Comparison of the Application of FNN and LSTM Based on the Use of Modules of Artificial Neural Networks in Generating an Individual Knowledge Testing Trajectory

Ekaterina Vitalevna Chumakova[1], Dmitry Gennadievich Korneev[2], Tatiana Alexandrovna Chernova[1], Mikhail Samuilovich Gasparian[2*], Andrey Aleksandrovich Ponomarev[2]

[1] Applied Computer Mathematics Department, Moscow Aviation Institute, 4 Volokolamskoe Shosse, Moscow 125993, Russia
[2] Applied Informatics and Information Security Department, Plekhanov Russian University of Economics, 36 Stremyanny Lane, Moscow 117997, Russia

Corresponding Author Email: Gasparian.MS@rea.ru

ABSTRACT

The paper considers the issues of implementing an adaptive testing system using artificial neural network modules, which should resolve the problem of intellectual selection of the next questions, thereby generating an individual testing strategy. An attempt is made to increase the accuracy of the artificial neural network in determining the level of difficulty of the next test question for two types of architectures – Feedforward Neural Network (FNN) and Long-Short Term Memory (LSTM) network. Parameters affecting the quality of education are analyzed. A modification of the input layer architecture of the FNN that allows for a significant increase in the accuracy of the networks is reviewed. To solve the problem of selecting the thematic block of the question, a hybrid module structure comprising the artificial neural network together with the algorithmic processing of the results it delivers is proposed. The feasibility of using an FNN compared to the LSTM network architecture is substantiated. The network input parameters are identified, and different architectures and network training parameters (weight update algorithms, loss functions, number of training epochs, packet size) are compared. The use of the FNN direct propagation network as part of a hybrid algorithmic module makes it possible to construct a trajectory with an individual testing length, regardless of the number of thematic blocks.

1. INTRODUCTION

Computer-based knowledge control systems have long been an integral part of educational technology. They are actively employed not only as a tool for the final assessment of the level of knowledge in a particular area but also as an instrument of current control to adjust the educational program, i.e., developing individual training profiles [1, 2]. The use of testing systems is not limited to the field of education. The tests are of increasing interest, for example, to the HR services of large companies [3, 4], both for hiring new professionals and for testing employees as part of continuing education [5].

In this regard, in our opinion, the greatest interest for research is computerized adaptive testing (CAT), at the core of which lies the use of intellectual methods. Particularly, the method includes the intellectual selection of questions based on the level of knowledge demonstrated by the test taker to generate an individual testing trajectory to reliably determine the level of knowledge with an optimal number of questions. Notably, the largest share of current theoretical and applied research into CAT relates primarily to the use of artificial neural networks (ANNs) [6]. For this reason, the topic under consideration is of theoretical and practical interest.

Research on the application of ANNs in testing systems is currently not as active as, for instance, in the fields of pattern recognition [7, 8] or text analysis [9]. In testing, ANNs are most often proposed to be used as the final scoring module; the test questions themselves are selected according to a

predetermined clear algorithm [10-12]. In several papers, there were attempts to solve the problem of intelligent question selection in the form of determining the level of difficulty of the next question based on the accuracy of the previous answer. For this purpose, the researchers propose to apply Feedforward Neural Networks (FNN) and recurrent Long-Short Term Memory (LSTM) networks [13-15]. As interesting ideas published in several works, we can note the use of open systems methods in the creation of neural networks, in particular, the creation of ANNs according to the modular principle [16].

At the heart of our research is an attempt to apply an approach that allows one to create a universal system structure with an ANN to determine the topic and difficulty of the next question, minding all the previous answers and the difficulty of all preceding questions, Also, the system should consider the relatedness of topics and response time as a marker of guessing or search for the answer.

At the previous stage of the research reported [17], we proposed the difficulty of the next questions to be determined using one of the two types of ANN, FNN, or LSTM, which did not show an accuracy above 85%. This prompted discussion about the further improvement of networks and served as a basis for a detailed study of approaches and methods to improve the effectiveness of ANNs [18-20]. An integral and important part of the testing system is the intelligent selection of the topic of questions as a basis for creating individual testing trajectories to reliably determine the testee's level of

knowledge in a specific area when constructing an adaptive testing system. In this context, the use of ANNs in the module will allow considering several important parameters, such as the relatedness of topics and the difficulty of questions already asked, when selecting the next question.

Therefore, the purpose of the study is to increase the efficiency of ANN and determine the feasibility of using a specific architecture for solving problems of this class, as well as building a universal module structure, independent of the number of thematic blocks of a particular test, containing an artificial neural network, and comparing the use of two types of architectures (FNN and LSTM) based on the use of modules of ANN in the formation of an individual knowledge testing trajectory.

To achieve this goal, it is necessary to solve the following tasks:

1. To conduct a series of ANN training cycles to determine the level of complexity of the question being asked using generally accepted approaches for tuning learning parameters aimed at improving the accuracy of the network and identify what factors affect the accuracy concerning this area.

2. To conduct training considering the identified factors and training parameters that affect the accuracy of the network.

3. To analyze the data of the testing process that affect the choice of the topic of the question and determine the type of ANN and its place in the module for determining the topic of the next question, regardless of the number of thematic blocks of the test.

4. To form a training sample with the involvement of experts and conduct a series of training cycles for ANNs of various architectures to identify the influence of various parameters of the learning process on the accuracy of the neural network model, as well as to determine the most appropriate network configuration.

2. METHODS

2.1 Study design

The study analyzed approaches and ways to increase the accuracy of neural networks in the learning process in terms of the possibility of their application in the system of adaptive testing of knowledge in technical disciplines. Such disciplines as "Databases", "Informatics", and "Computer Graphics", read in different educational institutions for students of different courses and specialties, were chosen. Four teachers from the Moscow Aviation Institute (National Research University) and the Plekhanov Russian University of Economics acted as experts involved to form a training sample.

Earlier, [17] proposed to determine the level of complexity of the next test question using an ANN of one of two network types based on an analysis of the application areas of various ANN types and, in particular, in problems of this type. It was decided to dwell in more detail on the study of two network types, namely FNN and LSTM; however, they did not show an accuracy of more than 85%.

At the initial stage, in addition to the study [17], to identify the reasons for the low efficiency of networks that determine the level of complexity of the next question, we carried out an additional series of training experiments. Using the results obtained in the study [17], the Keras high-level library was used; we used Adam as an optimizer in conjunction with the

MSE loss function (mean square error) as showing the best efficiency results in training networks concerning this area. During different training cycles for FNN (with 6-m-5 architecture) and LSTM (4-m-5) networks, the network structure and training parameters varied. Among the parameters affecting the accuracy were changing the number of neurons in the hidden layer, increasing the number of hidden layers, increasing the number of training epochs, changing the package size, and controlling the learning rate of the optimizer. We also analyzed the training sample.

To increase the sample size and its completeness, for an abstract thematic block, we decided to form a base with an equal numerical ratio of objects of different classes consisting of 8,075 complete testing trajectories (sequences of the complexities of six questions) was formed using specially developed software and expert (teacher) evaluation, based on which new training samples were obtained. We also decided to change the architecture of the input layer for the FNN network.

To identify the impact on the accuracy of learning networks on a new sample, several training experiments were repeated for FNN networks (with 7-m-5 architecture) and LSTM (4-m-5).

Next, we researched the process of intellectual selection of the thematic block of the next test question given the responses obtained in the previous stages and the thematic relatedness of these stages. To obtain a universal instrument, various technical disciplines with a different number of thematic blocks were considered, specifically the disciplines "Databases" with 10 blocks, "Informatics" with eight blocks, and "Computer graphics" with 12 blocks of topics.

Proceeding from the previous stages of research, for the task of determining the thematic block of the next question, we decided to examine an FNN with a constant architecture, i.e., one that is independent of the number of thematic blocks in the discipline. For the implementation of this universal approach, we propose a hybrid module scheme that contains an ANN determining the sufficiency of assessment of the block together with algorithmic processing of the results delivered by the ANN for their further interpretation in the system.

It is assumed that the system has a database that stores answers to questions with a characteristic of belonging to a particular thematic block. At each stage of testing (each next question), for all available topics, one can get some indicator of its evaluation in the range from 0 to 1 in terms of the correctness of answers to questions of various levels of complexity and their connection with other topics. In addition to grading, it is possible to consider other statistical indicators, for example, the maximum or minimum level of complexity of the questions asked. The cyclic preparation of questions and statistical indicators of the subject is assigned to algorithmic submodules, including the module for preparing data (obtaining average estimates, calculating the matrix of related topics, etc.), the module for mixing topics (for obtaining an arbitrary sequence of questions), the module for sampling data on the next topic, and the module selection of the next question for the topic.

The ANN in the module structure should give an opinion on the continuation or completion of the survey on the subject, i.e., the task of the network is reduced to a binary classification, which is interpreted as a recommendation for continuing the survey on a specific test topic. Topics identified by the ANN as graded drop out of the topic selection list, and the testing cycle ends as soon as there are no ungraded topics left. The

structure of the module is such that it assumes the use of ANNs only for direct propagation, while the number of input neurons depends on the number of parameters that must be considered when deciding whether to continue the survey on topics.

In the structure of the hybrid module, we studied an ANN with a 5-m-1 architecture. The paper justifies the choice of the indicated ANN architecture and compares the results of the ANN operation for various values of the parameter m. In the designed ANN, input neurons are supplied with the number of questions on the topic, the maximum and minimum difficulty of the questions already asked, the normalized value of the degree of testedness of the thematic block, and the number of questions recommended to establish the true level of knowledge on the theme.

All experiments were carried out using the Keras library.

2.2 Sample description

For the convenience of analysis when training networks to determine the complexity of the question and to increase the volume of the general sample, a base of 8,075 testing trajectories for one abstract thematic block was created. Each trajectory presupposes that the testee is asked five questions from this thematic block, the first one being of the average difficulty group and other questions of lesser or greater complexity at the discretion of the expert teacher, depending on the answers received to all previous questions and the relative response time. Information on each question asked that is stored includes its number in the trajectory, difficulty level, score for the answer, deviation of response time from "normal" the percentage deviation value from -1 to 2), and the difficulty of the next question, which is indicated by an expert.

Uniform completeness of the sample was achieved with the help of specially created software, which added to the base of samples those questions from the trajectory that had uniform scores and percentage deviations of response time. After each new question was added, the expert specified the level of difficulty for the next question proceeding from the previous scores, response times, and question difficulty.

The database of test trajectories thus created for the training of an LSTM network provided a general sample of 24,825 sets with a training sample of 80%, i.e., 19,860 training sets. The architecture of the trained network for five question complexity levels had the 4-m-5 structure, where each turn was presented with a question number, an answer score, question difficulty, and a time deviation from normal. The training sample was 13,780 sets, which constituted 80% of the general sample of 17,230 sets.

In preparing the general training sample, we limited the number of questions recommended for a complete assessment to the range of three to five, meaning that two to six questions could be asked on each topic. This produced a relatively small training sample of 1,410 sets and 180 sets in both the validation and test samples.

2.3 Research stages

At the first stage, to identify the reasons for the relatively low accuracy of networks that solve the problem of determining the complexity of a question, experiments were carried out with the simplest direct propagation network with one hidden layer of neurons and architecture 6-m-5 and, following general heuristic recommendations, for m = 15, 18, 21. SGD, Adam, NAdam, and RMSprop were compared as optimizers. The loss function MSE (mean square error) was used together with the optimizer. The training was carried out on a training sample of 1,500 sets, which was 80% of the general sample. Traditionally, training has taken place over a large number of epochs (300, 500, 700, and 1,000). Similar training experiments on the same general sample were carried out when switching to network architectures that included 2 and 3 hidden layers within 15-21 neurons, as well as for an LSTM network with a 4-m-5 architecture.

Based on the experiments, at the next stage, a study of emerging anomalies and possible methods for their elimination was carried out, and an analysis of data samples was carried out together with the architecture of the input layers of networks. Representative training samples were prepared.

At the next stage, we carried out repeated training cycles on new training sets and, in particular for the LSTM network; training was carried out for m=10, 15, and 21 hidden layer neurons for 50, 100, and 200 epochs. The experiments were also conducted via the Keras library with Adam used as the optimizer together with the MSE loss function as showing the best results for the tasks being solved.

For the modified FFN, training was carried out for the number of neurons in the hidden layer m = 15, 21. In addition, the size of the training sample allowed us to conduct experiments for two- and three-layer models, specifically 7-15-15-5, 7-21-21-21-5, and 7-21-21-21-5. The training was performed using the same instruments and approaches as for the LSTM network but over a greater number of epochs (200, 400, 600).

Further, the possibilities of using both FNN and LSTM networks for solving the problem of choosing a thematic block were analyzed. However, due to the features of the problem, it was proposed to use the FNN network as part of a universal hybrid module. For the general architecture of the ANN type 5-m-1, experiments were carried out for m = 10, 15, 20 neurons in the hidden layer. The duration of the experiments was 300, 400, and 500 epochs. An experimental attempt was made to optimize package size, for which purpose several experiments with various package sizes (8, 20, 32, 40) were conducted. Similar training experiments on the same general sample were conducted with the transition to network architectures with two and three hidden layers ranging between 10/20 neurons.

3. RESULTS

3.1 Increasing the effectiveness of an ANN determining question difficulty

Examination of the ANN determining the difficulty of the next question demonstrates that networks with varying architectures remained at an average accuracy of 83-85%. In this part of the study, we applied standard approaches to improving the accuracy of training, such as increasing the number of hidden layer neurons, the number of hidden layers themselves, and eras of training, as well as mixing data and changing the speed of training employing Keras. The results of the experiments suggest the need to analyze the general sample of examples for the completeness and difficulty of the model, as well as its expansion.

Given all previous experiments using the Keras library, including in work [17], the Adam optimization function and

the MSE loss function were chosen for training as showing the best results for the specifics of the tasks being solved.

As a result, several network architectures were trained for 50, 100, and 200 epochs on an increased training sample of 19,860 sets (Table 1).

Already after 50 epochs, the network demonstrated the accuracy of ≅98% (with m=10, 15) and 99% (m=21). Further increase in the number of training epochs yielded no overtraining effect. Figure 1 shows the learning curves of the 4-21-5 LSTM network, which demonstrated the best accuracy among the trained architectures.

Table 1. Results of training LSTM networks of different architectures

Accuracy model	Accuracy of training	Accuracy on the control set	Accuracy on the test set
4-10-5	97.75	97.9	97.81
4-15-5	98.86	98.75	98.82
4-10-10-5	98.91	98.23	98.37
4-21-5	99.34	99.4	99.36
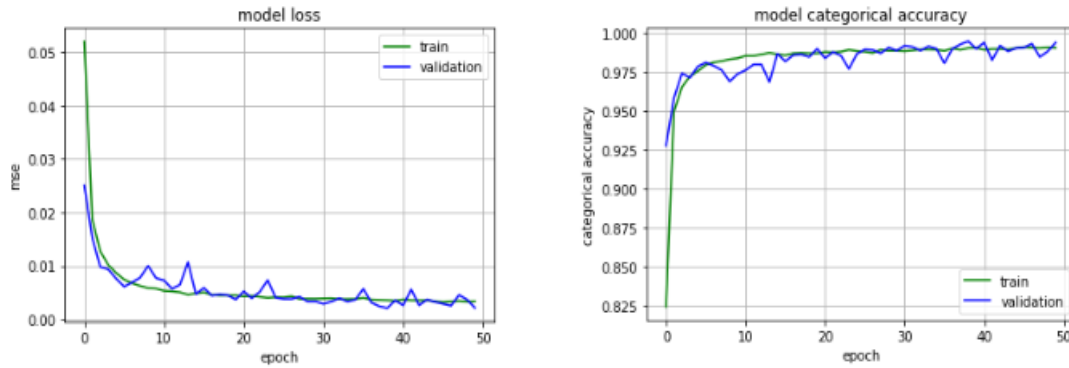
Figure 1. Learning curve of the 4-21-5 LSTM network

Table 2. Results of training FNNs of different architectures over 600 epochs

Accuracy model	Accuracy of training	Accuracy on the control set	Accuracy on the test set
7-15-5	94.19	94.44	94.51
7-21-5	96.54	96.78	96.36
7-15-15-5	96.34	95.99	96.34
7-21-21-5	97.42	97.23	97.36
7-15-15-15-5	98.85	98.79	98.75
7-21-21-21-5	99.12	99.19	99.07
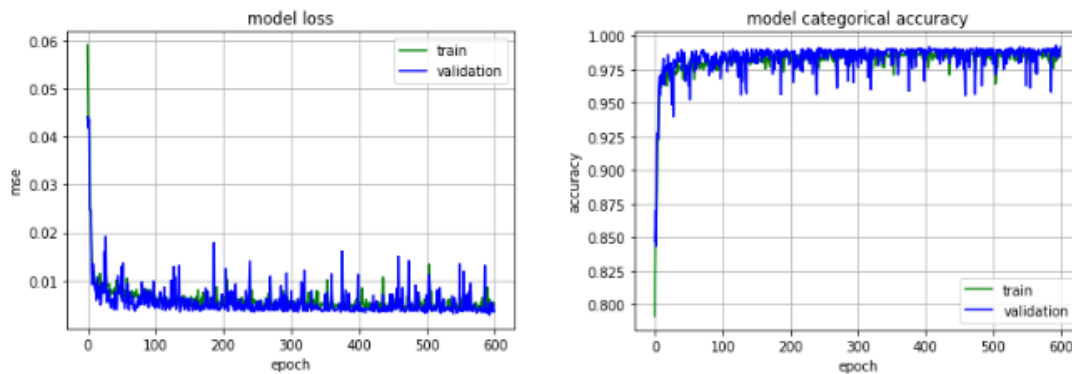
Figure 2. Learning curve of the 7-21-21-21-5 FNN

The LSTM network with two hidden layers of 10 neurons showed results comparable to those of the network with one hidden layer of 15 neurons, although its training took twice the time.

Preparation and analysis of the training sample for a 6-m-5 FNN proposed in the study [17] reveals the presence of repeating and inconsistent data due to the use of the averaged response time deviation as an input parameter. For example, the combination "guessing"/"finding an answer" (negative and positive deviations) is numerically similar to the normal rhythm of receiving answers. To resolve the identified contradiction, it was decided to conduct preliminary processing of data to bring it to an acceptable range, i.e., pre-

normalize the response time deviations to the interval [0; 1,6]. However, this solution was only partially effective, as it did not account for the sequence of deviations. As a result, the architecture of the input layer of neurons had to be slightly adjusted by adding to it one more neuron. Thus, in addition to the averaged values of the correctness of answers to questions and their difficulty, the number of questions already asked, the average deviation of response time, the score for the answer to the last question, and its difficulty, we also feed the normalized deviation of response time to the input of the ANN.

In this, the accuracy of the 7-m-5 architecture network with one hidden layer was 94 and 96% for the number of hidden layer neurons m=15 and 21, respectively. Increasing the

number of training epochs from 200 to 600 only slightly raised the network accuracy by 0.5-1%. Further improvement of the accuracy of the ANN was accomplished by adding additional hidden layers. Specifically, the volume of the training sample allowed us to conduct experiments for two- and three-layer models with the same number of neurons in the layer. The results of the experiments are shown in Table 2.

The last 7-21-21-21-21-5 FNN architecture showed the highest accuracy of 99% among all considered, and a learning time comparable to that of an LSTM architecture. The learning curves are depicted in Figure 2.

Both types of networks delivered similar results in determining the difficulty of the next question and also required practically the same time to train. At this stage, the LSTM architecture has an underlying advantage, as it does not require significant preliminary algorithmic preparation of data.

3.2 The structure of the hybrid ANN module for selecting the thematic block

When implementing the intelligent selection of the thematic block of the next question, first, we analyzed the opportunities for the application of two different ANN architectures, specifically FNN and recurrent LSTM network, to decide the difficulty of the next question.

Firstly, we attempted to obtain a universal network architecture that would not change with the transition from one subject (area of knowledge) to another. As a solution, an average number of difficulty levels and, accordingly, ANN outputs (in our case, ашму) was proposed. In this case, the authors of tests (test questions) will have to adjust to this gradation of difficulty levels by splitting or combining their groups of questions to meet the requirements of the system. This approach, however, is not feasible for topic selection, as each area of knowledge has its unique composition and number of topics, which is not always amenable to regrouping. In addition, a fixed number of topics in tests somewhat reduces the overall universality of the system and sets high requirements and preparation workload for test writers.

Secondly, to obtain the result with a constant and minimal number of input neurons, we considered the opportunity of using a recurrent LSTM network, which essentially preserves all of its previous states in time. For this reason, the application of this type of network to determine the topic of the next questions imposes high requirements on the training sample. Furthermore, the network essentially has to be able to save the number of previous states equal to the maximum number of test questions (which is around 50-100 states for each testing trajectory). This factor, among other things, greatly increases computation time.

To resolve the outlined issues, it was decided to employ a hybrid module combining a neural network that determines the degree of testedness of the given topic, and an algorithmic superstructure, which receives a recommendation from the ANN for each of the test topics. The functioning of the module is schematically depicted in Figure 3.

The idea behind the operation of the proposed module is that the ANN recommends whether or not to continue testing on the given topic. This recommendation is given for any topic based on a constant number of criteria regardless of the number of topics. The order in which the topics come to the input of the ANN is random, and their progression through the network continues until the first correct answer (the topic is

considered selected).

Viable criteria, i.e., the input network parameters utilized to decide whether to continue or to stop testing on the specific topic, include the total "degree of testedness" of the topic (taking into account its relation to other topics and answers to them), the number of questions already asked, the recommended number of questions that should be asked for completeness, and the maximum and minimum difficulty of the questions asked.
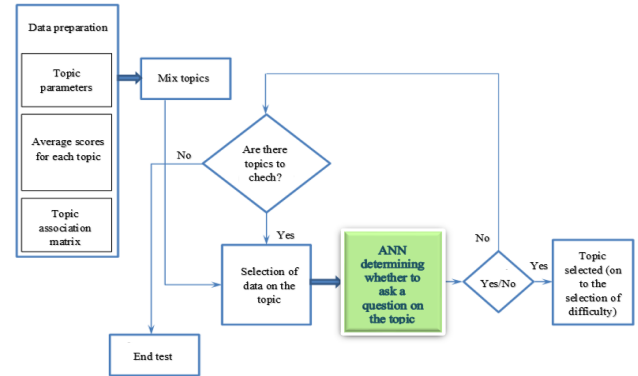

Figure 3. Functioning diagram of the question topic selection module

The proposed approach will enable the construction of individual testing trajectories based on not only specific topics and question difficulty, but also the number of questions both on a particular topic and in the entire test. This solution will thus provide an adequate assessment of the level of knowledge with a minimal number of questions, which will also be individual for each test taker.

The main input parameter is the "degree of testedness" of the topic, which should be some numerical value (preferably in the range from 0 to 1), considering the test taker's answers on related topics. To obtain a numerical representation of this indicator, it was proposed to use a matrix of topic relatedness, the elements of which are coefficients of relatedness in the range from 0 to 1 (0 not at all related, 1 completely related), and a vector of average scores on the topics. The relatedness matrix is symmetric, with ones on the main diagonal that set a 100% consideration of scores on questions in the current topic.

Multiplying the connectedness matrix (1) by the vector of mean scores gives a vector of values (for each of the topics), which is the sum of responses on the connected topics contributing to the total score according to the connectedness coefficients.

The values obtained in this way should be normalized to the range 0|1 (where 1 is the highest mark for the topic) by dividing them by the corresponding sums of all coefficients of relatedness for the particular topic. The normalized value is considered the value of the degree of testedness of the i-th topic Q_i (2).

$$Q_i = \sum_{j=1}^{N} m_{ij} o_i / \sum_{j=1}^{N} m_{ij} = (m_{i1} o_1 + \cdots + m_{ij} o_i + \cdots + m_{iN} o_N)/(m_{i1} + \cdots + m_{ij} + \cdots + m_{iN})$$

Thus, based on the described input criteria, the overall architecture of the ANN is 5-m-1.

$$\begin{bmatrix} m_{11} & m_{12} & ... & m_{1N} \\ m_{21} & m_{22} & & m_{2N} \\ ... & & & ... \\ m_{N1} & m_{N2} & ... & m_{NN} \end{bmatrix} * \begin{bmatrix} o_1 \\ o_2 \\ ... \\ o_N \end{bmatrix} = \begin{bmatrix} m_{11}o_1 + m_{12}o_2 + \cdots + m_{1N}o_N \\ m_{21}o_1 + m_{22}o_2 + \cdots + m_{2N}o_N \\ ... \\ m_{N1}o_1 + m_{N2}o_2 + \cdots + m_{NN}o_N \end{bmatrix} \qquad (1)$$
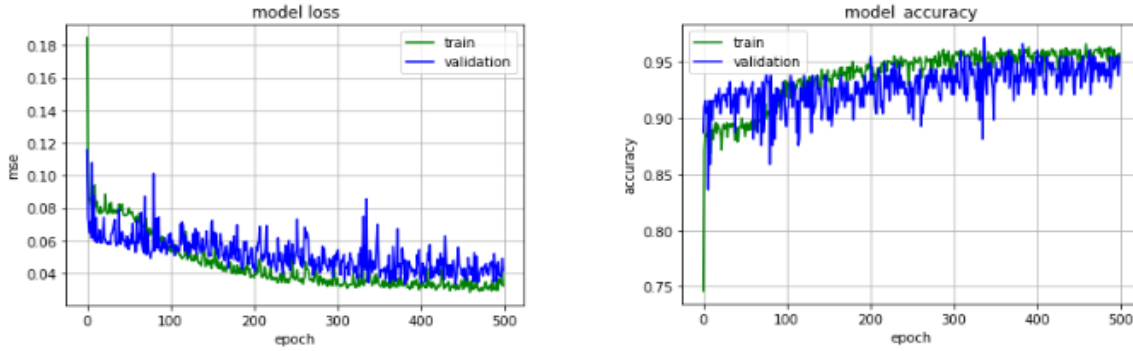

Figure 4. Learning curve of the 5-20-1 FNN

3.3 Training the ANN from the thematic block selection module

According to the known heuristic rules [21-23], with the available volume of the training sample, it is reasonable to train m = 10, 15, 20 neurons in the hidden layer. The training was performed for 300, 400, and 500 epochs. Raising the number of training epochs from 300 to 500 increased the result of the network by an average of 2%. The results of the training experiments are summarized in Table 3.

Table 3. Results of training FNNs of different architectures

Accuracy model	Accuracy of training	Accuracy on the control set	Accuracy on the test set
5-10-1	90.57	90.40	90.56
5-15-1	93.88	94.00	93.75
5-20-1	95.67	95.48	95.00

The learning curves of the ANN with 20 neurons in the hidden layer tasked with the necessity of further examination on the topic are given in Figure 4.

An experimental attempt was made to optimize package size. For all the considered network architectures, several experiments with different package sizes (8, 20, 32, 40) were conducted. However, none of these options demonstrated a higher generalization ability.

Training experiments performed on the same general sample for networks with two and three hidden layers within 10|20 neurons per layer did not yield a network accuracy greater than 93%.

4. DISCUSSION

In the course of the research, we have tested the methods of improving ANNs as applied to specific tasks associated with creating adaptive testing systems. The identified shortcoming of the set of input parameters of the FNN network required not only modification of its input layer but also changes to the overall structure of the entire system proposed in the study [17] when using a network of this type. The conducted studies suggest that there are two types of ANN well-suited for determining question difficulty – FNN and LSTM networks.

The accuracy of these networks was successfully raised from 85 to 99%. The results obtained do not allow us to make an unambiguous conclusion about which type of ANN is preferable to use in problems of this class. Meanwhile, the LSTM network has a simpler external architecture and does not require additional preliminary data preparation.

At the same time, the use of LSTM networks to determine the thematic block of questions is inadvisable due to their high computational resource requirements both for training (the training process takes a significant amount of time) and startup. In our case, these are supplemented by higher requirements for the training sample. This is due to the fact that the depth of long-term dependencies is variable for each test. Therefore, to obtain universality, it is necessary to train considering the longest test. In addition, it remains unclear how, by implementing the module only in the form of an ANN, one can consider the connectivity of thematic blocks and overcome their changing number when moving from test to test. As a result, it was decided that the ANN should decide on a simpler and less intelligent task, i.e., not determine the thematic block itself but only whether or not to continue the survey on a specific block. Following the topics is not entirely arbitrary with this approach, and the regulation of the testing process and the integration of the response received from the ANN are assigned to the algorithmic structure. The proposed hybrid module is universal for the number of thematic blocks in the test. The FNN included in the model, the accuracy of which reached 95%, allows one to not only account for the thematic relatedness of the topics but also regulate the optimal number of questions altogether.

As a result, we have identified the complete structure of an adaptive testing system that utilizes the proposed ANN models and thereby enables the generation of individual testing trajectories, minding all the answers received from testees and the difficulty of questions considering their thematic relatedness. In contrast, in previous solutions, only one previous step of the test is accounted for, and the accuracy of the recurrent network in testing does not surpass 75% [14]. A similar approach to the choice of the next question is in the study [24], where the ANN network was considered. The parameters of the five previous stages were statically considered, and the value of RMSE = 1.0634 was given as an estimate of the network training efficiency, which exceeds the error values obtained in this study.

5. CONCLUSION

As a result of the study, we conclude that the task of selecting the thematic block is best solved by an FNN. Furthermore, its use as part of a hybrid algorithmic module allows creating trajectories with individual lengths of testing regardless of the number of thematic blocks.

Overall, implementation of the proposed instruments will enable the organization of adaptive testing with an intellectual selection of questions depending on the level of the testee's knowledge to generate an individual testing trajectory to reliably determine the student's level of knowledge with the optimal number of questions.

The study presents a successful attempt to improve the efficiency of two types of ANNs, FNN, and LSTM, determining the difficulty of test questions asked. The obtained results do not prove any type of network to be more effective in solving this class of tasks. Both types of networks show high accuracy and comparable time of training. The choice of one specific type of network can be made based on a practical experiment on a real system.

Furthermore, the practical realization of the proposed adaptive testing system based on ANN and its approbation compared to an oral exam with the teacher may give insight into the effectiveness of using ANNs as part of testing systems in principle.

REFERENCES

[1] Liao, C.H., Wu, J.Y. (2022). Deploying multimodal learning analytics models to explore the impact of digital distraction and peer learning on student performance. Computers & Education, 190: 104599. https://doi.org/10.1016/j.compedu.2022.104599

[2] Tomasevic, N., Gvozdenovic, N., Vranes, S. (2020). An overview and comparison of supervised data mining techniques for student exam performance prediction. Computers & Education, 143: 103676. https://doi.org/10.1016/j.compedu.2019.103676

[3] Votto, A.M., Valecha, R., Najafirad, P., Rao, H.R. (2021). Artificial intelligence in tactical human resource management: A systematic literature review. International Journal of Information Management Data Insights, 1(2): 100047. https://doi.org/10.1016/j.jjimei.2021.100047

[4] Jøranli, I. (2018). Managing organisational knowledge through recruitment: Searching and selecting embodied competencies. Journal of Knowledge Management, 22(1): 183-200. https://doi.org/10.1108/JKM-12-2016-0541

[5] Say, R., Visentin, D., Cummings, E., Carr, A., King, C. (2022). Formative online multiple-choice tests in nurse education: An integrative review. Nurse Education in Practice, 58: 103262. https://doi.org/10.1016/j.nepr.2021.103262

[6] Wang, N., Wang, D., Zhang, Y. (2020). Design of an adaptive examination system based on artificial intelligence recognition model. Mechanical Systems and Signal Processing, 142: 106656. https://doi.org/10.1016/j.ymssp.2020.106656

[7] Coulibaly, S., Kamsu-Foguem, B., Kamissoko, D., Traore, D. (2022). Deep convolution neural network sharing for the multi-label images classification. Machine Learning with Applications, 10: 100422. https://doi.org/10.1016/j.mlwa.2022.100422

[8] Azgomi, H., Haredasht, F.R., Motlagh, M.R.S. (2022). Diagnosis of some apple fruit diseases by using image processing and artificial neural network. Food Control, 145: 109484. https://doi.org/10.1016/j.foodcont.2022.109484

[9] Toledano-López, O.G., Madera, J., González, H., Simón-Cuevas, A. (2022). A hybrid method based on estimation of distribution algorithms to train convolutional neural networks for text categorization. Pattern Recognition Letters, 160: 105-111. https://doi.org/10.1016/j.patrec.2022.06.008

[10] Rodríguez-Cuadrado, J., Delgado-Gómez, D., Laria, J.C., Rodríguez-Cuadrado, S. (2020). Merged Tree-CAT: A fast method for building precise computerized adaptive tests based on decision trees. Expert Systems with Applications, 143: 113066. https://doi.org/10.1016/j.eswa.2019.113066

[11] Pavlenko, D., Barykin, L., Nemeshaev, S., Bezverhny, E. (2020). Individual approach to knowledge control in learning management system. Procedia Computer Science, 169: 259-263. https://doi.org/10.1016/j.procs.2020.02.162

[12] Pominov, D.A., Kuravsky, L.S., Dumin, P.N., Yuriev, G.A. (2020). Adaptive trainer for preparing students for mathematical exams. International Journal of Advanced Research in Engineering and Technology, 11(11): 260-268.

[13] Petrovskaya, A., Pavlenko, D., Feofanov, K., Klimov, V. (2020). Computerization of learning management process as a means of improving the quality of the educational process and student motivation. Procedia Computer Science, 169: 656-661. https://doi.org/10.1016/j.procs.2020.02.194

[14] Jafri, S.S.M. (2007). Computerized adaptive testing using neural networks. A thesis presented to the deanship of graduate studies in partial fulfillment of the requirements for the degree Master of Science. King Fahd University of Petroleum & Minerals Dhahran, Eastern Province, Saudi Arabia.

[15] Zhadaev, D.S., Kuzmenko, A.A., Spasennikov, V.V. (2019). Peculiarities of neural network analysis of students' level of training in the process of adaptive testing of their professional competencies [Osobennosti neirosetevogo analiza urovnia podgotovki studentov v protsesse adaptivnogo testirovaniia ikh professionalnykh kompetentsii]. Vestnik Brianskogo gosudarstvennogo tekhnicheskogo universiteta, 75(2): 90-98.

[16] Grigorev, A., Mamaev, V. (2016). On the application of neural networks in knowledge testing. [O primenenii neironnykh setei v testirovanii znanii]. Nauchnoe priborostroenie, 26(4): 77-84.

[17] Chumakova, E.V., Chernova, T.A., Belyaeva, Yu.A., Korneev, D.G., Gasparian, M.S. (2022). Use of neural networks in the adaptive testing system. International Journal of Advanced Computer Science and Applications, 13(5): 20-27. http://dx.doi.org/10.14569/IJACSA.2022.0130504

[18] Bhandari, H.N., Rimal, B., Pokhrel, N.R., Rimal, R., Dahal, K.R., Khatri, R.K.C. (2022). Predicting stock market index using LSTM. Machine Learning with Applications, 9: 100320. https://doi.org/10.1016/j.mlwa.2022.100320

[19] Ang, K.M., Lim, W.H., Tiang, S.S., Ang, C.K., Natarajan, E., Ahamed Khan, M.K.A. (2022). Optimal training of feedforward neural networks using teaching-learning-based optimization with modified learning phases. In: Proceedings of the 12th National Technical Seminar on Unmanned System Technology 2020. Lecture Notes in Electrical Engineering, 770: 867-887. https://doi.org/10.1007/978-981-16-2406-3_65

[20] Lin, Z., Shi, Y., Chen, B., Liu, S., Ge, Y., Ma, J., Yang, L., Lin, Z. (2022). Early warning method for power supply service quality based on three-way decision theory and LSTM neural network. Energy Reports, 8(Suppl5): 537-543. https://doi.org/10.1016/j.egyr.2022.02.243

[21] Haykin, S. (2019). Neural Networks. A Comprehensive Foundation [Neironnye Seti: Polnyi Kurs]. Williams, Moscow.

[22] Hochreiter, S., Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8): 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

[23] Brauns, K., Scholz, C., Schultz, A., Baier, A., Jost, D. (2022). Vertical power flow forecast with LSTMs using regular training update strategies. Energy and AI, 8: 100143. https://doi.org/10.1016/j.egyai.2022.100143

[24] Viteepanya, B., Baikularb, P., Na-udom, A. (2021). Development of the next item selection procedure using artificial neural network with the item exposure control for computerized adaptive testing using simulated data. Journal of Faculty of Education Pibulsongkram Rajabhat University, 8(1): 37-48.