Methods of intelligent results ranking for the intermediate assessment of knowledge in mathematical disciplines

O N Masina¹, A A Petrov¹, O V Druzhinina² and S V Shcherbatykh¹

¹Bunin Yelets State University, 28, Kommunarov str., Yelets, 399770, Russian Federation
²Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, 44, building 2, Vavilov str., Moscow, 119333, Russian Federation

E-mail: olga121@inbox.ru

Abstract. We are developing methods of intelligent ranking of results for the intermediate assessment of knowledge on mathematical disciplines within the framework of a hybrid intelligent learning environment functioning. An extended structural diagram of a hybrid intelligent learning environment is presented and the methodology of forming a knowledge base is characterized. The study of new models of the pedagogical process is carried out using intelligent technologies. Clustering problems are solved and algorithms for analyzing test results based on the use of machine learning methods are proposed. The results of the proposed algorithms implementation are presented. The results of this study are of practical importance both for the development of methods for assessing knowledge in mathematical disciplines, and for identifying the abilities of students to research activities.

1. Introduction

Currently, in the learning process, it is important to design individual educational routes depending on the level of training and individual psychological characteristics of students in a hybrid learning environment [1–3]. Actual problems arise when creating a hybrid learning environment in the subject area associated with teaching mathematics in secondary school [4–5]. In this aspect, an important component of the educational activity of school is the systematic organization of continuous design and research activities of students based on the active use of digital technologies.

In [6] such a procedural model of design and research activities of a student, which is focused on preparing a school graduate for life in the emerging digital economy is developed. The design and research activity of a student is considered as a joint activity of the subjects of the educational process, including the subjects of the educational cluster “university-school”, to design a study aimed at solving problems of the surrounding reality based on the active use of digital technologies.

An important component of the educational activity of the school in modern conditions is the systematic organization of such continuous design and research activities of students, focused on their preparation for life in the digital economy [7].

In this paper, we develop methods for the intellectual analysis of test results within the framework of a hybrid intelligent learning environment (HILE) functioning. Section 2 provides an extended structural diagram of the HILE, describes the methodology for the formation of the knowledge base. In addition, algorithms for analyzing the test results are proposed, and a stochastic model of
competencies is synthesized. Section 3 presents the proposed algorithms implementation results. In Section 4, we discuss the practical significance of the results both for the development of methods for assessing knowledge in mathematical disciplines, and for identifying the abilities of schoolchildren for research activities.

2. Models and methods

The intelligent ranking methods of knowledge assessment intermediate results in mathematical disciplines are being developed as part of the creation of a hybrid intelligent learning environment (HILE). The purpose of the development of this HILE is to increase the automation degree of the pedagogical process by solving the following problems:

- The learning process modeling to identify qualitative effects that arise in the pedagogical process;
- Educational trajectories individualization through the automated generation of didactic and assessment materials;
- Training groups formation based on the results of intermediate testing;
- High creative potential of the student identification;
- Construction of infographics using modern computer tools.

The structure of the HILE is shown in figure 1.

![Figure 1. Structural diagram of a hybrid intelligent educational environment.](image)

We note that HILE is a closed system “HILE” – “students” with feedback. The knowledge module is a database and a logical solver required to generate assessment materials. The AI subsystem is responsible for collecting, analyzing, evaluating test results, as well as for the formation of individual educational trajectories. The base of educational trajectories is designed to save personalized information about the evolution of the knowledge level of each student. The tests generation module is responsible for generating assessment materials. It should be noted that the functioning of the HILE occurs in the mode of transactions, which are verified by the teacher.

An important issue in the organization of the pedagogical process is the problem of evaluating the results of tests performing. We are considering such a format of test results, which consists of blocks A1, A2, A3 of standard difficulty and block A4 of increased difficulty. The tests of blocks A1–A4 are of increasing difficulty.
Block A1 contains tests for which a solution algorithm is defined. Block A2 presents typical tasks of a constructive nature. Block A3 contains tests that require the transfer of knowledge to a new situation. Box A4 presents tests that require creativity.

In blocks A1, A2, A3 tests are presented, taking into account the fact that the correct answer corresponds to one point, the wrong answer corresponds to zero point.

In the Bespalco concept [8] a formalized indicator of a student's achievement the assimilation level of educational material (level of training) is the assimilation coefficient $k$. The specified coefficient is calculated as the ratio of the number of correct answers given by the student to the total number of correct answers (reference). The learning process can be considered completed at a given level of assimilation if the coefficient $k \geq \sigma$, where $\sigma = 0.7$. A student who demonstrates this level of mastering the material is able to improve his knowledge in the future in the process of self-education. When carrying out assessment procedures, the systematic nature of the mistakes made is taken into account, the need to carry out additional training.

The fixed grading scale for test results in mathematical disciplines can be represented as follows:

- An unacceptable level of assimilation, if correctly completed less than $\sigma N_i$ tests, where $N_i$ is number tests of block $A_i$, $i = 1, 2, 3$;
- A low level of assimilation, if correctly completed more than $\sigma N_i$ tests;
- The average level of assimilation, if correctly completed more than $\sigma N_i$ tests;
- A high level of assimilation, if correctly completed more than $\sigma N_i$ tests.

In comparison with the method described in [8], we propose to add a new block A4. The results of the control in block A4 may not affect the assessment of academic performance. At the same time, these results contribute to the identification of mathematical ability. The specified block is necessary to identify the creative potential in research activities. The test results of block A4 are evaluated by an expert on a continuous scale (for example, from $-1$ to $1$).

To refine the results of estimation on a fixed scale, we use a ranking based on machine learning methods [9]. The purpose of ranking is to divide students into groups according to test results. The number of such groups is predefined, but the structure of each group is determined as a result of the machine learning algorithm. We use unsupervised machine learning to solve clustering problems with further results sorting.

We propose a ranking algorithm that consists of the three steps and uses the notion of centroid. Centroids are determined using a cluster analysis algorithm ($k$-means, $k$-median, or genetic clustering algorithm). At the sorting stage, it is possible to weigh the individual components of the vector.

It is important to note that the proposed ranking algorithm cannot be used instead of a fixed rating scale. This is due to the fact that, regardless of students preparation level in the general group, this algorithm divides the results into 4 clusters. Within the author’s methodology, the ranking algorithm is applied in stages two times: for blocks A1, A2, A3, and then separately for block A4. We develop a stochastic model to simulate the process of test performing. For blocks A1–A3, the correct answer is determined as follows:

$$
A_{ij} = 1 \quad \text{if} \quad \Phi(\mu_j, \sigma_j) > l,
A_{ij} = 0 \quad \text{if} \quad \Phi(\mu_j, \sigma_j) \leq l,
$$

where $A_{ij}$ is the number of the $j$-th question of the $i$-th block, $\Phi$ is the probabilistic function of the normal distribution, which determines the level of mastering the competence, $\mu$ and $\gamma$ are the coefficients of the probability function, $l$ is the threshold of mastering the competence required for the correct answer.

The coefficients of the probability function are set by a matrix that determines the individual characteristics of each of the competencies. We proceed from the assumption that each question tests
only one competency. Note that it is of interest to study the generalization of this assumption to the \( n \)-dimensional case.

To evaluate the results of block A4, the formula:

\[
A_{ij} = \Phi(\mu_{ij}, \sigma_{ij})
\]

(2)

is applied, where the designations are explained above. For block A4, a specialized competency matrix is applied.

3. Results

To carry out computational experiments, we have developed a software package in Python 3 using the mathematical libraries Numpy, Scipy, Scikit-learn [10]. The k-means method is used for clustering. The questionnaire includes 60 exercises: 15 exercises per block, 3 exercises per competency. A total of 50 test results are considered. The difficulty of the exercises increases from block A1 to block A4. The results of ranking the estimates for blocks A1–A3 are shown in figure 2.

![Figure 2](image1.png)  

**Figure 2.** Test results by blocks A1–A3 according to a fixed grade scale.

The values \( k \) (mean values of correctness of answers) for blocks A1–A3 are presented along the \( x \)-axis, \( y \)-axis and \( z \)-axis (figure 2). Class membership is determined by color marks: red is “low”, yellow is “moderate”, blue is “good”, green is “excellent”. Figure 2 shows grading according to a fixed scale. Figure 3 shows the results of ranking based on machine learning. Note that there is a results redistribution at the cluster boundaries.

In figure 4 the results of the assessment are presented in relation to block A4 (research block).

The \( x \)-axis and \( y \)-axis show the average values for the A1–A3 blocks and the average value for the block A4, respectively. Class membership is determined by color marks: red is “low creativity”, yellow is “medium creativity”, green is “high creativity”. Note that low results in block A4 classify a student as “low creativity” regardless of the results of blocks A1–A3.

Further figure 5 shows the results of constructing a dendogram based on the test results full vectors for blocks A1–A4. Test vectors are sorted according to the mean value principle.
Figure 4. The results of ranking testing by blocks A1–A3, A4.

Figure 5. Dendrogram based on the results of testing blocks A1–A4.

Figure 5 the y-axis shows the identifier (ID) of the result and the distance according to the Ward metric [11]. The closest test results are grouped as branches. The analysis of typical errors by the method of constructing dendrograms allows one to draw conclusions about both the quality of the assessment materials and the violation of the rules for conducting the test.

4. Discussion

The development of systems for intellectual support of the pedagogical process is an urgent scientific direction. Solving the problems of analyzing the results of the intermediate assessment of knowledge in mathematical disciplines is an important stage in the development of HILE. Possible directions for further research are the development of new models of the pedagogical process, as well as solving the problems of intellectual interaction with a student in a natural language. We note the theoretical and applied interest in the problems of the development of HILE using cognitive modeling methods and in the development of new algorithms for data analysis.

The obtained results in Section 3 demonstrate the importance and applied significance of an in-depth analysis of the results of tests. Results with ID 49, 26, 0 demonstrate qualitative differences in identifying similar test results compared to the assessment based on the mean value.
The ranking of results is carried out in two stages. At the first stage, the results are divided into groups “low”, “moderate”, “good”, “excellent” according to the questionnaire survey in the framework of the compulsory educational program. At the second stage, the results are divided into the groups of “low creative potential”, “medium creative potential” and “high creative potential”.

Note that the analysis of test results is aimed at using the synergistic effect in education. On the one hand, close results of the survey indicate a similar level of competence development, which makes it possible to effectively carry out group correction of knowledge. On the other hand, the approach with the formation of groups based on dissimilar survey results can be effective in the problems of mutual learning of students. We plan to test this approach when forming creative groups with common scientific projects in mathematical disciplines.

5. Conclusion

The paper is devoted to the development of intellectual analysis methods of the test results in the framework of a hybrid educational intellectual environment. The study is aimed at the formation of individual educational trajectories of students and the identification of creative potential in the process of teaching mathematical disciplines.

The developed method of intellectual ranking of results for the intermediate assessment of knowledge in mathematical disciplines allows solving the problems of organizing the pedagogical process in the context of digitalization of education. The solution of clustering problems and the proposed approach to the analysis of test results contribute to the personalization of education. The software package developed taking into account the proposed methods and algorithms can be used in the technique of assessing knowledge and for identifying of the students abilities for research activities. A possible direction for the development of instrumental and methodological support for HILE is the implementation of interaction with the user in natural language and the involvement of various methods of machine learning (deep learning, reinforcement learning, etc.) [12].

Acknowledgments

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

References

[1] Basalin P D, Timofeev A E, Kulagina E A, Neymark E A, Fomina I A and Chemyshova N N 2018 Implementation of a hybrid intellectual learning environment of a production type. *Modern information technologies and IT-education* 14(1) 256–267
[2] Sleptsov A F and Sleptsova M V 2016 Intellectual educational environment: theoretical approaches and opportunities for implementation *Modern studies of social problems* 5 70–88
[3] Obukhov A S, Bulin-Sokolova E I and Semyonov A L 2014 Future teacher education. The direction of movement and the first practical steps. *Psychological Science and Education* 19(3) 207–225
[4] Druzhinina O V, Masina O N and Petrov A A 2021 Construction of differential mathematical models used in the development of a hybrid intelligent learning environment, taking into account the delay and control actions. *Continuum. Mathematics. Computer science. Education* 1(21) 69–80
[5] Druzhinina O V, Masina O N, Petrov A A and Shcherbatykh S V 2020 Application of intelligent technologies and neural network modeling methods in the development of a hybrid learning environment. *Journal of Physics: Conference Series* 1691 012125
[6] Minina I V and Petukhova T P 2019 Organization of design and research activities of a schoolchild in modern conditions. *Modern information technologies and IT-education* 15(4) 1031–1046
[7] González-Marcos A, Alba-Elias F, Navaridas-Nalda F and Ordieres-Meré J 2016 Student evaluation of a virtual experience for project management learning: an empirical study for learning improvement. *Computers & Education* 102 172–187
[8] Bespalko V P 1989 *Components of pedagogical technology* (Moscow: Pedagogy) 254
[9] Geron A 2019 *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems* (Boston: O’Reilly) 436
[10] Mckinney W 2017 *Python for Data Analysis, 2e: Data Wrangling with Pandas, Numpy, and Ipython* (Boston: O’Reilly) 231
[11] Ziv B-J, Gifford D K and Jaakkola T S 2001 Fast optimal leaf ordering for hierarchical clustering. *Bioinformatics* **17** 22–29
[12] Haykin S 1999 *Neural Networks: a Comprehensive Foundation* (Englewood Cliffs, NJ: Prentice Hall) 274