, Prospekt Vernadskogo 4, Simferopol, Republic of Crimea, 295007, Russia
² Saint-Petersburg Electrotechnical University “LETI”, Professor Popov Street, 5, St. Petersburg, 197376, Russia
³ Udmurt State University, Universitetskaya St. 1, Izhevsk, 426034, Russian Federation

E-mail: kazak_a@mail.ru

Abstract. In this paper, the most important aspect is touched upon - the principle of constructing an individual learning path in the system of open education. At the moment, there is no clear implemented construction method that satisfies both the learner and the teacher. The method proposed in this paper is based on the use of neural network technologies to determine significant input parameters that affect the success rate. At the first stage of determining an individual trajectory, proceeding mainly from theoretical ideas about the diagnosed set, a “draft” version of the list of parameters is formed. This option includes parameters that, according to the expert, should reflect the individual psychological differences of the subjects. The definition of a “draft” version of the initial set of diagnostic features is a difficult formalized task.

1. Introduction
In this paper, the most important aspect is touched upon - the principle of constructing an individual learning path in the system of open education. At the moment, there is no clear implemented construction method that satisfies both the learner and the teacher. The method proposed in this work finds a compromise of the problem posed: on the one hand, it does not burden the student with exhausting tests, and on the other, it does not completely shift the problem of constructing trajectories to the teacher [1, 2, 3, 4].

2. Materials and methods
In the paper, there were used general scientific and special research methods to solve mentioned aim. There are graphical analysis, statistic, techno-economic, expert estimation method. It was studied theoretical and applied papers on the research topic [1-26], also official statistics data. For developing the neural network used Matlab and Nnstar module.

3. Results
The principle of constructing an individual trajectory is proposed as follows: based on the predicted success rate of passing each module for a particular student, the modules are arranged in a chain in descending order of this indicator, which corresponds to the transition from simpler to more complex modules for this student. The method proposed in this paper is based on the use of neural network technologies to determine significant input parameters that affect the success rate. [6, 7, 8, 9, 10].
At the first stage of determining an individual trajectory, proceeding mainly from theoretical ideas about the diagnosed set, a “draft” version of the list of parameters is formed. This option includes parameters that, according to the expert, should reflect the individual psychological differences of the subjects. The definition of a “draft” version of the initial set of diagnostic features is a difficult formalized task.

Assessment of the selected parameters was carried out expertly, however, at the next stage of development, the use of psychodiagnostic tests and the results of the final control in the form of tests formed by the training system are implied. Such a system should reduce the time for creating training modules, control knowledge at various stages of training, and ensure the storage of material of the most common (hierarchical) structure [11,12,13,14,15].

The features of the developed system are the presentation of theoretical material in the form of a hierarchy of sections, some of which can be weighted (laborious) and assigned to a certain discipline (for example, power electronics) through a type of theory; can have practical tasks of various types, the result of which for each user is logged, where not only the result is recorded, but also the number of attempts and the date of execution. There are interdisciplinary relations, the connection of students with their courses, the composition of the courses [16,17,18].

The following parameters were selected as parameters of the modules and the student.

Student Options:

- full name;
- attentiveness;
- perseverance (endurance);
- independence;
- speed of perception;
- mentality;
- gender;
- initiative;
- organization.

Module parameters:
- course volume (hour);
- volume of lecture material (hour);
- the amount of practical work (hour);
- independent work (hour);
- number of illustrations;
- number of definitions;
- complexity (for a typical representative of the target group).

A separate database has been created to store data on students and modules. At the moment, it is an independent product along with software for working with it and is a specialized pre-processor. This approach is primarily due to the fact that initially the success of the new method for constructing individual learning paths was not known. Its introduction into an already working system was not advisable, the parameters of the students were determined expertly, and therefore, entering these data into the existing database is much more complicated, i.e. to requires its restructuring, in addition, the neuroimulator is not able to work with some databases and related tables.

The entire population is represented by the entire multitude of students studying electrical engineering disciplines from the moment they were last modified by one of the selected experts. During the work, statistics were collected for the student and test sample among students of the V.I. Vernadsky Crimean Federal University. To solve the problem, 78 students were selected from the general population. The selection criterion was students of the last two years of study who completed all the
modules according to the standard program (those who did not complete all the modules in the alternative form of passing or receiving an individual assignment).

45 students were randomly selected for training the network (248 examples), the remaining 33 students (112 examples) made up a test data set. To train the network, the number of examples must satisfy one of the rules $G > 2(b + l); G > (b + l)^2; G > 2(b + l),$ where $G$ is the number of examples for training the network, $b$ is the number of learner parameters, $l$ is the number of module parameters. It is believed that the higher the criterion that the set for training the network corresponds to, the more global the network will be (however, conflicting examples and methods for eliminating them should be taken into account). When distributing the examples to training and test, taking into account the criterion and preference for training the network, the parameters were distributed in a ratio of 1:2. The selection was carried out by selecting every third student, that is, every 6 records through 12 from the resulting table sorted by student code.

To simplify data entry, an automatic transition to a new parameter, “hot keys” and protection from incomplete information are provided.

The connection of the tables with the parameters of the student and the parameters of the module is carried out according to the “many to many” scheme by means of the communication table with the following fields

The model built in the Neural Network Toolbox (NNT). From the currently available neural network software products, this product is distinguished by the availability of targeted simplification of the neural network for the subsequent generation of a verbal description. It is the presence of developed capabilities to simplify the network, together with the construction of its verbal description, that gives the proposed product new consumer properties.

The model showed the best results when using the hyperbolic tangent function for the input, output layers and the hidden layer on which 3 neurons are located. As the learning algorithm, an elastic propagation algorithm was chosen. After training, a neural network test was run on a test set in order to calculate the average model error. In order to be able to compare the results of the two models, the model error was calculated using the following formula:

$$error = \frac{|V - V_n|}{V} \cdot 100\%$$

where $V$ - the actual value of the output parameter; $V_n$ - obtained forecast value.

The error was estimated according to the following principle: we consider the model suitable for building a forecast if the error is not more than 5%, while accurate results are achieved when the error value is less than 3%.

The error of the model built in the NNT relative to the real value is 2.89%. We consider this model suitable for forecasting.

Next, a study was made of the significance of the parameters using the built-in function of the NNT. The results, sorted in descending order, are presented in figure 1.

The rank of the significance of the input parameters for each network is determined, which will allow a more adequate analysis of the influence of the parameters on the output field. In most cases, the ranks of significance coincide with the average values of significance of the parameters, but sometimes they differ. The average value can be small (averaged over all networks), nevertheless, the rank shows the great importance of this parameter. The rank more precisely determines the order of the parameters according to the degree of their influence on the output field.

It can be seen from the figure1 that the most significant parameters are:

- initiative;
- organization;
- gender;
- date of birth in days;
- mentality;
number of illustrations;
complexity.

The minimum relevant parameters are:

- volume of practical exercises (hour);
- course volume (hour);
- attentiveness;
- volume of lectures (hour);
- perseverance.

Figure 1. Significance of parameters characterizing students.

From the results obtained, it is obvious that in the case of using elements of open education the concepts of “course volume”, “perseverance” (the ability to continuously assimilate more information) and “mindfulness” (the ability to assimilate details and basic aspects “on the fly”) are erased that first of all, it is connected with the features of open education - individual dosing of information and the lack of a rigid time frame (the student himself determines when and how much information he wants to study, guided by only the maximum time allotted for the module, which, as a rule, significantly exceeds the average required).

As for the parameters, initiative, organization, mentality, gender and date of birth, their significance is explained by the concept of professional suitability. It is proved that to work in certain professions a person must possess a number of qualities (this concept is most widely used in the selection of personnel
in law enforcement agencies). Such a parameter as the number of illustrations is a criterion, first of all, for visibility and, consequently, for material availability, which contributes to a better and much faster assimilation.

The principle of constructing an individual trajectory is assumed as follows: based on the predicted success rate of passing each module for a particular student, the modules are arranged in a chain in descending order of this indicator, which corresponds to the transition from simpler to more complex modules for a student.

The effectiveness of the method allows the transition to the next round of system development:

- creating a gateway for exporting data to create new neural networks;
- restructuring the database of the training system, taking into account the parameters of students and modules (and also taking into account the fact that for different courses this set will be different);
- creating a tool for importing verbal descriptions of a neural network into a training system with its subsequent use as a model for constructing trajectories.

The proposed method for determining the individual trajectory will allow to determine:

- a mechanism for constructing an individual trajectory based on the forecast of the success rate of the module;
- tests to determine the parameters of students and their implementation in the system in order to automate the process of collecting parameter data;
- a high percentage of correctly defined examples (from 70.8% to 84.3% across five networks) allows us to state the possibility of using this method for forecasting;
- the most significant input parameters affecting the success rate, most of which show a fairly stable effect (included in the top five most significant parameters in each of the trained neural network).

4. Conclusions
Based on the testing program included in the training system, comprehensive testing of students was conducted. The developed system is introduced into the educational process. In addition to the direct use of the computerized training system in educational institutions, it can be used at industrial enterprises for certification of personnel, in the selection of personnel. The individual components of the system can be used for making presentations, technical documentation, descriptions of technological processes.

Acknowledgments
The reported study was funded by RFBR, project number 20-08-00962

References
[1] Akhavan P, Ale E N, Fetrati M A and Pezeshkan A 2016 Major trends in knowledge management research: a bibliometric study Scientometrics 107(3) 1-16
[2] Bansal M, Gimpel K and Livescu K 2014 Tailoring Continuous Word Representations for Dependency Parsing In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics 2 809-15 (Baltimore, Maryland Association for Computational Linguistics)
[3] Baptista D and Morgado-Dias F 2013 A survey of artificial neural network training tools Neural Computing and Applications 23 609-15 10 1007/s00521-013-1408-9
[4] Baydin A G, Pearlmutter B A, Radul A A and Siskind J M 2015 Automatic differentiation in machine learning: a survey arXiv:1502.05767 [cs]
[5] Bengio Y, Ducharme R, Vincent P and Janvin C 2003 A Neural Probabilistic Language Model J. Mach. Learn. Res. 3 1137-55
[6] Bozkurt A, Ozbek E A, Yilmazel S, Erdogdu E, Ucar H and Guler E 2015 Trends in Distance Education Research: A Content Analysis of Journals 2009 - 2013 The International Review of Research in Open and Distributed Learning 16(1) 1-19

[7] Chen D and Manning C 2014 A Fast and Accurate Dependency Parser using Neural Networks Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing 740-50 (Doha Qatar Association for Computational Linguistics)

[8] Chetyrbok P V 2017 Preliminary systemic decomposition of big data for their classification using vector criteria dynamic management model of innovations generations Proceedings of 2017 20th IEEE International Conference on Soft Computing and Measurements (Saint Petersburg: Electrotechnical University 'LETI') 762-4

[9] Dorofeeva A 2018 Simulation of Advertising Activity in Tourist Business Proceedings of the 2018 International Conference Quality Management Transport and Information Security Information Technologies (IT&QM&IS) 836-8

[10] Kazak A N, Chetyrbok P V and Oleinikov N N 2020 Artificial intelligence in the tourism sphere IOP Conference Series: Earth and Environmental Science 421 042020

[11] Kazak A N, Gorobets D V and Samokhvalov D V 2020 Application of Simulink and SimEvents Tools in Modeling Marketing Activities in Tourism Lecture Notes in Electrical Engineering

[12] Kazak A N and Shamayeva N 2018 Separate Aspects of Smart Cities Security Proceedings of the 2018 International Conference Quality Management, Transport and Information Security, Information Technologies (IT&QM&IS) (Saint-Petersburg: Saint Petersburg Electrotechnical University 'LETI') 216-8

[13] Lukyanova Ye Yu 2017 BSC-oriented process management system task formalization for resort and spa sphere economic units Proceedings of 2017 20th IEEE International Conference on Soft Computing and Measurements, SCM 2017 (Saint-Petersburg: Saint Petersburg Electrotechnical University 'LETI') 754-6

[14] Means B, Toyama Y, Murphy R, Bakia M and Jones K 2009 Evaluation of Evidence-Based Practices in Online Learning Structure 15(20) 1-94

[15] Ostaev G Y, Kotlyachkov O V, Markovina E V, Kravchenko N A, Mironova M V, Nekrasova E V, Konina E A and Alexandrova E V 2019 Integrated budgeting at agricultural enterprises: functionality and management decision making Amazonia Investiga 8(22) 593-601

[16] Ostaev G Y, Suetin S N, Frantsisko O Yu and Alexandrova E V 2020 Assessment of the effectiveness of the management of agricultural organizations in management accounting Amazonia Investiga 9(28) 260-71

[17] Skiadas C H and Skiados Ch 2019 Chaotic Modelling and Simulation Analysis of Chaotic Models Attractors and Forms (Boca Raton: CRC)

[18] Sokolowsky J A and Banks C M 2011 Principles of Modeling and Simulation: A Multidisciplinary Approach (New York: Wiley)