III. ALTERNATIVE METHODS, FORMS AND TECHNIQUES IN DISTANCE LEARNING

USE OF ARTIFICIAL NEURAL NETWORKS IN THE SELECTION OF EDUCATIONAL CONTENT ON AN E-LEARNING PORTAL

Barbara Dębska
Rzeszów University of Technology, Department of Computer Chemistry
al. Powstańców Warszawy 12, 35-959 Rzeszów, Poland
Email: bjdebska @prz.edu.pl

Agnieszka Kubacka*
Institute of Technology in Department of Informatics,
State Higher Vocational School in Krosno,
Wyspiańskiego 20, 38-400 Krosno, Poland
Email: kubackaagnieszka@interia.pl
*Corresponding author

Abstract: *An individualized learning path is a new solution that has been introduced in didactic materials made available to students on the educational portals of our universities. This approach stems from the fact that students participating in the educational process, on the one hand, have different needs and expectations, and on the other hand, different possibilities and limitations in acquiring knowledge. The conducted research allowed recognition of the differences and the selection of the most appropriate of the three learning paths offered. Indication of the learning path is made by means of a classification system, for the construction of which artificial neural networks were used. On the basis of the tests carried out, it was shown that when selecting the learning path, the best results are provided by a multi-layer network, with one hidden layer that contains 9 neurons. The network was taught in 50 epochs, the activation functions of the hidden and output layer neurons were hyperbolic tangent and linear function respectively. Over 98 percent correctness was achieved in the classification of new students starting the education process. The innovation of the proposed solution is to demonstrate in practice the possibilities of individualizing the student education process and thus its adaptation to the educational needs and competency gaps of participants in this process.*

Keywords: educational portal, learning path, neural networks, classification of students

INTRODUCTION

Personalisation of internet portals has become an everyday reality. Internet users receive personalized advertisements, content in online publications, etc. An important factor in the current era of computerization, universal access to the Internet, and Lifelong Learning is the personalization of learning paths using multimedia intelligent education systems. A significant problem of such systems is the faithful reproduction of both the teacher's and student's activities in it, and the implementation of such software that would provide an adaptive teaching process. Education at a technical university places particular emphasis on practical application of knowledge, and the most commonly used method of education using an educational portal is blended learning. The goal of blended learning is to combine two modes of learning: face-to-face and on-line materials (e-learning). Time spent with the teacher is used primarily to engage students and allow them to gain interactive experiences. At the same time, participants can use the network rich in materials at any time of the day, from any place. This gives the learners with a plethora of other duties the flexibility they need in organizing their own learning process. Introduction of modern didactic means to the learning process and multi-sensory transmission of program content consolidates the knowledge acquired by students and makes them aware of the possibilities of using this form of education in the future (Baylari, Montazer, 2009; Anand, Mobasher, 2005).

The issue of selecting the learning content for the group of students discussed in this work concerns primarily the didactic materials made available on educational websites of universities. In this case, it is necessary to use methods that, in addition to knowledge acquired about the students, can take into account the different and often changing levels of knowledge they possess about particular topics, as well as the variable degree of its further acquisition. Systems designed to solve the problems of classification, in addition to being able to learn from examples, should also improve their performance as new experiences are accumulated, i.e. during the expansion of the student knowledge base that was built prior to the launch of the classification process. Therefore, it seems that the most effective classification algorithm that meets these requirements can be built using advanced data mining methods, including Artificial Neural Networks (ANN). Networks quickly learn to recognize user preferences in different data sets. ANN are characterized by the fact that they have the ability to map even the most complex functions. For this reason, non-linear models can be created easily and rather simply, without the need for the user to formulate hypotheses. This non-linear nature of the network increases the possibilities of its applications, which is especially important in the case of the need to select the right paths for new groups of students, which change with the following year's recruitment to the university. The article describes the operation of such an example of an experimental system that ensures the creation of personalized e-learning courses, allowing the selection of optimal learning paths.

1. NEURAL NETWORK

Neural networks allow control of the problem of multidimensional data, which in other methods of artificial intelligence makes it difficult to model nonlinear functions that have a large number of independent variables. In practice, neural networks themselves create the necessary models by automatically learning from representative data in which the dependency that is interesting to the user is hidden. Based on this data, the necessary structure of connections between data is automatically created in the network's memory, which takes the form of so-called weighting factors. On the basis of the thus created structure, the network performs all functions related to class forecasting, i.e. the use of the created model (Tadeusiewicz, 1999, Tadeusiewicz et al. 2007). Another important feature of artificial neural networks is resistance to errors and damage. This is manifested by the ability to act in spite of the existence of certain damage to the network and when the data analysed is disrupted, uncertain or incomplete. However, the most important feature that characterizes neural networks is the ability to learn (Duch, 2000; Kurzyński, 2008; Markowska-Kaczmar, Kwaśnicka, 2005).

In network learning, one of the most commonly used methods can be used:

- supervised learning (with a teacher) – consisting of networks providing examples of inputs and corresponding outputs,
- unsupervised learning (without a teacher) – characterized by the fact that the system itself has to discover the features that appear in the input data; in the learning process mechanisms are used to discover similarities between patterns and cluster identification,
- reinforced learning (with a critic) – the output patterns are not provided; however, there is an external observer who assesses the results of the network processing and uses them in the process of selecting weights.

The specificity of applying each of the mentioned methods lies in the fact that the result of the learning algorithm is used for changes (so-called tuning) of network parameters to improve its performance (Duch, 2000; Kurzyński, 2008; Markowska-Kaczmar, Kwaśnicka, 2005).

The most common method is supervised learning, which is implemented through the use of the information contained in the training set. Network learning is incremental (iterative). Learning rules refer to specific network parameters that are equivalent to the weights of synaptic connections in biological neurons. In a single step of network learning, the vector of traits *x* of the next teaching object is given to its input, the responses of the network *y* are observed, and after comparing it with the given value *d*, the appropriate correction of the network parameters takes place. These parameters are called weights. In the first step, their values are chosen randomly and then changed in such a way as to minimize the error between the response of the network *y* and the expected value *d*. Such a single step is called an epoch or a learning cycle. These epochs are repeated many times until the condition of stopping the algorithm is met. This condition may be the execution of a set number of epochs or the obtaining of the required quality of network operation by achieving a certain error threshold defined at the beginning.

The quality of network operation is most often measured by the mean square error. The network learns until such a time as the error value falls below a certain set value. It is also possible that both conditions must be met in order to stop the algorithm. After completing the learning, the values of the weights are remembered, and the network can be tested and also used to predict the class of new objects.

One of the tasks for which neural networks are used is classification, both non-reference and reference. Non-reference classification serves to analyse the structure of the surveyed collectives. Reference classification, on the other hand, consists in assigning each examined case to one of the classes. The number of these classes and their characteristics are known before the classification process begins. A network that has the task of classifying objects should receive on the input the values of the variables that describe them, and on the output the expected class identifier (e.g. number, label). The method of their interpretation depends on the type of classification problem and the way the output variable is represented (Lula, 1999; Lula, 2006; Kurzyński, 2008; Tadeusiewicz, 1993, Tadeusiewicz, 2007).

To solve the problem classifying of students for individual learning paths, one-way multi-layered networks were used, taught using the supervised method. Multi-Layer Perceptron (MLP) networks consist of a layer of input and output neurons. Between them there is at least one layer of hidden neurons that mediate the transmission of signals between the input nodes and the output layer. This layer is a collection of neurons that are not interconnected. Connections only occur between neurons in different layers (Markowska-Kaczmar, Kwaśnicka, 2005; Galushkin, 2007; Kurzyński, 2008).

A characteristic feature of an MLP network is the flow of information in one direction: from neurons located in the input layer, through hidden layers, to the neurons of the output layer. Neuron connections between individual layers are full, connected on an all-with-all basis. However, there may be a case when some connections do not occur (so-called partial connections). The advantage of a one-way network is that it always has stable behaviour (Tadeusiewicz, 1993; Wilde, 1997; Duch, 2000).

Teaching the neural network is based on the automatic search for such weighting factors in all neurons throughout the network, which guarantee the smallest value of the total error of forecasting committed by the network. During the learning process, this error is systematically reduced, resulting in a gradual improvement in the network's performance.

To create a neural network that aims to classify students who study on the e-Student portal (the virtual education portal of the Stanisław Pigoń State Higher Vocational School in Krosno, shared on the Moodle platform) and attend the course *Algorithms and data structures* (Dębska, Kubacka, 2014), the *STATISTICA Neural Networks* software was used. Its undoubted advantage is the built-in Automatic Network Designer module, which in most cases can effectively conduct the process of searching for the appropriate network model. All that needs to be done is selection of the type of network, the number of networks to be taught and the number of networks to be remembered, the activation function for the hidden and the output layer, and the range of weight reduction factors. The designer generates a set number of networks with various functions of neuron activation. The statistics of these networks can be analysed, then the best network can be selected and used for further research (it is possible to write the structure of the taught network in C/C++). The discussed module of the STATISTICA system was used in the described studies (Bishop, 1995; Welstead, 1996; Lula, 1999, Lula 2006).

2. DATA AND METHODS

The study whose results are discussed in this article~~,~~ was conducted on didactic materials prepared to assist in teaching the course *Algorithms and data structures*. The teaching was conducted using the blended learning method during the second semester of the first year of a *Computer Science* course. The choice was not accidental, because this is the first subject for which e-learning materials were developed and the students had already benefited from blended learning since 2010. The subject was divided into 5 thematic modules. Each module was in turn divided into lessons. Their number in the module depends on the complexity of the problem and the amount of material discussed during each unit. One lesson unit may include one or more topics. Each module ends with a test checking the student's knowledge, followed by classification, and a positive result provides the student with access to didactic materials from the next module. The content of didactic materials placed in the lessons corresponded, as to the degree of difficulty, to the requirements set in the sheet for the subject *Algorithms and data structures*, which documented what knowledge and skills should be obtained by a student who has achieved results at the assessment level of 3.0 (satisfactory), 4.0 (good) and 5.0 (very good). Selection of the subject of the lessons for subsequent learning paths was directed not only towards optimal adjustment to educational needs, but also the need to adapt the didactic materials to the competency gaps of the learning participants observed during classes conducted using the traditional method. In the conducted research on the personalized educational system, it was assumed that the selection of individual learning paths would take place in two stages. In the first stage initial selection of learning paths is carried out. The student's classification takes place on the basis of historical data, i.e. grades from subjects whose knowledge is necessary to study the current subject matter and the results of the placement test. Prior to the pre-classification process, these data were normalized using the *minimax* function. After initial acceptance, the student may proceed to independent study of the subject on the optimal learning path selected for him. As a result of the initial classification, 111 students were qualified to study on the first learning path, on the second path there were 57 students, while 29 students were assigned to the third path. Debska and Kubacka (2016) described the initial classification as well as the second and subsequent stages, which in this case were implemented using the cluster analysis method, with the Matlab system used for calculations. The conducted cluster analysis using the unsupervised method showed that three groups of people can be distinguished in the surveyed group of students, representing three well-separated learning paths. These results became the basis of the present study, whose aim was to develop a classification model using artificial neural networks, which can be used for classification of new students.

Data for classification using ANN were downloaded from the Moodle platform databases. These are the results obtained by students after the completion of the first part of the subject *Algorithms and data structures*, which discusses the subject of the time and computational complexity of algorithms. The objective of the classification algorithm was to assign each student to the appropriate class, on the basis of which training materials from the next section (RAM Machine) were to be made available to them. The contents contained therein are intended to meet their individual needs.

The input vector of the neural network consisted of 197 cases, each of them containing the following attributes:

- the mark obtained in the test,
- test solution time,
- number of approaches to solving the test,
- the number of the learning path which the student is currently on.

The output variable in the training set was the identifier of the class to which the student was assigned by a teacher based on the results that the students obtained during classes conducted using the traditional method. In this set of observations, 27 cases belonged to third class, 91 to second and 79 to first. The correct selection of the class was also confirmed by an unsupervised method of cluster analysis. In order to reduce the subjectivity of the assessment, the personal data of students was replaced with consecutive case numbers.

The initial parameters that were set during network design were as follows:

- minimum number of hidden neurons,
- maximum number of hidden neurons,
- error function,
- activation function of the hidden layer neurons,
- activation function of the output layer neurons.

Using these parameters, various models of artificial neural networks were created. Each of these networks was taught and tested on the same cases.

The training set was divided into three subsets: learner, testing and validation (in the ratio of 70:15:15). The first one is used to teach the network, the second for the final evaluation of the prognostic quality and selection of the best network (testing), and the third to verify the network learning process, which consists in controlling the effects of the learning algorithm over its duration (validation). Such a division of the training set, with an artificial subset allocation, which will not be used to teach the network, is a way to protect the network against over-matching. This subset is used to check whether the network has properly conducted the process of generalization of knowledge "received" from the training data (Statistica, 2009).

3. RESULTS OF THE NEURAL NETWORK LEARNING PROCESS

As a result of the operation of the generator, more than a dozen neural networks with different parameters were obtained. The best was achieved by a network containing 9 neurons in the hidden layer. The learning quality for this network is 99.28, testing quality 96.55 and the quality of validation is 100.00. The network was taught in 50 epochs, the activation function of the hidden layer of neurons has a hyperbolic tangent and the output layer is linear. Figure 1 presents a chart of the learning process for this network. On the x-axis, subsequent learning cycles are marked, and on the y-axis the participation of incorrect classifications. The learning and test sample chart shows that the lowest error value was achieved in the 50th cycle of this network's learning.

The second network, with only slightly worse parameters, is a network that also contains 9 neurons in the hidden layer. It differs from the previous network in the activation function. In this case, both functions are a hyperbolic tangent. The learning quality for this network is 98.56, while testing and validation are 96.55 and 100.00 respectively. Figure 2 shows a graph of the learning process of the network. It can be seen that the network needed 53 epochs to learn.

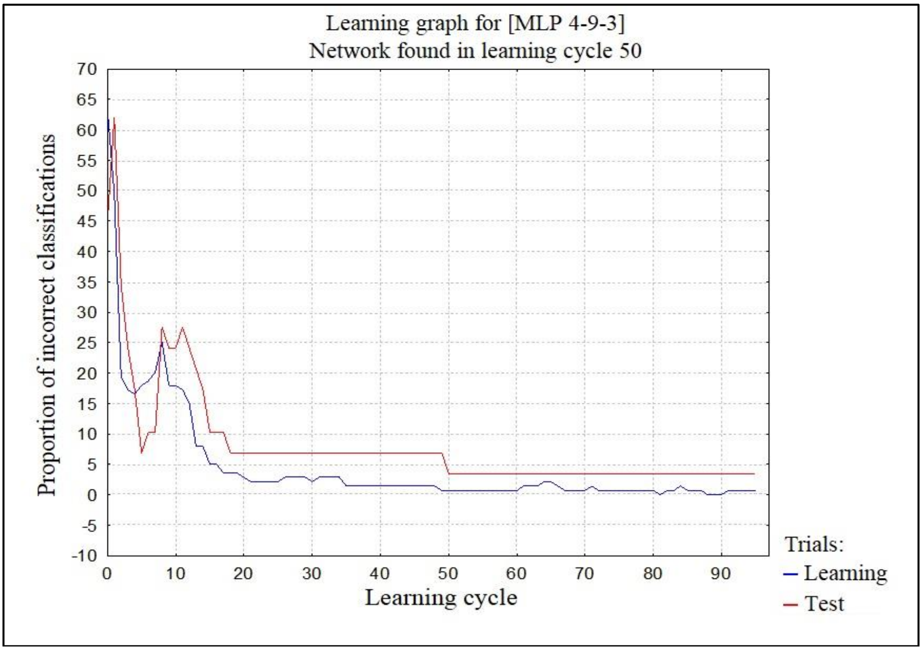

Figure 1. Learning graph for the best-performing network

Source: Own work

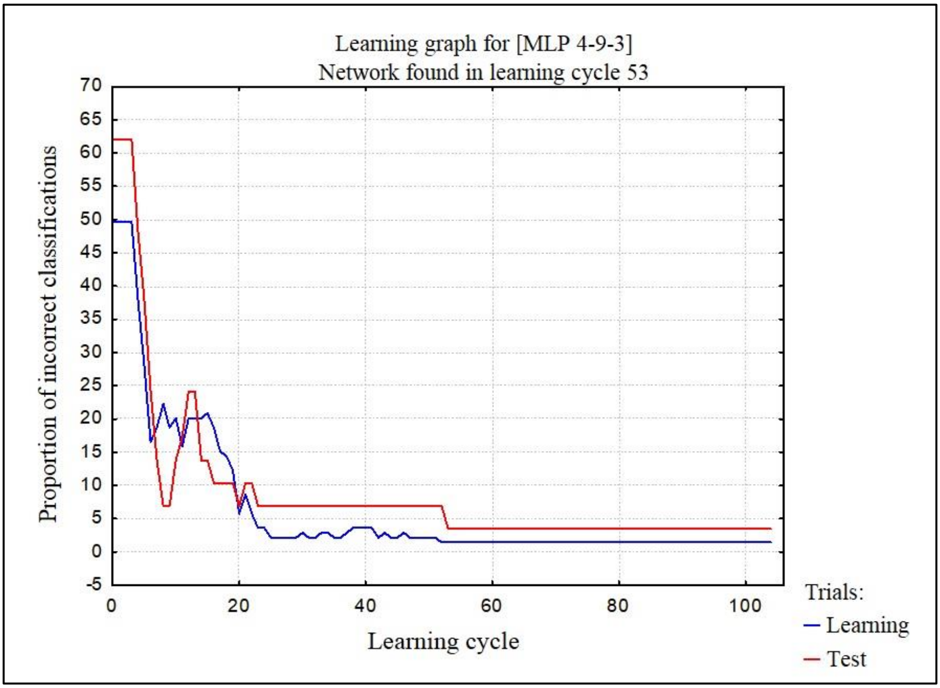

Figure 2. Learning graph for the second network.
Source: Own work

The next, third network, contained 8 neurons in the hidden layer, which are activated by the hyperbolic tangent function. The output layer neurons are activated by a linear function. The quality of learning, testing and validation for this network was analogous to network number 2. From the graph presented in Figure 3, it can be seen that the network needed 31 epochs to complete learning.

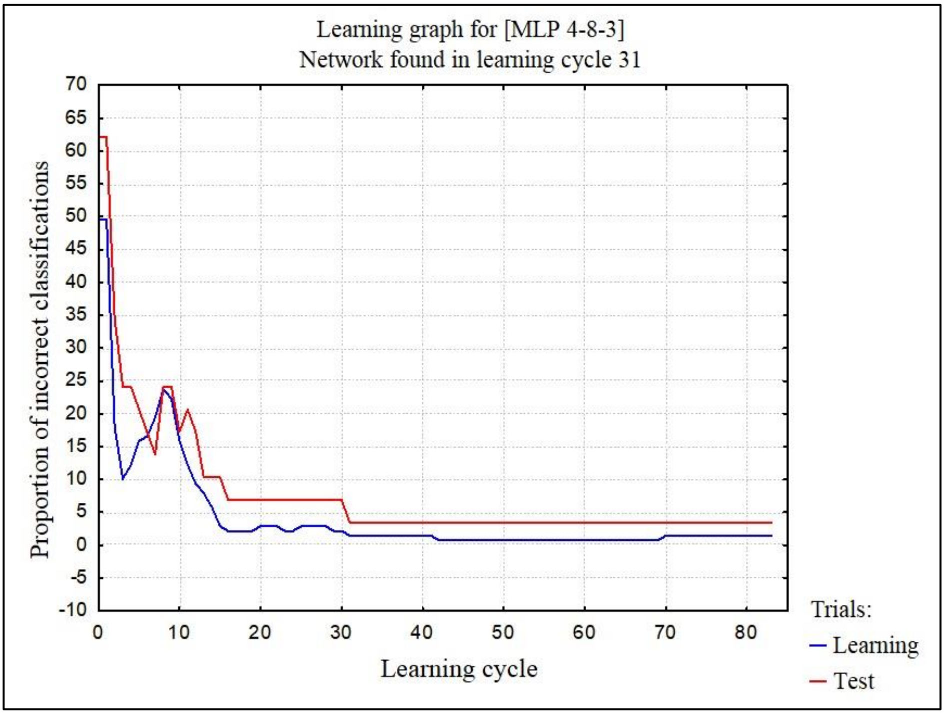

Figure 3. Learning graph for the third network

Source: Own work

The last of the analysed networks is characterized by the same quality parameters of testing and validation as in the previously described networks. However, it is characterized by the lowest value (slightly lower than the others) of the learning quality parameter among the discussed networks. It is 97.84. The network needed 26 learning cycles to learn (Figure 4).

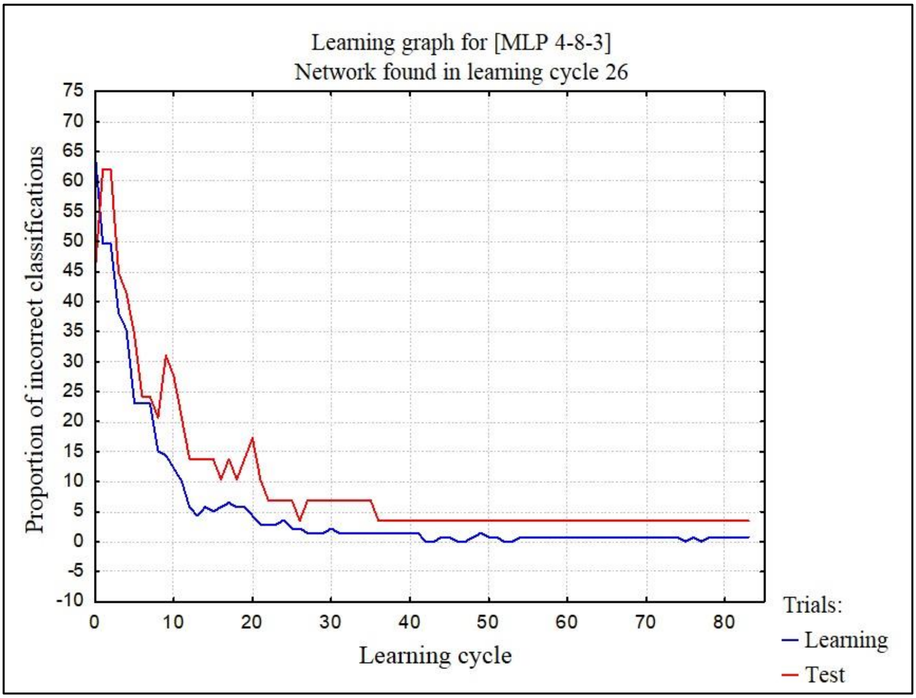

Figure 4. Learning graph for the fourth network

Source: Own work

4. NETWORK TESTING

In fact, the network's ability to generalize results is demonstrated by the correct result of its operation for new data. This is a confirmation of the legitimacy of implementing a given network. Therefore, external testing was used to check the performance of the discussed networks. All four networks discussed above were qualified for testing.

The data source used for the tests were also Moodle database tables. The data stored in them are the results obtained by students of three subsequent years, after completing four consecutive modules of the subject *Algorithms and data structures*:

- *Time and computational complexity of algorithms,*
- *RAM machine,*
- *Dynamic structures,*
- *Sorting algorithms.*

After the completion of each section, the students were again classified into one of three learning paths. The data sets in each test contain 197 cases for the first cycle of learning (the collection includes two student recruitments for 2015/2016 and 2016/2017) and 57 cases for the second cycle (academic year 2017/2018). The students in the second group constitute the so-called control set, i.e. they are completely new cases that were not taken into account when building the classification model.

Table 1 presents a summary of the classification results performed by individual networks for the first and second learning cycles. Their placement in the table corresponds to the order in which they are discussed, and to facilitate identification of the network, the column "*learning cycle number*", which was different for each of the networks in question, was placed in the table. The Section 1 test 1 column is empty because it was data used to teach the network.

The results obtained during the first and second learning cycles are similar. The results obtained during the tests confirm that the best of the generated networks is a network with 9 neurons in the hidden layer activated by the hyperbolic tangent function and with output neurons with a linear activation function. For students of the last year, the correctness of the classification based on the model of the first network is greater than 98.5%. Thus, the created model is predictive and can be successfully used to generate learning paths on the e-Student portal.

Table 1.

Summary of classification results for two consecutive learning cycles

1Net-work No.	Number of network neurons	Number of learning cycles	Section 1 test 1		Section 1 test 2	
			Cases classified correctly	Cases classified incorrectly	Cases classified correctly	Cases classified incorrectly
1	MLP 4-9-3	50	-	-	98.53%	1.47%
2	MLP 4-9-3	53	-	-	96.21%	3.79%
3	MLP 4-8-3	31	-	-	96.47%	3.53%
4	MLP 4-8-3	26	-	-	91.21%	8.79%
Net-work No.	Number of network neurons	Number of learning cycles	Section 2 test 1		Section 2 test 2	
			Cases classified correctly	Cases classified incorrectly	Cases classified correctly	Cases classified incorrectly

1	MLP 4-9-3	50	97.46%	2.54%	99.03%	0.97%
2	MLP 4-9-3	53	95.43%	4.57%	96.01%	3.99%
3	MLP 4-8-3	31	95.43%	4.57%	96.47%	3.53%
4	MLP 4-8-3	26	93.4%	6.6%	88.97%	11.03%
Network No.	Number of network neurons	Number of learning cycles	Section 3 test 1		Section 3 test 2	
			Cases classified correctly	Cases classified incorrectly	Cases classified correctly	Cases classified incorrectly
1	MLP 4-9-3	50	96.95%	3.05%	98.97%	1.03%
2	MLP 4-9-3	53	95.43%	4.57%	95.53%	4.47%
3	MLP 4-8-3	31	95.43%	4.57%	95.63%	4.37%
4	MLP 4-8-3	26	94.42%	5.58%	91.28%	8.72%
Network No.	Number of network neurons	Number of learning cycles	Section 4 test 1		Section 4 test 2	
			Cases classified correctly	Cases classified incorrectly	Cases classified correctly	Cases classified incorrectly
1	MLP 4-9-3	50	97.97%	2.03%	99.1%	0.9%
2	MLP 4-9-3	53	97.46%	2.54%	96.03%	3.97%
3	MLP 4-8-3	31	96.95%	3.05%	96.03%	3.97%
4	MLP 4-8-3	26	90.86%	9.14%	90.23%	9.77%

Source: Own work

As a result of classification using neural networks, each student was assigned to their individual learning path. Paths are selected individually for each person, after they pass the test for a given section. Table 2 summarizes the number of students using teaching materials available in modules 2-5, after completing modules 1-4 and qualifying for one of the three learning paths. In both learning cycles one can notice a decrease in the number of people on the first learning path. In each module, the smallest number of people were taught on the third path.

Table 2.

Numbers of students assigned to individual learning paths in the first and second learning cycles

Learning module passed	path 1		path 2		path 3	
	Cycle 1	Cycle 2	Cycle 1	Cycle 2	Cycle 1	Cycle 2
1	79	20	91	25	27	12
2	89	18	76	26	32	13
3	91	19	70	26	36	12
4	83	16	68	27	46	14

Source: Own work

CONCLUSION

The four neural networks described in this paper have been tested on new, previously unknown data. The high results obtained during the tests show that these networks have the ability to generalize.

Based on the analysis of the results of the conducted research, it can be concluded that neural networks are a very good tool for indicating the learning path on which students will find the teaching materials best suited to their particular abilities and potential limitations in acquiring knowledge but enabling them to acquire professional competences. The students' knowledge was tested in an examination in the subject *Algorithms and Data Structures*. The multi-layer network with one hidden layer proved to be the best. In the input layer there are 4 neurons, in the hidden layer 9, and in the output layer 3. The functions of neuron activation in individual layers were: hyperbolic tangent, for the hidden layer and the linear function for the output one. The number of neurons in the output layer was determined by the specificity of the tool used to create it (*Statistica Neural Networks package*) and is equal to the number of classes to which cases are assigned.

The assumption regarding the number of neurons in the hidden layer calculated on the basis of Kolmogorov's theorem has been confirmed. Based on this theorem, it was calculated that a sufficient number of neurons in the hidden layer is 9. The best parameters for the quality of learning, testing and validation were obtained for the network designed in just this way. Tests carried out for external data (new groups of students) confirmed that this network achieves the best results in the classification of new cases. Confirmation of the thesis of the correct assignment of the students to the learning paths was provided by the answers they gave during a discussion on the structure of the didactic materials made available to them.

As a result of the conducted research, it has been shown that a neural network with one hidden layer has a sufficient structure to solve the problem of assigning students to individual learning paths. This does not mean that a network with more layers would not cope better with the discussed problem, but the very good results of the testing phase achieved by the discovered network meant that further expansion of the network, with subsequent hidden layers, was discontinued. The test results confirmed a paradigm with which many researchers of neural networks agree, which states that networks with one hidden layer are sufficient to solve most classification problems (Markowska-Kaczmar, Kwaśnicka, 2005; Kurzyński 2008; Osowski, 2000).

Analysing the above facts, it can be concluded that it is reasonable to implement a neural network with the best parameters obtained during tests to generate individual, evaluating learning paths.

REFERENCES

Anand S.S., Mobasher B. (2005). *Intelligent techniques for Web personalization*, Intelligent techniques for Web personalization, Springer Verlag, Hamburg.

Bishop C. M., (1995). *Neural networks for pattern recognition*, Clarendon Press, Oxford

Baylari A., Montazer G.A. (2009). *Design a personalized e-learning system based on item response theory and artificial neural network approach*. Journal of Expert Systems with Applications, Volume 36 Issue 4: 8013-8021, 2009, http://dx.doi.org/10.1016/j.eswa.2008.10.080.

Dębska B., Kubacka A. (2014). *Generation of individual learning path in distant education*, Int. J. Cont. Engineering Education and Lifelong Learning, Vol. 24, No. 1: 53-61

Dębska B., Kubacka A. **(**2016). *Classification algorithms in the personalization of education portals,* Int. J. Cont. Engineering Education and Lifelong Learning, Vol. 26, No. 1: 105-116

Duch W., Korbicz J., Rutkowski L., In Tadeusiewicz R. (Eds.) (2000). *Neural Network*. Tom 6. Academic Publishing House Exit, Warszawa [In Polish].

Galushkin, A. I. (2007). *Neural network theory*, Berlin Heidelberg, Springer-Verlag.

Kurzyński, M. (2008). *Artificial intelligence methods for engineers*, State Higher Vocational School Publishing House in Legnica, Legnica [In Polish].

Markowska-Kaczmar U., Kwaśnicka H. (2005). *Neural networks in applications*, Wroclaw University of Technology Publishing House, Wrocław [In Polish].

Lula, P. (1999). *Selected applications of neural networks and the Statistica Neutral Networks program*, StatSoft, Kraków [In Polish].

Lula, P. (2006). *Neural networks - course materials*, StatSoft, Kraków [In Polish].

Osowski, S. (2000). *Neural networks for information processing*, Warsaw University of Technology Publishing House, Warszawa. [In Polish].

STATISTICA (*data analysis software system*) version 9.0. Electronic manual (2009). StatSoft, Inc.

Tadeusiewicz, R. (1993). *Neuron network,* Academic Publishing House Exit, Warszawa [In Polish].

Tadeusiewicz, R., Gąciarz T., Borowik B., Leper B., (2007). *Discovering properties of neural networks using C # programs*, Polish Academy of Learning, Kraków. [In Polish].

Welstead, S.T. (1994). *Neural network and fuzzy logic applications in C/C++*, New York: John Wiley & Sons.

Wilde P., (1997). *Neural Network Models. Theory and Projects*. Springer-Verlag London.