Adaptive Testing Model as the Method of Quality Knowledge Control Individualizing

Nonna Shapovalova[0000-0001-9146-1205], Olena Rybalchenko, Iryna Dotsenko, Svitlana Bilashenko, Andrii Striuk[0000-0001-9240-1976] and Levan Saitgareev

Kryvyi Rih National University, 11, Vitaly Matusevich Str., Kryvyi Rih, 50027, Ukraine

{shapovalovann09, ellinaryb, irado441, SvitlanaViktorivnaBilashenko, andrey.n.stryuk}@gmail.com

Abstract. The mission of the work is to develop and theorize the efficiency of application of the knowledge control system on the basis of adaptive testing technology, which combines the specifics of the professional and educational activity and the monitoring of the quality of training and the possibility of self-control of students, to develop a set of test assignments in the discipline "Artificial Intelligence Systems". Object of research is a software tool for monitoring students' knowledge in higher educational establishment. The subject of research is the development of software for an adaptive knowledge control system using machine learning device. Research goals: to develop a set of test case of different levels of complexity; to determine the structure, architecture and specificity of the application of the machine learning algorithm for the formation of a variable level of testing complexity for each student; develop appropriate software, guidelines and recommendations for adjusting and distributing issues by level of complexity. The result of the work is a complex of split-level application-oriented tasks for current and module control in the discipline "Artificial Intelligence Systems", web-oriented software that allows you to quickly monitor the quality of students' knowledge and is appropriate for use in online and mixed mode of training.

Keywords: adaptive testing, machine learning, psychological types of personality.

1 Introduction

In the modern educational space there are many forms and methods for controlling students' knowledge and skills. The effectiveness of both systems of the educational process and the quality of specialist training depend on the proper management of the assessment. Adaptive testing is a technology for determining the level of students' knowledge, where each next question is automatically selected based on the answers to previous questions and a predetermined level of complexity. The main difference between adaptive testing and classical tests is the dynamic (in real time), and not the static definition of the list of questions. In this case, the choice of the next question is determined by the personal characteristics of each individual student.

The use of the system of adaptive control of knowledge makes it possible to solve a number of urgent tasks such as the creation of subject complexes of test tasks; creation of tools for individualized diagnostics of the level of knowledge and degree of material digestibility; the formation of a visual representation and interpretation of test results. A feature of the use of adaptive technologies in education is that the quality control of knowledge occurs constantly in the process of teaching the discipline, and not only during the modular or session control. This fact allows the teacher to have operational and objective information about the quality of learning material, its correct understanding of the independent work of students and the like. Thus, adaptive learning systems, in particular adaptive testing, make it possible to optimize the learning process.

2 Research apparatus

The aim of the study is to theoretically substantiate, develop and experimentally test the system of adaptive testing on the discipline “Artificial Intelligence Systems” in the professional training of bachelors in the specialty 121 – “Software Engineering”.

Objectives of the study:

1. To analyze the state of the problem of developing and using adaptive software for quality knowledge control.
2. To substantiate the choice of methods for implementing the system of knowledge quality control, taking into account the individual characteristics of the student.
3. To develop a universal system of adaptive control of knowledge in a particular discipline.
4. To fill the software-instrumental environment with content of test tasks of various levels of complexity to determine the quality of knowledge of bachelors in the specialty 121 – “Software Engineering” in the discipline “Artificial Intelligence Systems”.

The object of research is the development of web-based software with an adaptive interface for determining the level of knowledge.

The subject of research is the development of an adaptive system for testing bachelors in the specialty 121 – “Software Engineering” in the discipline “Artificial Intelligence Systems”.

Research methods: analysis of sources from the investigated topic, methods for determining the psychological type of a person, methods of artificial intelligence for the individualization of knowledge control; modeling the learning process of classification algorithms; formalization of the constructed models; empirical method for determining the optimal parameters of the training model; method of object-oriented design and programming.

The practical significance of the obtained results is that a complex of tests in the discipline “Artificial Intelligence Systems” of various degrees of complexity has been developed for the bachelors in the specialty 121 – “Software Engineering”. The complex of test tasks can be used to determine the quality of knowledge in various disciplines of the vocational training cycle, involving the mastering of artificial intelligence technologies; the developed software can be used as an ICT tool in the learning process of any discipline.

3 Theoretical foundations of adaptive testing

The system of students’ knowledge control, which is an important component of any form of education, still needs to be updated and developed such knowledge control tools that meet the requirements of objectivity, comparability, predictability of assessment results and have clear criteria for assessment procedures.

The organization of the educational process is currently characterized by the widespread use of information and communication technologies, primarily the introduction of testing using computer technology. On the issues of computer testing to determine the level of students’ knowledge, a lot of research has been conducted, software tools have been created for generating and testing tests. Automated testing systems provide an increase in the efficiency of the educational process and are economical, but the lack of an individual approach to students, taking into account their personal characteristics, significantly reduces the objectivity, comparability and predictability of assessment results.

T. Hodovaniuk notes that the observance of the principle of an individual approach to teaching in higher education requires taking into account the level of intellectual development of a student, constant analysis of his academic and life experience, taking into account the level of independence and volitional development of each person. This approach is aimed at achieving students a common goal, but in different ways. The essence of this principle of learning is the multi-level independent activity of students, which purpose is to facilitate the assimilation of educational material in accordance with individual mental abilities and the existing level of students’ knowledge [12].

As M. Mazorchuk notes, traditional automated testing systems have a significant drawback: each subsequent test task is generated by random selection from the entire set of tasks of a particular topic [17]. With such an algorithm, situations where a student with a low level of knowledge is presented with complex tasks are not uncommon; the result is an almost complete lack of answers. Polar situation will be such that when a student with a high level of knowledge will be offered light tests. In this case, the tested person will not be able to realize their abilities. That is, to ensure objectivity and comparability of the results, the average complexity of the test task must correspond to the predicted level of student training. Such a selection of tasks can be accomplished using adaptive testing.

By definition of A. Malygin, adaptive testing is a scientifically based method of controlling the level of knowledge of students, which is implemented, using automated processes for generating, presenting and evaluating the results of performing adaptive tests. Each subsequent test question is automatically selected based on the responses received to previous questions and a predetermined level of difficulty [16]. The main difference between adaptive testing and classical tests is the dynamic (in real time), and not the static definition of the list of questions. The choice of the next question is determined by the personal characteristics of each individual student.

Adaptive testing makes knowledge control procedures effective due to an individual approach and offering the student tasks corresponding to his level of training [19]. According to S. Zahrebelnyi, a tested face can be presented with fewer tasks with preservation of the diagnostic ability of the whole volumetric test [32]. Due to the adaptive approach, it is possible to significantly reduce the complexity and testing time.

The issue of adaptive testing was studied by many scientists and practitioners (A. Malygin [16], S. Zahrebelnyi [32], M. Brus [32], M. Mazorchuk [17], Yu. Koltsov [14], N. Dobrovolskaya [14], and others).

The scientific base of adaptive testing is the modern theory of tests – Items Response Theory (IRT), its provisions and parametric methods are examined in the research by G. Rash [21] and others.

The main idea of IRT is to justify the possibility of effectively predicting test results for tasks of different levels of complexity, which is a necessary requirement for adaptive test control systems.

The forecast is based on the following statements:

1. There are dormant parameters of the personality that are unattainable for direct observation. In testing this is the level of preparation of the tested person and the level of difficulty of the task;
2. There are indicative variables associated with dormant parameters available for direct observation. The values of the indicator variables give information about the value of the dormant parameters;
3. Dormant parameter, it is estimated to be one-dimensional. This means that, for example, a test has to measure knowledge in one subject area.

To implement the adaptive testing algorithm in this work, the one-parameter model IRT G. Rash is used. The model reflects the probability of success of the test as a function of one parameter – the difference in the level of training of the subject and the level of difficulty of the task [21].

$$P\left(u_{ij}=1\middle|\theta_j,\delta_i\right)=\frac{e^{\theta_j-\delta_i}}{1-e^{\theta_j-\delta_i}} \tag{1}$$

where u_{ij} – an estimate for the j-th for i-th task; θ_j – level of preparation of j-th subject; δ_i – characteristic of the i-th point of the test.

Model (1) is a logistic function, and its graph depicting the probability of a correct answer from the latent characteristic θ is called the characteristic curve of the task (item characteristic curve (ICC). In [2] V. Avanesov noted that test questions, besides the fact that they must differ in the level of complexity, must meet the following requirements: brevity; manufacturability; correct form; correct content; logical form of expression; uniformity of rules for evaluating responses and can be presented in various types: closed (multi-alternative and single-alternative), open, to establish correspondence between elements, to establish the correct sequence, situational test items.

M. Bondarenko, V. Semenets, N. Belous devoted their research to the question of evaluating test tasks of different types [4].

Formally, the model for evaluating the results of adaptive testing can be represented as a differential equation [1]. The respondent needs to perform N test tasks of the level of difficulty $d_i = [1, D]$, and the result of the test T depends on the results of the previous tasks and therefore changes continuously. For the assessment of knowledge, a continuous rating scale is used in the range [0, 1], and the assessment for the performance of each test task is measured by a coefficient $t_i \in [0, 1]$. In this case, it is expedient to introduce a coefficient B in recalculating test results into an arbitrary points system. So, the lowest score is 1, and the highest – B points. Since any knowledge assessment system based on test execution makes it possible to guess the correct answer, it is necessary to enter a guessing coefficient $g_i \in [0, 1]$. Taking into account all the parameters, the model for evaluating the results of adaptive testing can be represented as (2).

$$f(N, d_i, t_i, g_i, B)=1 \tag{2}$$

For each of the types of test items, the coefficient t_i is calculated differently.

In this paper, the authors will propose to use closed single-alternative and multi-alternative test items, taking into account the possibility of guessing the correct answer. For evaluation of test tasks of different types, the technique proposed by N. Belous is used [4]. For problems with one correct answer, t_i is a binary value, takes the value 0 in the case of an incorrect answer and the value 1 in the case of choosing the correct answer option. In the case of test questions with several correct answers, the student may or may not have all the correct answers, or one or more incorrect answers, that is, to calculate t_i, it is necessary to consider not only the correctness of the answer to the task as a whole, but also the number of answer choices correctly chosen (2).

$$t_i = \frac{Qr_i}{QR_i + Qw_i}\ , \tag{3}$$

where Qr_i – the number of correct answers chosen by the respondent in the i-th task; QR_i – the number of correct answers in the i-th task; Qw_i – number of incorrect answers selected by the respondent in the i-th task.

The guessing coefficient for tasks with one correct answer among Q answers is calculated by the formula (3).

$$g_i = \frac{1}{Q} \tag{4}$$

In the case of several correct answers, the coefficient of guessing probability is calculated by the formula (4).

$$g_i = \frac{1}{2^Q} \tag{5}$$

The task selection algorithm is based on the feedback principle: if a student answers correctly, the next task is selected at the highest level of complexity, and if the answer is incorrect, a low level is selected than the one to which the student gave the wrong answer. During the first testing, the respondent level is set to R0=0, and the first question of the test will automatically be offered from a high level of complexity. Depending on the correctness of the answers to the proposed questions, the level of preparedness is listed. Thus, after completing the next test task, the following is selected such a level of complexity, which is calculated on the basis of the previous answer. An example of the procedure for selecting the level of complexity of each of the following questions is shown in Fig. 1. Despite all the advantages, in the concept of adaptive testing is not sufficiently implemented the assessment of the personal parameters of the test subject. Accounting in the system of adaptive learning traits of the personality of the student allows you to effectively achieve the goal of learning.

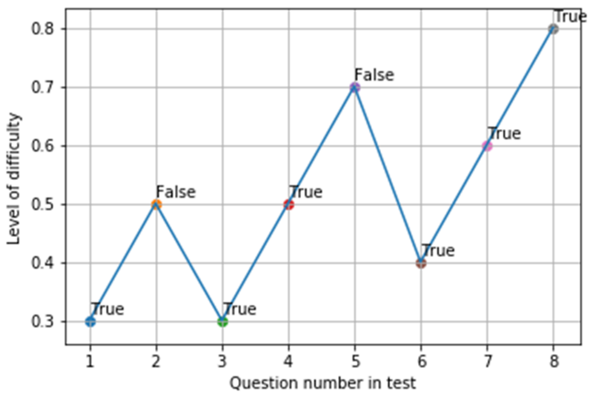


Fig. 1. Procedure for selecting the level of complexity of questions in the test

During the study, Yu. Koltsov and N. Dobrovolskaya came to the conclusion that it is reasonable to include the following qualities and personality characteristics to the core of the student's model: type of thinking; the form of knowledge representation is better perceived by the person; confidence when answering; level of learning. It is the consideration of these features that should influence the formation of the material that is submitted for testing according to the level of complexity and forms of knowledge representation [15].

Individual characteristics are a stable characteristic of a person that does not change at all or changes over a long period of time. They are determined using specially developed psychological tests. Persons who study may have an intuitive or theoretical-methodological thinking, different levels of anxiety during the response, motivation to learn and features of the processes of remembering and forgetting.

Accordingly, a task with different forms of knowledge representation can be offered for different categories of persons: analytical (analytical expressions, mathematical models, formalized descriptions), figurative (schemes, drawings, video fragments), heuristic (practical methods, heuristic descriptions).

4 Definition of psychological type

In the process of learning, innovative technologies are increasingly being introduced, one of which is the student-centered approach, which underlies many learning styles or models. The most popular learning model was the VAK model based on the psychophysiological features of information perception [30]. From the point of view of psychology, each person has his own psychological type, which is expressed in combination of character traits, describes its individuality and personality. Based on the psychophysiological features of perception of information, four psychological types are distinguished: audials, visuals, kinesthetic and digital (discretes). Visitors perceive most of the information through vision. This psychological type is the most common, since about 80-90% of information a person receives through vision. Audials are a rather rare type of people (5-7%), in which the auditory perception of information is more developed. The kinesthetic perceives reality through tactile sensations, in order to obtain any new knowledge or skills they must work it out on their own practical experience, but digitals (discretes) think in terms of functionality, using logic and numbers. The VAK model takes into account physiological properties when choosing the most appropriate way to perceive information depending on psychological type.

Psychologists and educators V. Barb and M. Milon announced the results of a detailed study of the introduction of VAK-techniques. They conducted an experiment, attracting 1000 students from southern California and proved that students learn the material more efficiently, and learning becomes preferable when taking into account the psychophysiological features of the perception of new information [3]. Studies by V. Barb have argued that when using the VAK model, the perception of information is facilitated, as a result, the learning process is improved. N. Fleming based on the model of V. Barb identified another psychological type [8] and the VAK model (visual / audial / kinesthetic) was expanded to the VAKR model (visual / audial / kinesthetic / reading). The teaching style according to the VAK methodology has become the most popular in the world teaching practice due to the simplicity of model building and ease of its use [28].

There are many techniques that allow you to determine a person's psychological type, but provided that this process is implemented as a separate software module, only those techniques that do not provide for visual contact with the subject will be considered. In order to select a method for determining a student's psychological type according to information perception, a number of psychological tests were considered: a test for determining the psychological type and propensity to work in groups or individually J. Reid [18]; diagnostics of the dominant perceptual modality of S. Yefremtsev [31]; a test for determining a certain integral indicator of the general abilities of V. Buzin, E. Vanderlik; methodology "Register of information assimilation style" A. Gregos (identifying priority methods for collecting information) [22].

For this work, the test "Diagnosis of the dominant perceptual modality of S. Yefremtsev" was chosen, as such, which most closely meets the following criteria: age of the test person, purpose, number of questions. The use of this test method makes it possible to determine the type of perception of the subject's information. In many scientific papers, in determining the psychological type of a person, the authors rely on the test of S. Yefremtsev [5; 10].

5 Data and methods of individualization of control knowledge

5.1 Structure of the main modules

Designing an adaptive knowledge control system consists of the implementation of three modules that meet the objectives of the implemented system and constantly interact in the process of using the system: data collection; selection of content; personalization. The data collection module consists of primary (input) testing and saving test results for the discipline. At this stage, a student's model is formed on the basis of his psych-type, type of information perception channel (visual, audial, kinesthetic, digital), as well as the accumulation of information about the student's knowledge of the degree of assimilation of certain concepts.

Depending on the patterns of student behavior in the system, its features, reactions to changes in the levels of complexity of selected tasks, information about the user's image is summarized and the content is formed for certain types. So, it can be schemes, graphics, parts of a program code, objects of augmented reality, etc. [27, 29]. The implementation of the content selection module is based on the task of determining the optimal parameters for the type of presentation of the image of a student through the implementation of a machine learning algorithm that simulates a typical case study task. The personalization module includes the differentiation of tasks according to the level of complexity, displaying the progress of each student's success, the test forecast based on the number of attempts to pass the test and the average result achieved during all attempts.

5.2 Classification methods

To implement content selection and personalization modules, it is necessary to solve the problem of classifying the image of each user in order to select the exact level of complexity of the next test task corresponding to the level of knowledge and the type of perception of a particular student. To solve this problem we will use the methods of machine learning (ML). ML is a rather large subdivision of artificial intelligence, studies methods for constructing algorithms capable of learning.

The investigated task belongs to the class of learning tasks by precedents (supervised learning). Each use case is a pair of "object – answer". It is necessary to find the functional dependence of the answers on the descriptions of the objects and build an algorithm that takes the description of the object at the input and gives the answer at the output. In the task of classifying a set of valid answers defined. They are called class labels. A class is the set of all objects with a given label value. To solve the problem of learning from precedents, the model of renewable dependence is fixed first of all. Then a quality functional is introduced, which value shows how adequately the model describes the observed data. The quality functional is usually defined as the average error of the answers given by the algorithm for all objects in the sample. The learning algorithm is looking for a set of model parameters in which the quality functional on a given training sample takes the optimal value [26].

The process of developing a specific ML model consists of stages:

- process of preparation (presentation) of data;
- algorithm design process;
- training process of the algorithm on the available data;
- algorithm validation process on test data

At the presentation stage, the rules for coding elements and forming data structures are determined [11]. The objects of learning are the vectors that are formed from the signs presented in numerical form. We will use the following features: the psychological type of the person tested; form of information; level of progress of a tested person. The sign “psychological type of a tested person” refers to a nominal type (signs with disordered states) and can take on the value of one of certain psychological personality types: visual, audial, kinesthetic, and digital. Put each type of signs in accordance with the numerical value: 1 – visual; 2 – audial; 3 – kinesthetic; 4 – digital. The sign “information presentation form” is also nominal, and in the numerical form it takes the following values: 1 – analytical; 2 – figurative; 3 – heuristic; 4 – audio. The sign “progress level” reflects the dynamics of the indicator of the quality of knowledge at the time of classification and is calculated as the growth rate of the assessment coefficient for each question in the test. Since the sign of the “progress level” has the ordinal type, it is necessary to match the actual value of the growth rate indicator with the ordered numbers, they can be compared with each other, but the distance between them is not defined. If the growth rate is negative, the sign “progress level” is taken as 1, with zero – 2, with positive – 3. As a class is the level of complexity of the next task. The specified feature is a target, that is, its value must be predicted, it is ordinal in type, and it takes on the values: 1 – easy level; 2 – medium; 3 – difficult.

The classification problem can be solved by many methods of ML, in particular, by the Bayes classifier, a decision tree, an algorithmic composition of decision trees, an artificial neural network, etc. [23]. The Bayes classifier is a probabilistic classifier, based on the Bayes theorem to determine the probability of the sample object belonging to one of the classes. To implement this model, it is necessary to assume independence of variables, what is a disadvantage of this method, because otherwise the probabilities of belonging to classes are not exact. The main advantage of the Bayes classifier is that to determine the parameters of the model, a small amount of data is required, as well as moderate use of the machine resource, high speed and simplicity [6]. A decision tree is a family of algorithms that are widely used in ML technologies. The structure of the tree is the “leaves” and “branches”. On the “branches” of the decision tree, the conditions on which the forecast depends depend on, in the “letter” the predicted value is recorded, and in other nodes – the conditions on which the tree forks. As a rule, binary decision trees are used. Among the advantages of the method is simplicity and reliability. The disadvantages include the fact that building an optimal tree is an NP-complete problem; decision trees are very easily retrained, which requires the use of additional resource-intensive measures to combat retraining.

Artificial neural networks are a mathematical model of the functioning of biological neural networks – networks of nerve cells of a living organism. As in the biological neural network, the main element of an artificial neural network is a neuron. Interconnected neurons form layers, the number of which can vary depending on the complexity of the neural network and the tasks it solves. The theoretical foundations of programming such neural networks are described in many papers [15; 20; 25]. The advantages of the method include high fault tolerance, orientation to parallel calculations, no need to formalize data. However, this type of method has drawbacks: the inability to reproduce the result, the complexity and empirical nature of the definition of network architecture [7].

To date, the most effective classifier is the machine learning method – gradient boosting, which belongs to the class of compositional methods. Boosting is a way to create compositions from decision trees, within which the basic algorithms are built sequentially, one after another, and each next algorithm is chosen in such a way as to correct the errors of an already constructed composition (6).

$$a_N(x) = \sum_{n=1}^{N} b_n(x), \quad (6)$$

where $b_n(x)$ – basic algorithms (decision trees) on the space of signs x.

To assess the quality of the algorithm, the loss function $L(y,z)$ is defined, where y is the true answer for each object from the sample x, z is the algorithm's prediction for the same object. The classification tasks use the logistic loss function (7).

$$L(y, z) = \log\big(1 + \exp(-yz)\big) \quad (7)$$

The first basic composition algorithm is based on the formula (8).

$$b_0(x) = \operatorname{argmax}_{y \in Y} \sum_{i=1}^{l} [y_i = y] \quad (8)$$

It returns the label of the most common sample class.

Learning the algorithm that is attached to the composition in each subsequent step is reduced to the task of learning from precedents. (9).

$$b_N(x) = \operatorname{argmin}_b \frac{1}{l} \sum_{i=1}^{l} (b(x_i) - s_i)^2, \quad (9)$$

where $\{(x_i, s_i)\}_{i=1}^{l}$ – the sample on which the training is performed, the vector s – the vector of shifts (10).

$$s = -\nabla F = \begin{pmatrix} -L'_z(y_1, a_{N-1}(x_1)), \\ \cdots \\ -L'_z(y_l, a_{N-1}(x_l)) \end{pmatrix}, \quad (10)$$

which should minimize the loss function (6) for each next algorithm.

At the stage of constructing the algorithm we will construct a classifier with the help of the freely distributed library of machine learning programming language Python – scikit-learn.

```
from sklearn.ensemble import GradientBoostingRegressor
```

```
boost = GradientBoostingRegressor(n_estimators=num,
    max_depth=d, random_state=42)
```

We perform the classifier training on training data, leaving 25% of the data as test data for model validation. In the cycle, we will consistently teach the algorithm on 100 decision trees, changing the depth of the trees, in order to empirically determine the optimal parameters of the model – the number of trees in the composition and their depth.

```
X_train, X_test, y_train, y_test = train_test_split(X, y,
    train_size=0.75, random_state = 42)
trees = np.array([5, 10, 15, 20, 30, 40, 50, 60, 70, 100])
depths = np.array([3, 5, 7, 10, 15, 22])
for d in depths:
    scores_train = []
    scores_test = []
    for num in trees:
        boost = GradientBoostingRegressor(
                n_estimators=num, max_depth=d,
                random_state=42).fit(X_train,y_train)
       scores_train.append(np.sqrt(mean_squared_error(
                y_train, boost.predict(X_train))))
       scores_test.append(np.sqrt(mean_squared_error(
                y_test, boost.predict(X_test)))
```

6 Software architecture and operation

6.1 Architecture

For the software being developed, client-server architecture was chosen, in which the client and server interact using the HTTP protocol. Data exchange within the framework of this protocol is performed according to the standard “request – response” scheme, which is the basis for transmitting data to the Internet [9]. HTTP encapsulates the entire process of serving web pages and provides the ability to specify a way to encode a message, so that the client and server can exchange binary data. Adaptive testing software has been developed taking into account the architectural style of REST, which was created on the basis of and together with HTTP. REST defines the restrictions on the use of HTTP, and describes a well-developed web application: reliable, properly working, and scalable, with a simple elegant design that can be easily changed [24].

In the process of developing web applications, the authors used the concept of separation of logical sections of code, which contributes to higher cohesion both in the initial development and in the constant support of any system. A clear distinction between client and server levels makes modular code sections easily manageable. In addition, the task was to clearly separate the data and display the markup of the site:

the data is not embedded in the page, but is delivered in a textual JSON data exchange format based on JavaScript. This is consistent with the modern concept of unobtrusive JavaScript, which separates the behavior, structure and presentation of the page.

Flexibility and code reuse is a logical result of good code organization. Flexibility exists at many stages of the application life cycle, when code sections can be developed in relative isolation (application programming interfaces (API), created clients of mobile devices, and new versions of program sections are released). The method of software development was chosen the scheme of separation of application data, user interface and control logic – MVC (Model – View – Controller). A key advantage of the MVC approach is that the components are loosely coupled. Each separate part of a web application running on the Django framework has a single key purpose and can be changed independently without affecting other parts. For example, a developer can change a URL for a specific part of a program without affecting the base implementation. A designer can change the HTML code of a page without touching the Python code that creates it. The database administrator can rename the database table and indicate changes in one place, and not search and replace multiple files [13].

6.2 Description of software operation

In Fig. 2 shows a diagram of the system activity and the basic logic of the software, starting with the user's authorization, to the implementation of the adaptive testing function with the result output. After registration, the user is offered a survey to determine the individual characteristics of the person. Based on these data, the system selects tasks with an emphasis on a certain form of user perception.

At the output, the system offers the user a personal selection of questions of varying complexity and presentation form. The test is considered to be passed successfully, in the case when the subject reaches a set number of points. If the user has exhausted the limit of the number of questions in the test, without gaining the required number of points, the test is considered uncollected. The system takes into account how much time is spent on passing the test, as well as the fluctuation factor between answer choices, if the subject chose different answer choices before the final decision. On the basis of preliminary data on the level of completed tasks, the optimal level of the next question is predicted for a particular subject.

7 Results

A database of test questions with different levels of complexity was developed for the discipline “Artificial Intelligence Systems”, which is taught for 7-8 semesters to bachelors in the specialty 121 – “Software Engineering”. The discipline consists of 5 informative modules and 19 thematic tests. Each test contains from 5 to 8 questions of two types: with one correct answer (one-alternative question) and several correct answers (multi-alternative).

Developed web-based software AdaptEd with an adaptive interface that allows applying testing in the discipline directly during lectures on students’ mobile devices. AdaptEd system allows creating online courses of disciplines due to the fact that the teacher can fill the system with lecture materials, additional materials, tasks for laboratory or practical work, test tasks, etc. Figure 3 show screenshots of AdaptEd at the stage of the entrance examinations of students and testing on the topic “Regression models” of the discipline “Artificial Intelligence Systems”. Each student can see the progress of testing success in the personal account of the system. At this time, organized pedagogical experiment, which should show the advantages of using adaptive testing to the standard one.

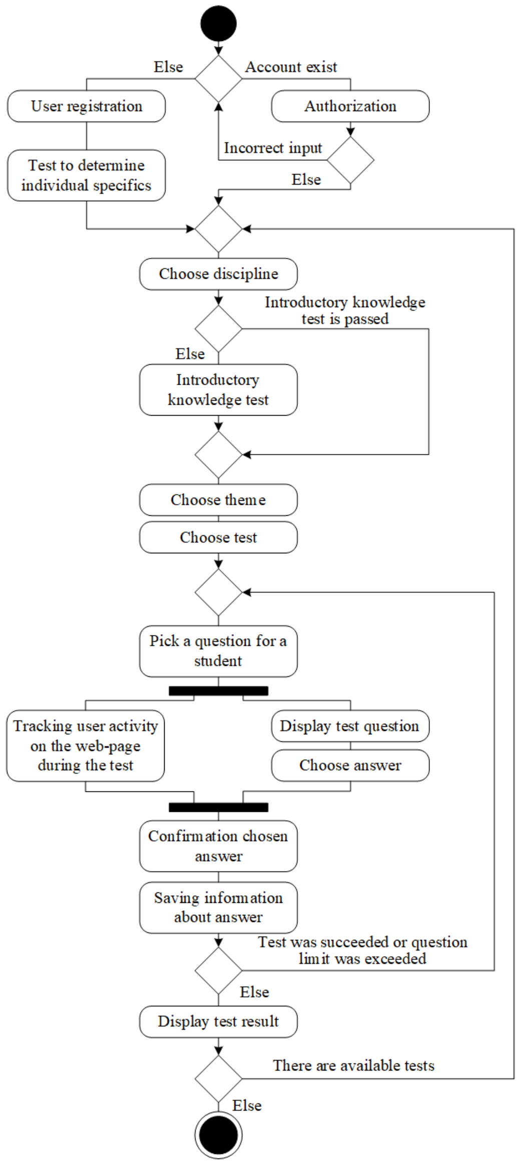
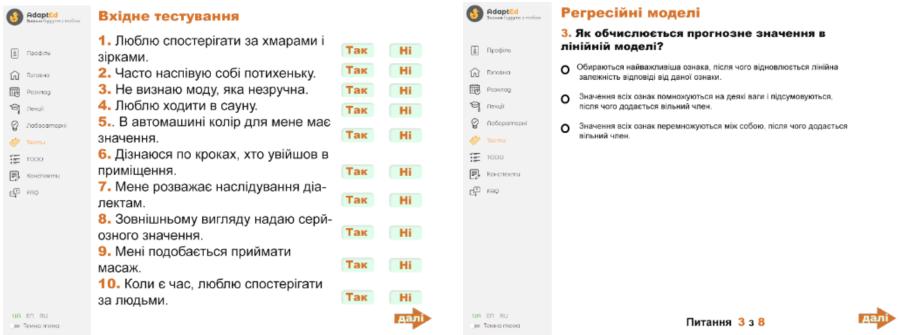


Fig. 2. System activity diagram



a) b)

Fig. 3. a) Input testing. Definition of psychological type; b) Test for the topic "Regression Models" for the discipline "Artificial Intelligence Systems"

8 Conclusion

In the process of studying the problem of developing web-based software for an adaptive testing system for bachelors, the current state of using computer adaptive testing methods was analyzed, the choice of methods for implementing a knowledge quality control system was justified, taking into account individual student characteristics, a universal system for adaptive knowledge control was developed, and a complex was developed-oriented test tasks of different levels of complexity for the current and module control in the discipline "Artificial Intelligence Systems" for bachelors in the specialty 121 – "Software Engineering".

The results of the study lead to the following conclusions:

1. Adaptive testing, based on the individual indicators of the subject, is an effective method of controlling the knowledge of students of higher educational establishment.
2. The implementation of adaptive knowledge testing should be carried out under conditions of systematic control, individualization by personality type and level of knowledge.
3. The software implementing the adaptive testing system should be in constant access for the subject, therefore the client-server interoperability model is implemented.
4. The database of test tasks must have enough of them for each level of complexity and be presented in a different form of information presentation.

The results of experimental use of the developed software for adaptive control of students' knowledge in the discipline "Artificial Intelligence Systems" showed that the testing system should be supplemented with tasks with a detailed answer. In addition, it is desirable to expand the base of tests with tasks with imaginative and heuristic forms of knowledge representation. These comments define ways to further improve the developed system of adaptive quality control of knowledge:

- study the methodology and develop a mechanism for assessing the detailed response using the methods of dormant-semantic analysis;
- implement a software module that will provide the opportunity for the subject to demonstrate tasks with multimedia content.

References

1. Amelkin, V.V.: Differential equations in Applications. Nauka, Moscow (1987)
2. Avanesov, V.S.: Difficulty of a Test and Testing Tasks. Management of School. 40 (1999)
3. Barbe, W.B., Milone M.N.: What we know about modality strengths. Educational Leadership. February, 378–380 (1981).
4. Bondarenko, M.F., Semenets, V.V., Belous, N.V., Kutsevich, I.V., Belous, I.A.: Assessment of Testing Tasks of Various Types and Their Complexity Determination. Artificial Intelligence. 4, 322–329. http://www.foibg.com/ibs_isc/ibs-12/ibs-12-p07.pdf (2009). Accessed 13 Feb 2019
5. Chelombitko, V.F., Mazhuga, M.O.: Use of Illustrative Materials for Creating Book and Electronic Multimedia Publications. Bionics of Intelligence. 1(86), 112–115 http://openarchive.nure.ua/bitstream/document/4852/1/Chelombitko_VF_112-115.pdf (2016). Accessed 10 Feb 2019
6. Domingos, P., Pazzani, M.: On the Optimality of the Simple Bayesian Classifier under Zero-One Loss. Machine Learning. **29**(2-3), 103–130 (1997). doi:10.1023/A:1007413511361
7. Flach, P.A.: Machine Learning: The Art and Science of Algorithms that Make Sense of Data. Cambridge University Press, Cambridge (2012)
8. Fleming, N.D., Mills C.: Not Another Inventory, Rather a Catalyst for Reflection. To Improve the Academy. **11**, 137–149 (1992)
9. Forcier, J., Bissex, P., Chun, W.: Python Web Development with Django. Pearson Education, Boston (2008)
10. Garipova, Z.F.: Determination od the type of education material perception by Primary Students. In: North of Russia: Development Strategies and Outlook, Surgut (2016)
11. GOST: Data Organization in Data processing Systems. Terms and Definitions. GOST 20886-85.
12. Hodovaniuk, T.L.: Idividyal Training at Higher School. NPU im. M. P. Drahomanova, Kyiv (2010)
13. Holovaty, A., Kaplan-Moss, J.: The Definitive Guide to Django. Apress, New York (2009)
14. Koltsov, Yu.V., Dobrovolskaya, N.Yu.: Neural Network Models in Adaptive Computer Learning. Educational Technology & Society. **5**(2), 213–216 (2002)
15. Luo, F.L., Unbehauen, R.: Applied Neural Networks for Signal Processing. Cambridge University Press, Cambridge (1998)
16. Malygin, A.A: Adaptive Testing of Students' Learning Results at Distant Learning. Abstract of Cand. Sci. (Pedagogy) Dissertation (2011)
17. Mazorchuk, M.S., Dobriak, V.S., Honcharova, K.A.: Model-Based Test Parameter Assessment. In: International Scientific and Practical Conference "Problems and Outlooks of IT-Industry Development", pp. 123–125 (2010)
18. Perceptual Learning Style Preference Questionnaire https://ru.scribd.com/document/258495684/Perceptual-Learning-Style-Preference-Questionnaire (1984). Accessed 30 Jan 2019

19. Petrova, M.Ye., Mintii, M.M., Semerikov, S.O., Volkova, N.P.: Development of adaptive educational software on the topic of "Fractional Numbers" for students in grade 5. In: Kiv, A.E., Semerikov, S.O., Soloviev, V.N., Striuk, A.M. (eds.) Proceedings of the 1st Student Workshop on Computer Science & Software Engineering (CS&SE@SW 2018), Kryvyi Rih, Ukraine, November 30, 2018. CEUR Workshop Proceedings. **2292**, 162–192. http://ceur-ws.org/Vol-2292/paper19.pdf (2018). Accessed 21 Mar 2019
20. Principe, J.C., Euliano, N.R., Lefebvre, W.C.: Neural and Adaptive Systems. Fundamentals Through Simulations. John Wiley & Sons, New York (2000)
21. Rash, G.: Probabilistic models for some intelligence and attainment tests. University of Chicago Press, Chicago (1993)
22. Registry of the Information Intake style by A.R. Gregos. http://psytests.org/cognitive/gregos.html. Accessed 28 Jan 2019
23. Richert, W., Coelho, L.P.: Building Machine Learning Systems with Python. Packt Publishing, Birmingham (2016)
24. Saternos, C.: Client-Server Web Apps with JavaScript and Java: Rich, Scalable, and RESTful. O'Reilly Media, Sebastopol (2014)
25. Semerikov, S.O., Teplytskyi, I.O., Yechkalo, Yu.V., Kiv, A.E.: Computer Simulation of Neural Networks Using Spreadsheets: The Dawn of the Age of Camelot. In: Kiv, A.E., Soloviev, V.N. (eds.) Proceedings of the 1st International Workshop on Augmented Reality in Education (AREdu 2018), Kryvyi Rih, Ukraine, October 2, 2018. CEUR Workshop Proceedings. **2257**, 122–147. http://ceur-ws.org/Vol-2257/paper14.pdf (2018). Accessed 21 Mar 2019
26. Shapovalova, N.N., Rybalchenko, O.H. Kuropyatnyk, D.I.: Comparative Analysis of Methods of Optimizing the Functional Machine Learning Models Quality. Bulletin of Kryvyi Rih National University. 46, 104–112 (2018)
27. Striuk, A.: Design of Educational Objects of Augmented Reality. Transactions. Georgian Technical University. Automated control systems. 2(26), 127–134. http://gtu.ge/Journals/mas/Referat/N26_conf_unesco_2018_2_26.pdf (2018). Accessed 29 Oct 2018
28. Sviridenko, O.M.: Use of the Specific Character of Representative Systems of Students for Successful Information Uptake. Bulletin of National Aviation University, Pedagogy, Psychology. 1 (2009)
29. Syrovatskyi, O.V., Semerikov, S.O., Modlo, Ye.O., Yechkalo, Yu.V., Zelinska, S.O.: Augmented reality software design for educational purposes. In: Kiv, A.E., Semerikov, S.O., Soloviev, V.N., Striuk, A.M. (eds.) Proceedings of the 1st Student Workshop on Computer Science & Software Engineering (CS&SE@SW 2018), Kryvyi Rih, Ukraine, November 30, 2018. CEUR Workshop Proceedings. 2292, 193–225. http://ceur-ws.org/Vol-2292/paper20.pdf (2018). Accessed 21 Mar 2019
30. Wulansari, Y.: The Use of Visual Auditory Kinesthetic (VAK) Learning Model to Improve Students' Reading Comprehension: Graduating Paper. State Institute for Islamic Studies, Salatiga (2016)
31. Yefremtsev, S.: Diagnostics of Dominant Perceptive Modality. In: Fetiskin, N.P., Kozlov, V.V., Manuilov, G.M.: Social and Psychological Diagnostics of development of a person and small groups, pp. 237–238. https://psycabi.net/testy/289-test-audial-vizual-kinestetik-diagnostika-dominiruyushchej-pertseptivnoj-modalnosti-s-efremtseva (2002). Accessed 25 Jan 2019
32. Zahrebelnyi, S.L., Brus, M.V.: Adaptive Testing as a Means of Students' Knowledge Control at Technical Higher Educational Institutions. DHMA Bulleting. 1(22), 155–162 (2017)