

# Development of an integrated complex of knowledge base and tools of expert systems for assessing knowledge of students in mathematics within the framework of a hybrid intelligent learning environment

O.V. Druzhinina<sup>1</sup>, I.A. Karpacheva<sup>2</sup>, O.N. Masina<sup>2</sup>, A.A. Petrov<sup>2</sup>,  
<sup>1</sup>Federal Research Center «Computer Science and Control» of RAS,  
44, building 2, Vavilov str., Moscow, 119333, Russia,  
<sup>2</sup>Bunin Yelets State University,  
28, Kommunarov str., Yelets, 399770, Russia

Received: May 13, 2021. Revised: June 18, 2021. Accepted: June 21, 2021. Published: June 23, 2021.

**Abstract**—The paper is devoted to the development of instrumental and methodological support for the assessment of students' knowledge in mathematics in the framework of a hybrid intellectual learning environment. According to the structural scheme of the model of intellectual assessment of knowledge and creative potential of students in mathematics, the structure of the corresponding knowledge base is formalized. Knowledge base within the framework of data mining methods to create a module for monitoring and evaluating knowledge in mathematics has been developed. To model the educational process the authors propose methods for solving clustering, dimensionality reduction, and classification problems based on machine learning and artificial neural networks. The basics for the development of integrated knowledge base complexes and artificial intelligence tools for the educational process are presented. The obtained results are aimed at creating methods which ensure the processes of operational training, control and evaluation of knowledge.

**Keywords**—hybrid learning environment, digital transformation of mathematics education, knowledge base, intelligent technologies in education, control and assessment of knowledge, machine learning, data mining, high-level languages.

## I. INTRODUCTION

THE active introduction of information and communication technologies and electronic educational resources into the educational practice of Russia has actualized the search for effective high-tech solutions to the problem of high-quality management of the educational process. In addition the search for new, practically significant didactic mechanisms which increase the efficiency, comfort and objectivity of evaluation procedures in the context of digitalization of education is relevant.

In the context of digitalization, the Learning management system (LMS) is becoming widespread. Any LMS-system provides a set of learning management functions, including assessment procedures. These procedures include testing at all stages of training; providing conditions for making individual training routes; systematic monitoring of the results and the process of using training materials, etc. Though there is a problem related to the fact that LMS systems do not provide personification of the learning process according to the account of the students' individual typological characteristics.

The learning systems developed in recent years contain web technologies [1]. These web technologies contain learning materials which support the problem solving meta-theory in a way that initiates dialogue and research rather than offering a definitive answer [2]. Besides it learning systems implement the ideas of digital game-based learning (DGBL) [3] and also take into account learning styles [4].

The digital transformation of education creates additional

opportunities for each student to achieve the necessary educational results. Personalization of the educational route is carried out at the expense of: 1) using of artificial intelligence methods; 2) augmented and virtual reality tools; 3) the development of a hybrid intellectual learning environment in educational institutions.

A hybrid intelligent learning environment is a rich information and educational environment which includes expert systems, intelligent systems, neural networks, etc. Inside of a hybrid intellectual learning environment conditions are formed for the effective design of individual educational trajectories of students depending on the level of education and individual characteristics [5], [6], [7].

Nowadays the issues of creating a hybrid intellectual learning environment for specific subject areas, especially mathematics, are being updated. Features of using of intelligent technologies and methods of neural network modeling in the development of a hybrid intelligent learning environment are presented in a number of publications of the authors [8], [9], [10]. This paper proposes an approach to the design of an expert system in developing of an integrated complex of knowledge base and tools for assessing knowledge on the subject.

It is important to consider two circumstances during developing a knowledge base for evaluating learning outcomes in a subject (mathematics).

First, in relation to the field of artificial intelligence the term "knowledge" is interpreted as a type of information "reflecting the specialist's (expert's) experience of in a certain subject field, his/her understanding of numerous current situations and ways to move from one description of an object to another" [11]. According to A. A. Mitskevich, G. S. Pospelov knowledge should have the following characteristics: explainability of model elements with the help of the model itself (internal interpretability); setting the structure of their semantics (structuredness), expression of regularities, causal relationships and other quantitative relations of knowledge (connectedness); having the property to generate new knowledge (mutual activity) [12], [13]. T. A. Gavrilova considers that it is important the structure of knowledge contains not only data, but also metadata [14], [15].

Secondly, the term "knowledge base" generalizes the one "database" but it is not reduced to it. A knowledge base is a special technically structured set of knowledge of a specific subject field. This set of knowledge presupposes equipping with an intelligent interface [16].

Various approaches are used in designing expert systems and different solutions to the problem of forming knowledge bases are proposed.

The development of the knowledge base of the expert system based on conceptual modeling of a dynamically complex environment based on the synthesis of the activity approach and situational analysis is carried out in the study by A. B. Sorokin, E. V. Brazhnikova and L. M. Zheleznyak [17]. The specified knowledge base is a decision matrix that allows

checking decisions both on the acts of activity and on situations. Verification by acts of activity forms the functional properties of the model. The situation check generates the structural properties of the ones. The decision matrix is a decision-making model too. This model is based on the generalization of the results "on the object of the process, on the subject of the process, on the subject and object of the process, on the object and components of the process".

The methodology for designing an active expert system based on a continuously developing knowledge base was suggested in research by D. Yu. Muromtsev, V. V. Ermolaev and A. Yu. Kotok [18].

Other authors consider the problem of forming the knowledge base of a learning system based on a semantic network [19]. The semantic network contains the concepts of the subject field of the academic discipline that is being studied and the relations between the concepts. The proposed model allows formalizing the knowledge base of the learning system.

The knowledge base of the expert system is designed based on the use of a pedagogical model for assessing the quality of students' knowledge in another research [20]. This knowledge base includes not only basic questions on the subject but also includes auxiliary and additional questions. The algorithms for presenting additional and auxiliary questions are generalized. Ones can be used in developing of intelligent systems in various subject fields.

We consider the knowledge base of the expert system in the context of solving the problem of applying artificial intelligence algorithms to support the process of teaching mathematics. This paper is a continuation of the following researches [9], [21].

The novelty of this paper lies in the application of intelligent technologies and computer learning to the implementation of a didactic model of forming a bank of tasks. We offer a method for analyzing test results based on machine learning without a teacher.

## II. METHODS

A team of authors is developing specialized software for design a hybrid learning environment. The control and evaluation of knowledge in a hybrid learning environment is carried out through the use of the neural network knowledge base. The knowledge base is formed on the basis of the theory of representation of the student's activity in a particular academic subject in the form of successive levels of assimilation [7]. Each of the levels (student, typical, heuristic, research) reflects the development of the students' experience in this subject of study. The level assessment tools are tasks that require certain students' knowledge, skills and experience.

A specific type of tests is used to assess the formation of each level. The tests that involve algorithmic activity with a hint are used for the first level of assimilation. The tests for the ability to apply existing knowledge, skills and experience as well as to demonstrate readiness to perform tests according to

an algorithm, sample or rule are used for the second level of assimilation. The third level of assimilation is characterized by the tests that involve both understanding and clarifying the situation and the transformation of known methods of solutions and actions into new methods. The fourth level of assimilation is characterized by the tests that do not have an unambiguous solution and involve the ability to act in a certain subject field “without any rules” [22].

The developed knowledge base includes the tests aimed at diagnosing the assimilation of educational material in mathematics at all the levels presented above. Tests of the first and third levels will make up the database of the expert system knowledge base and are differentiated according to three blocks (B1, B2, B3). Diagnostics of the research level of training with the use of an intelligent automated learning system is solved by the research team as an independent test [23].

The quantitative indicator of achieving the level of learning (the level of assimilation) is the coefficient of assimilation  $K$ . The coefficient is calculated as the ratio of the number of correct answers to the standard. V. P. Bespalko justifies the following thesis in relation to the educational field of activity: the learning process can be considered completed at this level of assimilation, if the coefficient  $K$  is greater than or equal to 0.7[22], [24]. We use this condition in developing an approach to form a knowledge base.

Verification procedures involve evaluating the results of learning on a topic, section, or module. The results are evaluated both for each level (local assessment) and for the whole levels (generalized assessment). Thus the generalized estimate is the result of integrating the solution of problems of three levels of complexity. Besides we have distinguished four levels of assimilation of the study material: unacceptable, low, medium and high.

Students have been assigned to one of the groups according to the following criteria. The group with an unacceptable level of assimilation includes students who have completed less than 70% of the tests of each of the blocks B1, B2, B3. The low level of assimilation includes students who have completed more than 70% of the tests of one of the three blocks (either B1 or B2 or B3). The medium level of assimilation includes students who have completed more than 70% of the tests of any two blocks of three (either B1 B2 or B2 B3 or B1 B3). A high level of assimilation includes students who have completed the tests of all three blocks B1, B2, B3 by more than 70%.

Further we will consider an approach to the intellectual assessment of students' knowledge and creative potential in the framework of a hybrid learning environment based on constructing and training of neural networks. One of the important stages is the development of a model structural diagram for intellectual assessment of knowledge and creative potential in mathematics in the general secondary education system.

An important issue in the development of a hybrid learning

environment is the domain analysis. If we consider the educational process as a single system the complexity of the process lies in the presence of objective and subjective factors. To form a knowledge base of a hybrid intelligent learning environment based on artificial intelligence technologies it is necessary to have source data that representatively reflects the main laws of the functioning of the educational system besides didactic materials. The general structure of the learning process model using artificial intelligence and the knowledge base is shown in Fig. 1.

The teacher and the students are considered as subjects in this scheme. The knowledge module contains the knowledge of an expert who acts as a teacher or methodologist. The artificial intelligence subsystem includes a number of algorithms for data analysis and evaluation of results which are interpreted further as knowledge. The task of the teacher during working with a hybrid intelligent learning environment is to verify the data obtained on the basis of pedagogical experience.

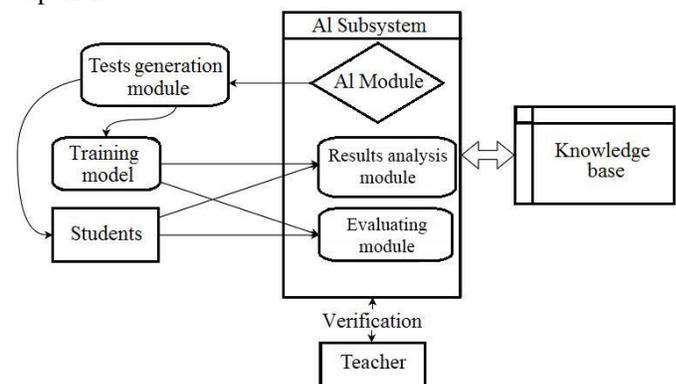


Fig. 1. The structure of a generalized model of the educational process using the artificial intelligence module and the knowledge base module

The teacher's influence on the system is excluded in using simulation. In this case the learning model is a model that can be represented in the form of a dynamic system, an artificial neural network, an automaton, etc. This model allows accounting the qualitative effects of the pedagogical process.

Using of artificial intelligence methods makes it possible to solve a number of problems of managing the pedagogical process. In particular the following tasks should be highlighted: clustering; dimensionality reduction; classification based on computing learning and artificial neural networks.

It is well-known that artificial neural networks are a universal tool for solving a wide range of problems. A characteristic feature of the neural network as a universal tool is the fact that the neural network is a flexible model for the nonlinear approximation of multidimensional functions. The advantage of the neural network approach to approximation is the ability to automate the process of model identification, selective sensitivity in the areas of data concentration, as well as smooth interpolation of characteristics in other areas [25].

Clustering is used in cases where managing of the pedagogical process requires the analysis of a large amount of data with a high dimension which can include the results of the

tests. The dimensionality reduction method can be used as an auxiliary method for clustering as well as a method for visualizing data and highlighting the main metrics. The allocation of metrics opens up a number of opportunities for implementing a more flexible, individual assessment of students. We consider such an approach to data mining which is based on classification using machine learning. This approach can be used in evaluating test results using the knowledge provided by experts. This knowledge is structured in the form of “test results” – “assessment”.

### III. RESULTS

We propose to apply the considered methods for developing of such modules in a hybrid intelligent learning environment which are associated with a special subsystem of artificial intelligence. These methods are aimed at solving such problems as developing of knowledge bases and tools for expert systems; developing of computer simulation tools and interactive control of subject knowledge, competencies and procedures; creating methods for engineering and presenting subject knowledge, competencies and procedures as expert systems; creating output mechanisms and users’ interface systems; providing of operational training due to a set of software and tools, methods, algorithms, techniques; implementing of algorithms and programs of mathematical support for intellectual support of learning processes, control and evaluation of knowledge.

An overview of the current state of data science theory and methods is contained in such source [26]. The essence and application of these methods to various problems have been particularly considered in the following studies [27], [28].

To develop an intelligent subsystem in a hybrid intelligent learning environment we use the language Python3 in conjunction with the libraries Scipy, Numpy, Sklearn, Matplotlib. The libraries Scipy and Numpy are general scientific libraries which implement a number of mathematical functions. The Sklearn library provides high-performance clustering and downscaling algorithms. The Matplotlib library is used to render the results.

It is important to note that in this study we implement the proposed methodological approach to the development of modules in the conditions of a hybrid intelligent learning environment using the indicated tools. Besides we create a problem-oriented software package.

From the groups of students who have passed the assessment based on clustering, a subgroup of students with potential opportunities for research activities can be distinguished. Such a group has been distinguished according to the results of the construction of dendrograms and scatter diagrams with Dirichlet tessellation. Examples of dendrograms and scatter diagrams are shown in Fig. 1-4.

Consider the following statement of the problem. Let the students take the “Algebra Test”, Grade 8. A number of didactic materials and a set of tests consisting of several blocks (for example, B1, B2, B3) have been developed for this

purpose.

Block B1 contains the tests that initially specify as the goal and situation as the actions to solve it or the result. These are tests for identifying, distinguishing, and classifying the studied material.

Block B2 presents the tests in which a student needs to apply existing knowledge and previously learned actions according to given conditions. These are substitution tests (“add...”, “list...”), constructive tests according to given active tests (“write down the formula”, “list the attributes...”), and typical tests.

In the tests of block B3 the goal is clearly stated but the situation is “uncertain”. These are the tests that require independent transfer of existing knowledge and skills to a new context.

The answers to the test tasks are formed in the form of a normalized vector where “0” corresponds to the fact that the task was completed incorrectly, “1” corresponds to the fact that the task was completed absolutely correctly.

The purpose of the clustering problem is to find several types of similar test results which can be useful in dividing a class into subgroups and forming group adjusted educational trajectories to eliminate typical gaps in the formation of knowledge.

The most common clustering algorithm is an algorithm based on the generalization of the k-means method (generalized k-means algorithm). The generalized k-means algorithm consists of the following steps.

- 1) Choose n points (centroids) which are the centers of clusters in the first approximation.
- 2) Calculate the centers of gravity among the k vectors near to the centroids.
- 3) If the centers of gravity correspond to the position of the centroids, the work of the algorithm will be finished. Otherwise, the centers of gravity are selected by the new centroids and the transition to step 2 takes place.

The described algorithm has the following advantages: the algorithm converges to a local minimum. There is no possibility of looping the specified algorithm. The disadvantages are the algorithm does not necessarily converge to the global minimum and requires a preliminary determination of the number of clusters.

Forming a complete dendrogram of the obtained test results allows to definite the clustering results. Then we analyze with constructing a dendrogram for the test results presented in Table 1.

Table. 1. Test results

| ID | B1    | B2    | B3    |
|----|-------|-------|-------|
| 0  | ++++  | +++++ | ----+ |
| 1  | +++++ | +++++ | ----+ |
| 2  | ----+ | ----+ | +++++ |
| 3  | +++++ | +++++ | ----+ |
| 4  | +++++ | ++++- | +++++ |
| 5  | ----+ | ----+ | ----+ |
| 6  | ++++- | +++++ | ++++- |

|    |       |       |       |
|----|-------|-------|-------|
| 7  | +++++ | +++++ | +++++ |
| 8  | -+-+  | ---+  | +-    |
| 9  | +++++ | ++++- | ++++- |
| 10 | +++++ | ----  | ----  |
| 11 | ++++  | ----  | ++++  |
| 12 | +++++ | +++++ | ----  |
| 13 | +++++ | ----  | +++++ |
| 14 | ++-++ | +---- | +---- |
| 15 | +++++ | ++++- | ++++- |
| 16 | +---- | +++++ | +++++ |

In Table 1, “plus” and “minus” signs indicate correct and incorrect answers, respectively. The dendrogram of the test results is shown in Fig. 2.

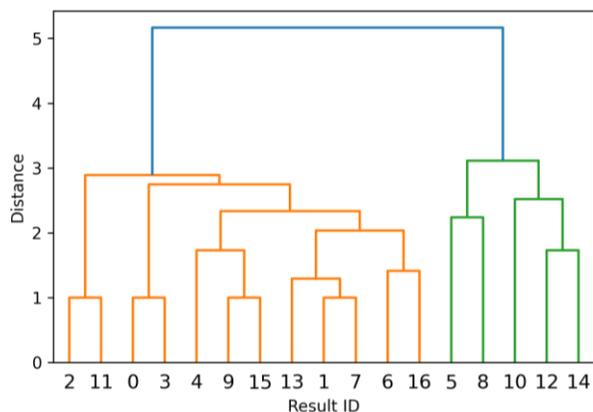


Fig. 2. Dendrogram of testing with distances obtained by Ward's method

The dendrogram presented in Fig. 2 shows that the algorithm divides the test results into two large subgroups. The branches in the left part correspond to high and medium test results, respectively, while the branches in the right part represent abnormally low results.

Then we will consider the dendrogram based on the aggregated test results (Table 2) in the same group of students.

Table 2. Aggregated test results

| ID | B1 (general) | B2 (general) | B3 (general) |
|----|--------------|--------------|--------------|
| 0  | 0.8          | 1.0          | 0.6          |
| 1  | 1.0          | 1.0          | 0.8          |
| 2  | 0.8          | 0.8          | 1.0          |
| 3  | 1.0          | 1.0          | 0.6          |
| 4  | 0.8          | 0.8          | 1.0          |
| 5  | 0.4          | 0.6          | 0.4          |
| 6  | 0.8          | 1.0          | 0.8          |
| 7  | <b>1.0</b>   | <b>1.0</b>   | <b>1.0</b>   |
| 8  | 0.4          | 0.4          | 0.4          |
| 9  | 1.0          | 0.8          | 0.8          |
| 10 | 1.0          | 0.0          | 0.0          |
| 11 | 0.8          | 0.6          | 1.0          |
| 12 | 1.0          | 1.0          | 0.0          |
| 13 | 1.0          | 0.8          | 1.0          |
| 14 | 0.8          | 0.8          | 0.2          |
| 15 | 1.0          | 0.8          | 0.6          |
| 16 | 0.8          | 1.0          | 0.8          |

The dendrogram of the aggregated results is shown in Fig. 3.

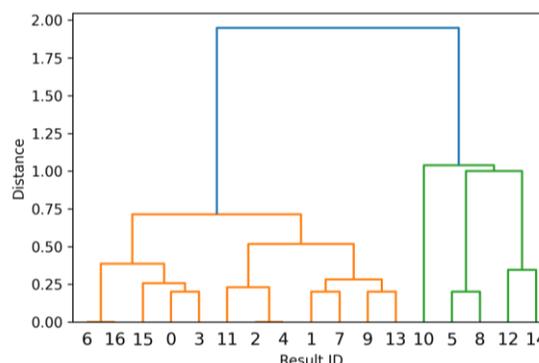


Fig. 3. Dendrogram of aggregated testing metrics with distances obtained by Ward's method

The results presented in the Fig. 3 show that the general principle of clustering the results has not changed. The dendrogram construction algorithm still divides students into two subgroups according to the test results but the internal arrangement of the results has changed a little. According to the implementation in the Fig. 3 method of clustering it can be noted that the differences between the positive test results in comparison with picture are less observed.

In addition, we have analyzed the data using the generalized k-means algorithm. Fig. 4 shows the results of data processing taking into account the construction of centroids of test results by the k-means method.

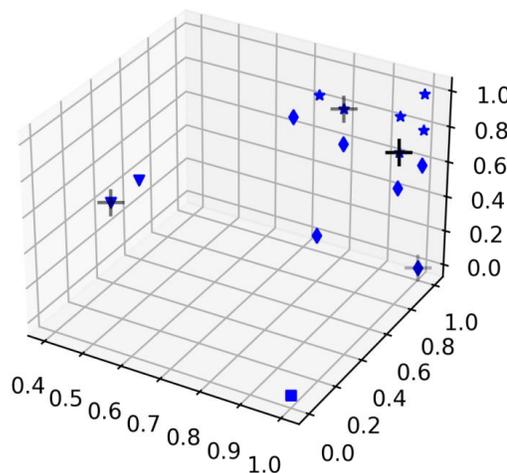


Fig. 4. Construction of centroids taking into account the results of knowledge assessment by the k-means method

In the Fig. 4 markers “triangle”, “square”, “rhombus” and “asterisk” represent the test results for the marks “unsatisfactory”, “satisfactory”, “well”, “excellent”. The “crosses” correspond to the obtained centroids. The marks are calculated according to Table 2.

To implement the scoring module, we use a feed forward artificial neural network with a hidden layer. The architecture of the neural network is shown in the Fig. 5.

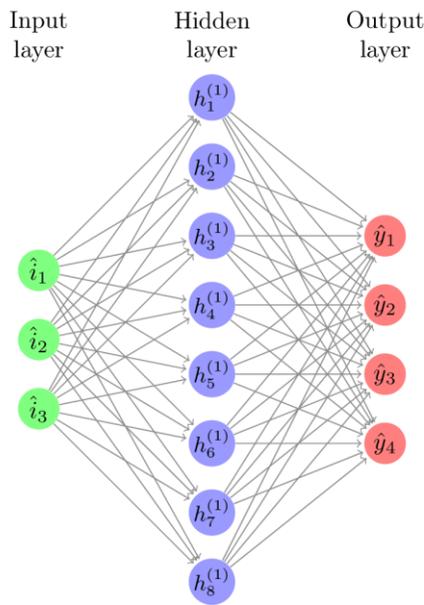


Fig. 5. Artificial neural network architecture for evaluating test results

The values of estimates for each block (features) are fed to the input of the neural network. These values are obtained either by direct estimation (see Table 1) or using dimensionality reduction algorithms using autoencoder networks, or using the principal component method [29]. Using of the tangential activation function for neurons in the hidden layer is assumed. A tangential activation function like the Heaviside function can be used for the neurons of output layer.

The activation of each neuron from the output layer means one of the following estimates: “unsatisfactory”, “satisfactory”, “well”, “excellent”. If several neurons are activated the neuron with the highest value is selected.

An artificial neural network with the specified architecture can be used to assess the results of tests using expert knowledge of the form “test results” – “mark”.

#### IV. DISCUSSION

It should be noted that corrected marks of the results of knowledge assessment are obtained using the k-means method. That is why, we emphasize that the scale of ratings conformity has changed significantly in comparison with the strict ranking of ratings.

Taking into account the used clustering methods based on k-means and dendrograms a data analysis module is developed in the conditions of a hybrid intelligent learning environment.

Using of the for assessing the results of tests within the framework of a hybrid intelligent learning environment methods proposed in the study allows achieving a higher degree of adaptability of the educational process to the level of knowledge of students taking into account goal-setting and promptly adjusting educational trajectories to achieve optimal learning outcomes. Note that the proposed clustering algorithm in the future is aimed both at using test results with the help of expert knowledge and at identifying the creative potential of

gifted students.

Besides an important area of further research is the individualization of the educational process by dividing it into academic groups. The clustering results can be used to form groups with similar characteristics of mastering didactic material among students. It is also possible to consider the reverse approach in which academic groups are created from students with different characteristics of mastering the material. It is also possible to consider the problem of forming academic groups in such a way that these groups are created from students with dissimilar learning characteristics.

#### V. CONCLUSION

The article presents the experience of developing instrumental and methodological support for assessing students' knowledge in mathematics in the framework of a hybrid intelligent learning environment. The significance of the obtained results is considered to be used in developing of an integrated complex of knowledge base and artificial intelligence tools for assessing students' knowledge in mathematics within the framework of a hybrid intelligent learning environment.

The author's contribution is the application of new intelligent technologies and machine learning to the implementation of the didactic model of forming of a bank of tests. We offer a method for analyzing the results of the tests based on machine learning without a teacher.

Obtained results are aimed at developing a set of tests for the knowledge base of the expert system based on the idea of representing the students' activities on a specific academic subject in the form of learning material levels. It is noted these results make it possible to substantiate an approach to the formation of a cognitive problems and to develop criteria for assessing the levels of educational material assimilation. The proposed clustering algorithms find their place in the chain of the educational process “HILE” – “Student” – “Teacher”. These algorithms can be used in solving problems of educational trajectories individualization. It is important to note that the approach to clustering the results of knowledge assimilation is developing in addition to the well-known assessment methods according to which the role of the teacher is decisive.

The proposed schemes of intellectual assessment of knowledge models and creative potential of schoolchildren in mathematics allow formalizing the structure of the corresponding knowledge base. The sufficient effectiveness of the approach to the development of a knowledge base within the framework of data mining methods for creating a module for monitoring and assessing knowledge in mathematics is demonstrated.

It is important to emphasize that further research is developing and implementing of new algorithms formed into the corresponding modules of the intelligent learning environment especially hybrid clustering and classification algorithms as well as hybrid neural network algorithms.

#### ACKNOWLEDGMENT

The reported study was funded by RFBR, project number 19-29-14009

#### References

- [1] G. J. Hwang, P. H. Wu and C. C. Chen, "An online game approach for improving students' learning performance in web-based problem-solving activities", *Computers & Education*, vol. 59, no. 4, pp. 1246–1256, 2012.
- [2] D. H. Jonassen, "Toward a design theory of problem solving", *Educational Technology Research and Development*, vol. 48, pp. 63–85, 2000.
- [3] D. Gibson, C. Aldrich and M. Prensky, *Games and simulations in online learning: research and development frameworks*. Information Science Publishing, 2007.
- [4] S. Graf, *Adaptivity in learning management systems focusing on learning styles* (Ph.D. Dissertation). Vienna, 2007.
- [5] P. D. Basalin, A. E. Timofeev, E. A. Kumagina, E. A. Neymark, I. A. Fomina and N. N. Chernyshova, "The implementation of a hybrid intelligent learning environment of the product type", *Modern Information Technology and IT-education*, vol. 14, no. 1, pp. 256–267, 2018.
- [6] A. F. Sleptsov and M. V. Sleptsova, "Intellectual Educational Environment: Theoretical Approaches and Implementation Opportunities", *Contemporary Research on Social Problems*, no. 5, pp. 70-88, 2016.
- [7] A. S. Obukhov, E. I. Bulin, E. I. Sokolova and A. L. Semyonov, "The future of pedagogical education. Direction of movement and first practical steps", *Psychological Science and Education*, vol. 19, no. 3, pp. 207–225, 2014.
- [8] O. V. Druzhinina, E. V. Igonina, O. N. Masina and A. A. Petrov, "Aspects of the use of prototyping and artificial intelligence technologies as part of the digital transformation of the educational process", *Modern Information Technology and IT-education*, vol. 16, no. 1, pp. 65-74, 2020.
- [9] O. V. Druzhinina, O. N. Masina, A. A. Petrov and S. V. Shcherbatykh, "Application of intelligent technologies and neural network modeling methods in the development of a hybrid learning environment", *Journal of Physics: Conference Series*, vol. 1691, no. 1-6, p.012125, 2020. Available: <https://iopscience.iop.org/article/10.1088/1742-6596/1691/1/012125>.
- [10] ISO/IEC/IEEE 24765-2010(E), *Systems and software engineering. Vocabulary. IT-Regie. The business of IT for the business*. Available: <https://www.smaele.nl/documents/iso/ISO-24765-2010.pdf>.
- [11] L. A. Kandybovich and A. V. Mudrik, *Mehnot a Student of Knowledge (Terminological dictionary-reference)*. Minsk: Harvest, 2010.
- [12] A. A. Mitskevich, *Knowledge Bases and Expert Systems*. Moscow: MIU, 1989.
- [13] G. S. Pospelov, *Artificial Intelligence*. Applied Systems. Moscow: Knowledge, 1985.
- [14] T. A. Gavrilova, *Knowledge Base of Intelligent Systems*. St. Petersburg: Piter Publishing House, 2000.
- [15] T. A. Gavrilova and N. V. Gulyakina, "Visual methods of working with knowledge: an attempt to review", *Artificial Intelligence and Decision Making*, no 1, pp. 15-21, 2008.
- [16] F. S. Voroiisky, *Informatics. A New Systematized Explanatory Dictionary (Introduction to Modern Information and Telecommunication Technologies in Terms and Facts)*, Moscow: FIZMATLIT, 2003.
- [17] A. B. Sorokin, E. V. Brazhnikova and L. M. Zheleznyak, "Designing a knowledge base for expert systems based on the conceptual structure of an act of activity", *High-performance Computing Systems and Technologies*, vol. 4, no. 1, pp. 190-196, 2020.
- [18] D. Yu. Muromtsev, V. V. Ermolaev and A. Yu. Kotok, "Methods for designing a knowledge base for active expert systems", *Problems of Modern Science and Practice*. University named after V.I. Vernadsky, no. S2(52), pp. 92-96, 2014.
- [19] Zh. I. Batyrkanov and K. D. Boskebeev, "Knowledge base on the basis of a semantic model", *Bulletin of the Kyrgyz State Technical University by name of I. Razzakov*, no. 23, pp. 20-25, 2011.
- [20] G. M. Yavorskaya, "Development of an expert system using a pedagogical model for assessing the quality of students' knowledge", *Avtomatika. Informatics*, no. 1-2(22-23), pp. 99-101, 2008.
- [21] I. A. Karpacheva, V. S. Merenkova and S. N. Dvoryatkina, "Management of the learning process: from models to intellectual management", *Psychology of Education in Multicultural Space*, no. 3(51), pp. 95-109, 2020.
- [22] V. P. Bepalko, *Components of Pedagogical Technology*. Moscow Pedagogy, 1989.
- [23] E. I. Smirnov, S. N. Dvoryatkina and S. V. Shcherbatykh, "Intelligent management in mathematical modeling of the research activity of schoolchildren", *Scientific and methodical journal "CONTINUUM. Mathematics. Computer Science. Education"*, no. 3(19), pp. 48-61, 2020.
- [24] V. P. Bepalko, "Typical pedagogical errors of testing in education", *Shkolnye tekhnologii*, no. 5, pp. 3–11, 2012.
- [25] S. Haykin, *Neural Networks: A Comprehensive Foundation*. Prentice Hall, 1999.
- [26] A. Karpatne, "Theory-Guided Data Science: A New Paradigm for Scientific Discovery from Data", *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 10, pp. 2318–2331, 1 Oct. 2017.
- [27] A. Geron, *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. 2nd New edition. O'Reilly, 2019.
- [28] W. Mckinney, *Python for Data Analysis, 2e: Data Wrangling with Pandas, Numpy, and Ipython*. 2nd ed. Edition. O'Reilly, 2017.

- [29] D. Lakshmi Padmaja and B. Vishnuvardhan, "Comparative Study of Feature Subset Selection Methods for Dimensionality Reduction on Scientific Data", IEEE 6th International Conference on Advanced Computing (IACC), 2016, pp. 31–34.

**Creative Commons Attribution License 4.0  
(Attribution 4.0 International, CC BY 4.0)**

This article is published under the terms of the Creative Commons Attribution License 4.0

[https://creativecommons.org/licenses/by/4.0/deed.en\\_US](https://creativecommons.org/licenses/by/4.0/deed.en_US)