Perseptron for assessing the students’ task performance in learning a foreign language according to the competencies using bi data

N A Prom, E A Litvinova, A V Shokhnekh, T G Yovanovich, N I Lomakin

Russia, 400005, Volgograd, Prospect V I Lenin, 28

natalyprom77@mail.ru, srdjan@live.ru, shokhnekh@yandex.ru, elen-gordienko@yandex.ru, tel9033176642@yahoo.com

Abstract. The article presents the Perceptron neural network program developed by the authors, which enables assessing in points the quality of the task performed by students in a foreign language. The theoretical basis for the artificial intelligence systems applied in the learning process has been considered. The studies have shown that the online education market has grown in the world, and the trend marked persists. The trends in use of neural networks in learning have been summarized. Analysts noted that the use of artificial intelligence systems has been becoming more widely used in the learning process. It has been hypothesized and proved that a neural network makes it possible to give a grade for the quality of a control text read by a student according to the requirements of competencies using the BI data resulted from processing with an artificial intelligence system. The algorithm proposed by the authors uses the capabilities of the PhyPhox mobile application and the Deductor platform. The algorithm is based on the following scheme. The teacher records a piece of a text into a file on the PhyPhox mobile application; its digitized version is then used as a reference to compare the quality of the assignment done by a Master student. The resulting reference file is exported to XL and then imported into the Deductor for further processing. The students do the same; their control files are paired with the reference, combined in XL and imported into the Deductor for processing. In the final, based on the comparison of the coefficients of correlation between the control and reference files, the Perceptron program assigns a score corresponding to the “excellent,” “good,” “satisfactory” and “not satisfactory” grades. The neural network formed is trained on data sampling, where they add semester points and competence points. The relevance lies in the fact that the program enables automating the learning process and replacing the teacher’s work at the stage of the assignment control according to the competencies, which is of practical importance.

1. Introduction

The study presented was conducted due to the fact that the market of online education is constantly growing in the world, and this trend persists. In order to automate the work of a foreign language teacher (English), an attempt was made to computerize the process of assessing the quality of students’ reading a text according to the requirements of competencies at the end of the control week.

The relevance is determined by the fact that the algorithm proposed automates the learning process and replaces the teacher’s work at the stage of assignment quality control according to the competencies, which is of practical importance.

It has been hypothesized and proved that a neural network applied makes it possible to assess the quality of student’s reading the control according to the requirements of competencies, using the BI data obtained as a result of processing by the artificial intelligence system.

2. Results and discussion
To develop the AI-system program for digital control and assessment of the compliance of a student’s oral speech according to the requirements of competencies, initial data were taken, i.e., two digitized sound files, one of them was a reference file recorded by the teacher.

2.1. Results

The algorithm proposed by the authors used the capabilities of the PyPhox mobile application and the Deductor platform [1]. The algorithm was based on the following scheme. The teacher recorded a piece of a text to a file on the PhyPhox application, which was then used as a reference for comparing the quality of the test done by a Master student. The resulting reference file was exported to XL, then into the Deductor for further processing. The students came in the same way; their control files were combined in pairs in XL and imported into the Deductor for processing.

In the final, based on the comparison of the correlation coefficients between the control and reference files, the Perceptron program, previously trained on data sampling, assigned a score corresponding to the marks “excellent,” “good,” “satisfactory” and “not satisfactory.”

The sound track was digitized according to the parameters: time (s), Period (s) and Frequency (Hz) with the help of the free-of-charge mobile application Phyphox (Fig. 1).

Figure 1. A text fragment, the quality of reading the text was assessed; and a mobile phone screen with the file recorded

A text fragment, the quality of reading the text was assessed; and a mobile phone screen with the file recorded.

The voice file obtained in a digital format for control was exported to an XL file and then into the Deductor (Table 1).

Table 1. A fragment of the initial data, i.e., two digitized files in XL
The next step of the algorithm was processing of the data imported into the Deductor. In particular, the digitized files were aligned in length (by filling in gaps), processed by quantization for subsequent visualization in a multidimensional diagram (by grouping) and used to calculate multiple correlation coefficients for subsequent comparison of their identity in the Perceptron in order to assign an appropriate estimate (Table 2).

Table 2. A fragment of initial files containing 996 lines in the Deductor

|    |       |       |       |       |       |       |       |
|----|-------|-------|-------|-------|-------|-------|-------|
| A  | B     | C     | D     | E     | F     | G     | H     |
| 1  | 0.197 | 0.0382| 261.7486| 0.146 | 0.000209| 4790  |       |
| 2  | 0.228 | 0.0383| 263.1868| 0.187 | 0.006367| 157,049|       |
| 3  | 0.248 | 0.0383| 263.1868| 0.228 | 0.011696| 855,3571|       |
| 4  | 0.284 | 0.03779| 264.6409| 0.269 | 0.003946| 253,4392|       |
| 5  | 0.331 | 0.03841| 260,3251| 0.308 | 0.00207| 483,8384|       |
| 6  | 0.365 | 0.03862| 258,9198| 0.349 | 0.00207| 4790  |       |
| 7  | 0.406 | 0.04008| 249,4792| 0.389 | 0.00207| 4790  |       |
| 8  | 0.444 | 0.05219| 191,146| 0.626 | 0.00207| 4790  |       |
| 9  | 0.486 | 0.04551| 219,7248| 0.865 | 0.00146| 6842,857|       |
| 10 | 0.528 | 0.03841| 260,3251| 0.906 | 0.00146| 6842,857|       |
| 11 | 0.565 | 0.03039| 323,6486| 0.945 | 0.00188| 5322,222|       |

2.2 Discussion
The files loaded in the Deductor were aligned in size using the “Fill-in gap” function and processed by the quantization method for later visualization in a multidimensional diagram.

The quantization in signal processing is known to be a partition of the range of signal reference value into a finite number of levels and rounding of these values to one of the two levels closest to them. It should be noted that in this case, the signal value might be rounded to either the nearest level or to a lower or higher level of the nearest ones, depending on the encoding method. In practice, this provided an effect of “compressing” the data (Table 3).

Table 3. A fragment of data after quantization
The authors found it appropriate to visually compare the resulting sound files digitized that were quite different in parameters, i.e., time (s), period (s) and frequency (Hz) (Fig. 2).

In the algorithm presented, the initial data was transformed and used in the BI format. So, the voice files in digital format received from students to be assessed were alternately exported into an XL file that already contained a reference file, then imported into the Deductor, processed there by quantization and compared. Thus, the information was converted to the BI, later used in the algorithm developed.

Business intelligence (abbreviated as BI) is the designation of computer methods and tools for organizations that translate transactional business information into a human-readable form suitable for business analysis and means for mass work with such information processed [2].

In order to compare the quality of a control file with the reference one, a neural network in the Deductor was formed to conduct the correlation analysis (Fig. 3).
Figure 3. Values of the correlation matrix

The values obtained were accumulated in a table. The correspondence of the correlation coefficients values to the teacher's estimates obtained in the control is presented in Table 4.

Table 4. The correspondence scale of the assessment level to the correlation parameters in the Perceptron training sample

| Period (s) | Frequency (Hz) | Competence Score | Semester Score | Evaluation | Evaluation_OUT | Evaluation_ERR |
|------------|----------------|------------------|----------------|------------|----------------|----------------|
| 0.750      | 1              | 35               | 40             | 3.5        | 90             | 5.97036517906575E-7 |
| 0.0142     | 0.0541         | 25               | 30             | 75         | 75             | 5.682354364545420E-7 |
| 0.068      | 0.014          | 15               | 20             | 50         | 50             | 5.2587925969675E-11 |
| 0.082      | 0.0025         | 5                | 10             | 40         | 40             | 3.747394781388E-7 |

The Perceptron formed during the research study has five neurons in the input layer, i.e., Period (s)_0, Period (s)_1, Frequency (Hz), competence score and semester score; one hidden layer with two neurons; in the output layer there is one output parameter-the estimate (Fig. 4).

Figure 4. The Perceptron graph

With reference to the field of education, the term “competence” means a set of legal rights and obligations (duties) of state and/or authorized non-state bodies, institutions or organizations for making legal decisions in the course of legislative or law enforcement activity in education [3].

The score according to the competencies of a graduate student was put down by the teacher according to the “Fund of appraisal funds” of the Steering document for the discipline “Business Foreign Language (English)” in the categories “know,” “know how” and “have skills” [4].

The federal and state authorities’ competence established in Articles 28 and 29 of the Education Law of the Russian Federation in federal subjects of the Russian Federation is exhaustive and can be changed only by law.

The semester points were given to the graduate by the teacher according to the control list in the dean's office.
Thus, in the algorithm proposed for scoring the test, a trained neural network, the Perceptron [5] was used. When using the “what-if” function based on the parameters entered the neural network, we finally got an estimate (Fig. 5).

![Figure 5. Scoring using the “what-if” function](image)

Thus, the Perceptron program for assessing the task performance by students in a foreign language according to competences based on the BI-data according to the results of the neural network has been successfully formed.

It should be noted that the program presented differs from the “Neural network for assessing the students' competence in the discipline of Organization of entrepreneurial activity” previously proposed because it is more advanced regarding the automatic generation of a sound file and effective operation of the data transmission system [6].

The studies showed that in modern conditions the use of artificial intelligence systems in the learning is increasing.

For example, Parla startup raised € 2 million investment for the development of a virtual English teacher. The developers threatened to replace teachers and tutors with artificial intelligence soon. The Parla developers predicted the death of traditional teaching methods. Having replaced the English teacher, they plan to release applications for the study of physics, mathematics, biology and other subjects.

The analysts at J’son & Partners Consulting found out that the online education market has grown by 23% in the world and according to Global Market Insights, makes $ 165 billion or about 3% of the global education market for the past five years.

The study of the Russian market of online education and educational technologies established that the market will grow annually in the country by at least 5%. In Russia, in 2016, the language learning segment using games was worth $ 315.7 million. According to Skyeng, English accounts for 76% of the Russian market for learning foreign languages, German 10% and French 7%. According to Technavio, the volume of the global digital English language learning market in 2016 was $ 4.83 billion and is projected to rise to $ 13.8 billion by 2021 [7].

The main purpose of the Parla program was to create learning techniques that are comparable with speaking of native speakers. So, you can understand the properties of the language, synonyms, wordplay and collocations [8].

The experts’ opinions regarding the future of such technologies vary. We observed several trajectories of the development of artificial intelligence applied in education:

- machine learning and neural networks for the compilation of educational content for the users’ interests;
- selection of teacher-student pairs using computer vision systems and machine intelligence; and
- cognitive computations [9].

Cognitive computations based on neural networks and deep learning use knowledge of cognitive science to create systems that simulate human thought processes. However, instead of focusing on a unique set of technologies, cognitive computing encompassed several disciplines, including machine learning, natural language processing and human-computer vision and interaction [10, 11].
A promising direction is the use of innovative IBM open source projects [12]. Deep machine learning is in the trend, for example, Tensor Flow that is an open source software library for numerical computation using data flow graphs, in other words, the best way to create deep learning models [13, 14].

3. Conclusions
Based on the study conducted, we can draw the following conclusions.
Firstly, the use of artificial intelligence systems has broad prospects; therefore, research of automating the control of students' knowledge in a foreign language based on a neural network is important [15, 16].
Secondly, it was hypothesized and proven that based on the neural network, the quality of the control text read by a student can be assessed, according to the requirements of the competencies, using the BI data obtained as a result of processing by the artificial intelligence system.

References
[1] Basegroup Advanced analytics without programming 2018 URL: https://basegroup.ru/deductor/description
[2] Business Intelligence 2018 URL: https://ru.wikipedia.org/wiki/Business_Intelligence
[3] Federal Center for Educational Legislation 2018 URL: http://www.lexed.ru/obrazovatelnoe-pravo/knigi/yagofarov2005/422.php
[4] Steering document “Business Foreign Language (English),” Field of education 09.04.01 Informatics and computer engineering. Programs “Information and software of automated systems,” “Methods of analysis and synthesis of design solutions” and “Information and software of production automated systems,” Master's level of training. Full-time form of education. The term of study is normative. Volgograd 2017 32 p.
[5] Serebryakova S V, Milostivaya A I 2017 Semantic emergence as a translation problem Vestnik Volgogradskogo Gosudarstvennogo Universiteta-seriya 2-yazykoznanie 16 (3) 48-57
[6] Ergunova O T, Lizunkov V G, Malushko E Yu, Marchuk V I, Ignatenko A Yu 2017 Forming the system of strategic innovation management at the high-tech engineering enterprises IOP Conference Series: Materials Science and Engineering 177 (1) 012046
[7] Perseptron 2018 URL: https://ru.wikipedia.org/wiki/
[8] Lomakin N I, Lukyanov G I, Plaksunova T A, Litvinova A V, Loginova E V, Kozlova E A, Skobor E A, Odintsov A E, Zavyalova M N, Korochinskaya S V 2017 Certificate of registration of a computer program 2017618928 Russian Federation. Neural network for assessing the students’ competence in the discipline Organization of entrepreneurial activity in the field of education 38.04.01 Economics according to the program Economics of a firm and industry markets VSTU
[9] Shamne N L, Shishkina E V 2017 German popular scientific medical online media: structural and functional aspects Vestnik Volgogradskogo Gosudarstvennogo Universiteta-seriya 2-yazykoznanie 16 (2) 143-151
[10] Malushko E, Maletina O, Lizunkov V, Tsybaneva V 2016 Use of virtual learning system for educating students with disabilities and special needs International Multidisciplinary Scientific Conferences on Social Sciences and Arts 481-487
[11] Fedotov N G, Moiseev A V, Syemov A A, Lizunkov V G, Kindaev A Y 2017 New Methods of Three-Dimensional Images Recognition Based on Stochastic Geometry and Functional Analysis IOP Conference Series: Materials Science and Engineering 177 (1) 012047
[12] How artificial intelligence replaces living teachers 2018 URL: https://www.rbc.ru/newspaper/2017/09/21/59c25e659a7947f3f6210ac80
[13] Tim Jones M 2018 A guide for beginners in the fields of artificial intelligence, machine learning and cognitive computing URL: https://developer.ibm.com/articles/cc-beginner-guide-
machine-learning-ai-cognitive/

[14] IBM open source projects 2018 URL: https://developer.ibm.com/open/code/

[15] Lizunkov V, Politsinskaya E, Malushko E, Kindaev A, Minin M 2018 Population of the world and regions as the principal energy consumer International journal of energy economics and policy 8 (3) 250-257

[16] Awesome Tensor Flow 2018 URL: https://github.com/jtoy/awesome-tensorflow