Я.К. Салал, С.М. Абдуллаев

МОНИТОРИНГ КАЧЕСТВА ОБРАЗОВАНИЯ И ВНЕДРЕНИЕ ИНДИВИДУАЛЬНОГО ОБУЧЕНИЯ: ДЕМОНСТРАЦИЯ ПОДХОДОВ И АЛГОРИТМОВ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ОБРАЗОВАТЕЛЬНЫХ ДАННЫХ

Как показывает анализ данных международных тестирований знаний учащихся, среднее и высшее образование многих странах переживает кризис, который в том числе вызван не адекватностью системы мониторинга качества традиционного и дистанционного образования. В свою очередь, развитие такой системы мониторинга требует разработки методов классификации и квантификации, необходимых для прогноза индивидуальных и коллективных успехов учащихся. В этой статье теоретически и экспериментально показано, что наиболее перспективным подходом, решающим одновременно обе задачи прогноза, является создание гетерогенных ансамблей, состоящих из нечетного числа различных базовых классификаторов, таких как деревья решений, простейшие нейронные сети, наивный байесовский классификатор и другие. Проведя обучение и тестирование 11 различных бинарных классификаторов на шести различных выборках образовательных данных, нами показано, что индивидуальный детерминированный прогноз таких ансамблей превосходит по точности прогнозы как отдельных базовых классификаторов, так и однородных ансамблей созданных по технологиям бэггинга и бустинга. Преимущество гетерогенных ансамблей созданных из трех, пяти и семи классификаторов становится определяющим, если учесть, что при естественном дисбалансе выборки образовательных данных, полезным прогнозом классификатора может считаться только такой прогноз, точность которого превосходит относительную частоту класса объектов имеющих наибольшую повторяемость в выборке данных. Главным преимуществом гетерогенного ансамбля является возможность трансформации детерминированного прогноза в вероятностный прогноз, когда вместо отнесения объекта к тому или иному классу, даются вероятности его принадлежности к отдельным классам. На основе этого нами предложен новый метод бинарной квантификации, когда индивидуальные вероятности принадлежности к каждому из классов объектов суммируются по отдельности, а полученные в итоге суммарные вероятности интерпретируются как относительные частоты объектов в выборке. В результате экспериментов показано, что такая бинарная ансамблевая квантификация по точности заметно превосходит традиционный метод «классифицируй и считай».

Прогноз индивидуальных и коллективных успехов учащихся; анализ образовательных данных; классификация и квантификация; несбалансированные выборки; гетерогенные ансамбли; детерминированный и вероятностный прогноз.

Ya.K. Salal, S.M. Abdullaev

MONITORING OF THE EDUCATION QUALITY AND IMPLEMENTING OF INDIVIDUAL LEARNING: DEMONSTRATION OF APPROACHES AND EDUCATIONAL DATA MINING ALGORITHMS

The quality monitoring system for traditional and distance education requires the development of machine learning classification and quantification techniques necessary to predict individual and collective student performance. This article theoretically and experimentally shows that the most promising approach that simultaneously solves both forecast tasks is to create heterogeneous ensembles consisting of an odd number of different base classifiers, such as decision trees, simple neural networks, naive Bayesian classifier and others. By training and testing 11 different binary classifiers on six different samples of educational data, we show that the individual determined forecast of such ensembles exceeds the accuracy of forecasts of both individual base classifiers and homogeneous ensembles created by bagging and busting technologies. The advantage of heterogeneous ensembles is decisive when we deal with the imbalance of sample characteristic of

educational data. In these cases, only the forecasts with accuracies exceeding the relative frequency of the class of objects dominating in the sample of data can be considered as useful forecasts. The main advantage of the heterogeneous ensemble is the ability to transform the deterministic forecast into a probabilistic forecast, when instead of referring the object to a particular class, the probability of its belonging to individual classes is given. On this basis, we have proposed a new method of binary quantification, where individual probabilities of belonging to each of the classes of objects are summed up separately, and the resulting total probabilities are interpreted as relative frequencies of objects in the sample. As a result of experiments, it is shown that such ensemble binary quantification is significantly superior to the traditional "classify and count" method.

Individual and collective student performance forecasts; educational data mining; classification and quantification; imbalanced datasets; heterogeneous ensembles; deterministic and probabilistic forecast.

Introduction. The main problem of the higher and secondary education system in the fast-changing world needs is an inability to rapid self-improvement of the centuries-old structures of the nature of the teacher-pupil relationship, including the knowledge transfer, individual and collective student performance evaluation. It is clear that the school management type must be reformed also.
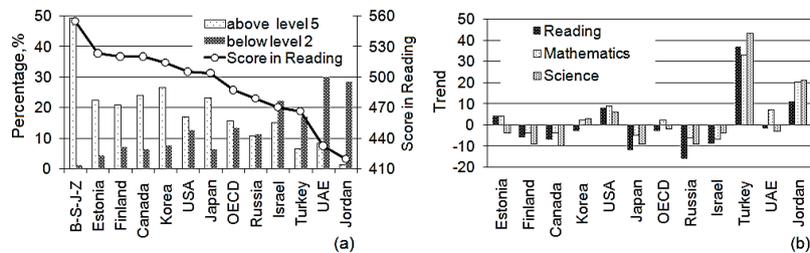

Fig. 1. Indicators of education systems in selected countries. (a) Assessment of reading quality (line) and percentage of 15 years old strains with low and high reading scores; b) Trends in mathematics, science and reading (2015-2018). Beijing, Shanghai, Jiangsu and Zhejiang cities of China) were abbreviated by B-S-J-Z

Success in this race for the future of nations depends on many socio-economic, ethnic, religious, historic, ecologic and any more factors. There are no simple dependencies here. This illustrates Figure 1a, which is based on recent worldwide testing of PISA [1]. The figure provides the score in reading of 15 years-old schoolchildren (line) in some selected countries and the corresponding percentage of students which demonstrate the ability to understand the text above level 5 (below) of complexity and student's light (dark) shaded bars. It is obvious that in the countries with the best grades (major cities of continental China, Estonia, Finland, Canada, and Korea) the number of students with a high (low) level of knowledge is above 20 % (below 10%). The largest economies of the U.S. and Japan have similar or slightly weaker results, but overall, they are better than the average value of the Organization for Economic Cooperation and Development (OECD). Russia, which has slightly weaker rates than the OECD average, has an almost equal percentage of such students of about 11 %. We have shown in [2] that this situation with secondary education leads to the fact that even into the most prestigious faculties of Russian higher schools come a very small number of applicants, demonstrating the real ability to the mathematics of bachelor level. For comparison, in the Middle East and Turkey, where overall educational achievements are lower than in Russia, there is a tendency to sharply increase the portion of weaker learners up to 20–30 %. Moreover, if in Israel schools it is adjacent to 15 % of strong learners, in Turkey, United Arab Emirates (UAE) and Jordan it is contradicted by a very small part of students with significant

knowledge. It should be said that the other Arab countries listed in [1] have much worse indicators than Jordan. Iraq was not part of the PISA testing, but alternatively, we test that in the relatively favorable environment of education of Al-Diwaniyah city the proportion of students with low knowledge of mathematics and computer science may be as high as 60 % [3]. We hypothesized that low knowledge is the result of the lack of a quality monitoring system.

Note that the economic indicators of countries are not decisive for the success of the education system: gross per capita income of top-five Estonia ($33,000) only slightly more than Turkey and Russia ($28,000), and less than Israel ($38,000) or UAE ($70,000). Moreover, the quality of national education (Fig. 1a) itself does not speak of its trends (Fig. 1b). As the figure from all selected countries, only Turkey, Jordan and the USA significantly improved their performance in reading, mathematics, and sciences during 2015-2018, and the rest of countries reduce their scores. The largest drop of ~10% of quality indicators is observed in Russia. We demonstrate that in such a situation, the degradation of traditional education can be partially slowed by a combination of individual and project-based learning [2], with the leading role of individual study. We have also made the case for more widespread application of computer systems for individual learning (CSIL) which can be seen as adaptive multi-agent systems with a detailed design of educator and student agents to develop an adequate pattern of their learning interactions, i.e. pedagogy. No one doubts the need for individual education and CSIL in our country [4, 5]. However, the procedure for introducing these forms of education raises many questions. Firstly, individual education approaches in higher schools are contradicted by observed trends to increase in the number of students in groups and to the increasing number of such groups leading by the same educator. These processes lead to the homogenization of pedagogical efforts. Sometimes this homogeneity is elevated to the rank of law and takes the form of electronic slavery when the professor fills in countless forms, surveys, and tables that have nothing to do with real work. In contrast, the CSIL requires a certain freedom of the teacher, and, in essence, changes the existing system of relationships at the university. On the other hand, the CSIL can take on a significant part of the load, freeing up the teacher's time for more complex pedagogical procedures. Secondly, it is obvious that the existing e-learning systems are not yet set up for a detailed analysis of the individual characteristics of the student and his relationship with a professional mentor. In many such systems, attention is paid only to the formal ability of the student to receive the course material. In such electronic training, even if it is well structured and meets the requirements of the standards [5], the human factor is excluded.

In this paper, we cover two related issues, while simultaneously predict the quality of education and implement of individual learning. To do this, we will first look at the system of monitoring the quality of education (section 1), determining that these tasks are solved in the subsystem of the prediction of individual and collective success of students Then in sections 2 and 3 we will show that the most appropriate methods for solving these problems are heterogeneous ensembles of supervised classifiers and quantifiers.

1. Education quality monitoring. The Education Quality Monitoring (MQE) general chart is shown in fig. 2. Based on this diagram, we discussed [3] the reasons for the low-quality of secondary education in the Republic of Iraq and we came to the conclusion that a three level of MQE system are needed. Work [3] discussed the background of the MQE of the Republic of Iraq at the national, local and regional levels. The main cause of weak human development indicators has been shown to be low-quality education due to the lack of national MQE in Iraq. We concluded that the national level of monitoring, the issue of forecasting is associated with a big data problem where the collection of initial information is a costly and time-consuming procedure, in addition to

that decision-making with feedback is a long-term process that may take years to complete. A regional monitoring system may be in a better situation. The need of this MQE at province level was demonstrated by the analysis of socio-demographic data and student performance at three preparatory schools in Al-Diwanyah city, Iraq. Despite the appropriate educational environment in capital of province, at least 4⁄5 of the number of students lost more than one year of study and 2⁄3 of them showed a low level of knowledge of two subjects of 160 students. Significant differences in the level of knowledge brought by different schools have also been found [3].
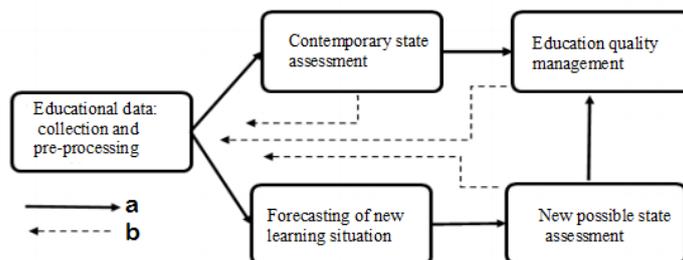

Fig. 2. Diagram of the educational quality monitoring and management with (a) direct and (b) feedback information flows

These observations have led us to conclude that the ongoing national-wide reform which introducing school-based management cannot be successful without the local MQE and their interaction with regional MQE. Particularly, we have demonstrated that decentralized national education systems with school-based management do not automatically lead to a rise of education quality. We conclude that at the local and regional level MQE, a key part of system is the prediction of a new situation with the quality of learning, determined by the individual and collective achievements of students, which in turn are educational data mining problems resolving by the classification and quantification methods [6–7]. Obviously, the mass-line courses implement [8] and recent ubiquitous implementation of distance learning as consequence of COVID-19 pandemic, also requires the application of these methods.

2. Supervised classification and sample imbalance. From the position of machine learning algorithms, the forecast of individual student performance can be produced by classification algorithms which trained on dataset containing previously collected data and then applied to new dataset. The task of the algorithm with the specified accuracy is to determine the class of the object. In our case, these are the student's future grades on the certain discipline. Without loss of generality, we will consider the problem of supervised classification into two classes: c+ is the class of student's which knowledge ratings are above certain threshold and c- is the class of students which educational performances are below this threshold.

Let's give 1) a training sample $D_1 = \{(\mathbf{x}_1, y_1), \dots (\mathbf{x}_{n1}, y_{N1})\}$, where the vector $x_i \in X_1$ consists of m attributes of the feature space of the student $X_1 \subset X$, (whereX more general feature space) and y_i denotes the corresponding class of this object, $y_i \in Y = \{c_+, c_-\}$ and 2) a classification algorithm that explores the training set and constructs a binary classifier model h, linking features and two classes of student performance (formula 1).

$$h: X_1 \to \{c_+, c_-\}. \tag{1}$$

It is assumed that this binary classifier applied to another independent and identically distributed sample of vectors $x_k \in X_2 \subset X$, will correctly indicate its individual classes $\hat{y}_k = c_+$ or $\hat{y}_k = c_-$. As a result of this, we obtain a forecast sample

115

$\widehat{D}_2 = \{(\mathbf{x}_1, \hat{y}_1), \dots (\mathbf{x}_N, \hat{y}_N)\}$. A simple evaluation of the effectiveness of the classifier is obtained by comparing the class $\hat{y}_k$ predicted by the classifier for the object x_k with a known value y_k from test sample $D_2 = \{(\mathbf{x}_1, y_1), \dots (\mathbf{x}_N, y_N)\}$ and calculating the relative sum of correct forecasts CF. This rating is called the overall accuracy A of the classifier. (formula 2)

$$A = \frac{\text{Correct forecasts}}{\text{Total forecasts}} = \frac{\sum_{k=1}^{k=N} CF(k)}{N}, \text{where } CF(k) = 1, \text{if}\hat{y}_k = y_k \text{or } CF(k) = 0, \text{if}\hat{y}_k \neq y_k. \quad (2)$$

The A is a fairly used as indicator of student performance forecasts, generally based on various algorithms of decision trees, Naive Bayes (NB), Multi-Layer Perceptron (MLP), k-Nearest Neighbors (kNN) and others algorithms [9]. Typically, the accuracies A about of 0.7-0.8 or more were reviewed in [9] for these classifiers. If one compares these forecast accuracies with value of $A_r \approx 0.5$ of the forecasts produced by the "random" classifier, which virtually tossing a coin to select the class of an individual student, then the accuracy in the range of 0.7-0.8 will seem to us quite acceptable. In addition, any forecast with an accuracy of more than accuracy of random forecast of 0.5 is considered useful. This would be a true conclusion when training and test samples contain approximately equal two numbers N+ and N-, where N+ is number of students belonged to class c+ and N- is number of students belonged to class c-. It is usually said that these datasets are balanced. However, detailed study of the student performance forecasts [7] showed that reviewed above high values of A were most often caused by imbalance of educational data samples. Indeed, our entire education systems are set up in such a way that the number of students N+ is much higher than N- Then, if the binary classifier mistakenly categorizes all new unknown students as a class of c+, the number of correct predictions in (2) will be equal to the N+, and the accuracy of this weak classifier A_w will be equal to relative frequency p_+:of dominant class c+:

$$A_w = \frac{\text{Correct forecasts}}{\text{Total forecasts}}\Big|_{\text{weak}} = \frac{N_+}{N_+ + N_-} \equiv p_+ = 1 - p_- \quad (3)$$

where recurrence of minority class c- in sample is denoted by p-. Thus, if an imaginary "optimistic" forecaster says that all unknown students will belong to the dominant class, the justification for his predictions will be more than random Ar=0.5! On the other hand, the p+ vs. p- imbalance results that binary classifier will be trained on a larger amount N+ of data corresponding to features of dominant class, and, as consequence, on the strongly unbalanced samples p+>>p-, the minority class of academic performance is virtually ignored by forecasting procedure. These effects can be seen from Table 1 where are demonstrated the results of forecast of 10 different types of classifiers trained by using Weka 3.8 [10] on the samples of 100 students derived from 6 experimental unbalanced dataset which, as we are sure [7, 11–15], have reasonable pre-processing quality.

Table 1

Accuracies (in %) of forecast of base and ensemble classifiers* trained and tested in the same conditions

#row	# Column Source of data	1 P+	2 NB	3 MLP	4 SVM	5 KNN	6 J48	7 AB1	8 RF	9 H₃	10 H₅	11 H₇
1	Diwaniyah, Iraq [7]	62	83	88	81	81	97	93	92	94	90	99
2	Assam, India [11]	66	61	91	74	70	100	98	100	98	88	100
3	Kashmir, India [12]	66	90	100	92	89	86	88	87	96	93	94
4	Nasiriah, Iraq [13]	72	75	73	80	77	71	67	71	76	76	78
5	Porto, Portugal [14]	77	86	79	82	75	87	86	85	86	87	87
6	Wasit, Iraq [15]	90	95	88	89	85	88	93	91	97	89	91
	Average	72,2	81,7	86,5	83,0	79,5	88,2	87,5	87,7	91,2	87,2	91,5

Comparing the frequency of the dominant p+ class in the test sample (column 1) with the accuracy of the NB, MLP, Support-Vector Machine (SVM), decision tree J48 (columns 2-6), you can see for each unbalanced sample of data, there is a classifier showing accuracy (crossed out) below Aw=p+ base line limit. Significantly, that in the last sample (row#6) with strong imbalance ratio of 9:1, all accuracies of classifiers were above 85%, but only one NB give useful forecast.

Two main types of solutions have been proposed to resolve the problem of training set imbalance: i) by artificially balancing using the re-sampling procedures [16] or ii) by so-called boosting procedure: sequential training of ensemble of classifier., where each new supplemental classifier, usually decision tree, is trained on the residual sample contained only the instances (students) which incorrectly classified by the prior ensemble member [17]. The classic type of such ensemble presented in the [13] is AdaBoostM1 (AB1, column #7). As you can see, with more accuracy in forecasts, this method also does not always satisfy the condition of the A>Aw. Another type of ensemble applicable to overcome imbalance problem is random forest RF which create a number of decision tree classifiers on various sub-samples of the dataset and uses averaging to improve the accuracy (column #8). Remarkably, that both the usual J48 and ensembles AB1 and RF, simultaneously showed poor results on the same dataset (row#4).

3. Classification and quantification by heterogeneous ensembles. The ensemble approaches of AdaBoost and Random Forest mentioned above are inherently homogeneous ensembles when new ensemble members are created by using the same technology. As it is illustrated by the above example of the homogeneous ensembles are not totally winning versions of the forecasting technology. Here we show a few basic advantages of the forecast with the help of heterogeneous ensembles.

3.1. Deterministic forecasts. First, the use of different basic algorithms brings us closer to the real-world situation, when the examination boards invite experts with experience in various fields. And as in real-world case, the technology for obtaining ensemble forecast results is quite simple and robust. Following the work [7], we will replace the labels "c+" and "c-" of binary classification by integers +1 and -1, and will consider the forecast of the binary classifier of model h_t for vector $\mathbf{x_i}$. It is clear that which will take two values $h_t(\mathbf{x_i})=y_t \in \{-1,1\}$. Let's try to create an ensemble $H_T(\mathbf{x})$ contained the odd number T of such binary classifiers $h_t(\mathbf{x_i})$. Let's assume, also, for simplicity that each of ensemble member gives equally important forecast for each of students in dataset. In this case, as is demonstrated by expression 3, the result of forecast $S(\mathbf{x})$ will be equal to the sign of the algebraic sum of forecasts of ensemble members $h_t(x_i)$:

$$S(\mathbf{x_i}) = \text{sign}\ [H_T(\mathbf{x_i})] = \text{sign}\ (\textstyle\sum_{t=1}^{T} h_t(\mathbf{x_i})), \quad \text{where } T = 2n + 1 \text{ and } [H_T(\mathbf{x_i})] \in \{\pm 1, \cdots, \pm T\}. \quad (4)$$

As can be seen this simple voting ensemble gives an erroneous forecast only is case when n+1 independent ensemble member gives non-correct forecast. Obviously, this deterministic forecast is no different in form from the predictions of the data of basic algorithms and homogeneous ensembles in Table 1. Therefore, it is easy for us to illustrate the advantage of a heterogeneous ensemble H_3 which contain as members J48, MLP, and NB classifiers. This ensemble as seen from column #9 gave predictions with accuracy above Aw on all samples. If we unify by (4) the individual results of all five basic classifiers J48, MLP, NB, SVM, and kNN to the ensemble H_5, we will see that the quality of forecasts of H_5 will deteriorate slightly due to the weak forecasts of individual members on the sample #6. Obviously, if you add the results of two homogeneous ensembles AB1 and RF to the previous set, the resulting meta-ensemble H_7 will once again have a high performance. However, for the sample #6, the accuracy of 91 % was obtained when H_7 determined 99 students to the class of 1, and only one to the class -1. Note that exactly this situation was occurred earlier with RF forecasts.

Thus, the obvious advantage of the presented technology of heterogeneous ensemble in that it gives higher and more stable results of the forecast, allows easily to incorporate new ensemble member: any classifiers or other ensembles. It is obvious that you can consciously select sets of ensemble members, for example, conducting a preliminary ranking of the classifiers on the stability of forecasts [7].

3.2. Probabilistic forecasts. It is obvious that the student's future grades on a particular subject or overall results depend on a set of factors. Two main ones can be identified: the student's cognitive abilities and social motivation. Obviously, there are different combinations of these factors (see Table. 2), and if on the basis of some specially questionnaires to allocate their strength (e.g. low, medium, and high) the task of binary classification would come down to the forecast of the future state of relatively small group of students marked by "?". Cognitive abilities should be considered as combination of cognitive functions which are not homogeneous in relation to the subject being studied; the presence of most cognitive functions is highly dependent on the age of the individual [18]. Thus, the peak of the basic cognitive functions contributing to programming falls on 17-18 years with a possible range of 15 to 21 years [19].

Table 2

Possible student performance due to combinations of student cognitive abilities with contemporary motivation strength

Abilities/Motivation	low	medium	high
low	2	2?	3
medium	2?	3	4
high	3	4	5

The above reasoning explains to some extent the following assumptions about the probabilistic forecast. We hypothesized that the basic student dataset could contain implicit information about the Abilities/Motivation combinations presented in table 2. The simplest example of this combination is that: a 19-year-old student begins to study professional programming, despite all the other positive props and strong motivation, his age signals that the cognitive functions required to the programmer are most likely to be in the stage of deactivating. Obviously, most of experts, and especially mathematical algorithms, will not notice the last fact and predict the cloudless future of this student. But it is possible that someone from the ensemble of experts (classifiers) will notice this and other alarming facts indicating below average abilities (see table 2) and will classify the student as a group requiring special attention. Unfortunately, in the process of voting by a simple majority (formula 4), this correct prediction of one of the ensemble members will be ignored. Obviously, the association between a mathematical classifier and an expert may seem tendentious. However, classification algorithms extract useful information from the feature space in different ways, and then there is a chance that one of the ensemble members will pay attention to suspicious combinations of features. When adopting such a hypothesis, for a more detailed forecast it would be necessary to take into account the opinion of all members of the ensemble and move on to the probabilistic forecast.

Here's how to do it: suppose that each of the **xi** students has individual natural probabilities to be categorized as class +1 and -1, respectively, $P_{+1}(\mathbf{x_i})$ and $P_{-1}(\mathbf{x_i})$. From formula 3, you can see that the deterministic forecast of the ensemble member $h_t(\mathbf{x_i}) = \pm 1$ can be converted into a probability forecast where each of the T individual classifiers gives its share of the deposit 1/T to one or the opposite probability. So, ensemble can generate two values of probability $P_{+1}(\mathbf{x_i})$ and $P_{-1}(\mathbf{x_i})$, for every student. Let's express them as a percentage, and note that the difference of these probabilities is equal to the algebraic sum of deterministic predictions divided by the T number of members of the ensemble (formula 5):

$$P_{+1}(\mathbf{x}_i) - P_{-1}(\mathbf{x}_i) = \frac{H_T(\mathbf{x}_i)}{T} \times 100\% \equiv \Delta P(\mathbf{x}_i) \qquad (5)$$

where designation $\Delta P(\mathbf{x}_i)$ is introduced. Given that the sum of individual probabilities is 100%, probabilities are easy to calculate by formula 6:

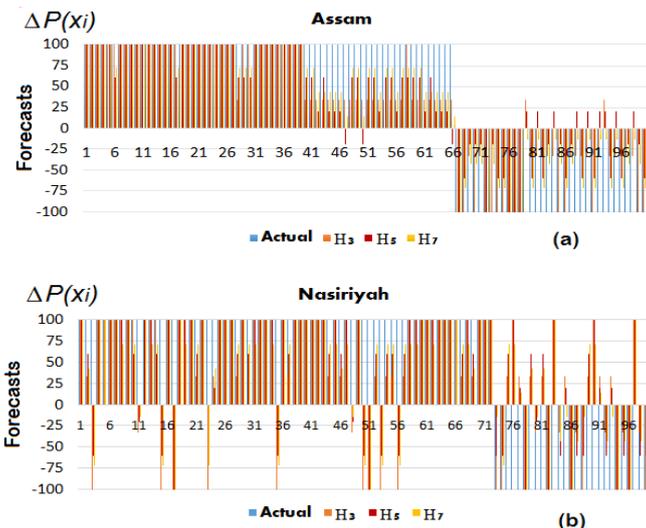$$P_{\pm 1} = 50\% \pm \frac{\Delta P(\mathbf{x}_i)}{2}. \qquad (6)$$

Fig. 3. Probabilistic forecasts $\Delta P(x)$ produced by ensembles of three H_3, five H_5 and seven classifiers H_7 using data of a) sample#2; and of b) sample#4

The two examples of predicted probability differences $\Delta P(\mathbf{x}_i)$ for the two classes of students and their real successes are presented on fig. 3. The positive and negative values of the forecasting difference $\Delta P(\mathbf{x}_i)$ and real difference ± 100 % mean that the student is classified as a member of group with satisfactory and non-satisfactory achievements. As we can see from the visualizations of probability forecasts, there are two different situations. In (fig. 3,a) is demonstrated the ensemble forecasts for sample #2 when the differences $\Delta P(\mathbf{x}_i)$ of erroneous forecasts of class +1 are rare and relatively small whereas $\Delta P(\mathbf{x}_i)$ of failed forecast of class -1 is more frequent and their values are relatively large. Although in this sample, H3, H5 and H7 made fewer errors (see table 1) than in other cases, however, these two patterns with low (high) frequency of errors of class +1 (-1) and small (large) values of differences $\Delta P(\mathbf{x}_i)$ are typical for forecasts on other samples #1. #3 and #5. The case of sample # 4 (fig.3, b) where values of $\Delta P(\mathbf{x}_i)$ of erroneous ensemble forecasts are significant in both classes is atypical and from (fig. 3,b) and Table 1, it is clear that student evolution in this case is the less predictable.

3.3. Ensemble quantification. The supervised quantification is a relatively newer section of machine learning [20]. The goal of the trained quantifier is to predict the frequency of objects belonged some classes in an unseen test sample X_2, which is ideal for predicting the collective success of students. In the case of two classes, the binary quantifier q gives a prediction of class probabilities $\widehat{p_+}, \widehat{p_-}$

$$q: X_2 \rightarrow \{\widehat{p_+}, \widehat{p_-}\}. \qquad (7)$$

The task of binary quantification (7) was first formulated by Forman [21], who proposed to solve it using base classification algorithms: Classify and Count (C&C). Some of the features of ensemble quantification we have outlined in the works [6, 7].

Indeed, in the case of a deterministic of base or ensemble forecast, if one counting the sum N+ of forecasts of class +1 then it can obtain the estimation of class probability $\widehat{p_+}^{C\&c}$ and $\widehat{p_-}^{C\&c}$.(formula 8):

$$\widehat{p_+}^{C\&C} = \frac{N_+}{N}; \quad \widehat{p_-}^{C\&C} = 1 - \widehat{p_+}^{C\&C}. \tag{8}$$

Taking into account the previous discussion, formula 7 suggests that a more accurate C&C classifier will give both best accuracy and the best prediction of +1 class frequency $\widehat{p_+}^{C\&C}$. All things being equal, this means that ensemble quantification is more accurate (Table 1) than basic C&C quantification. However, the opposite sense is wrong: the best quantifier does not necessarily become the best classifier. Among other things, we've shown that ensemble quantifiers based on probability forecast (formulas 5 and 6) give more accurate distribution assessment than a deterministic forecast. The following assumption may be put in the foundation of binary quantifiers which use probabilistic output of heterogeneous ensemble. If each of the objects in the group has the probability of P_{+1} (formula 5) being in the state of +1, then a group of such objects will have a mean cumulative probability $\widehat{p_+}^{Q}$ to be in this state (formula 9):

$$\widehat{p_+}^{Q} = \frac{1}{N}\sum_1^N P_{+1}(x_i). \tag{9}$$

Forecast values $\widehat{p_+}^{Q}$ of ensemble quantifiers of Q3, Q5, Q7 calculated according to (formula 9) compared to the C&C forecasts of ensemble classifiers H3, H5, H7 for two cases of highly unbalanced samples are shown in fig. 4. As seen in 5 cases, the quantifiers gave better predictions of true repeatability than C&C classifiers, and only in one case the accuracy was equal.
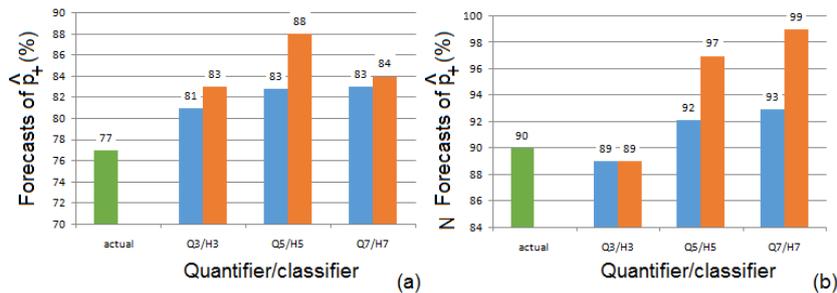

Fig. 4. Quantification: class frequency forecasts of quantifiers Q3-Q7 and ensemble classifiers H3-H7 comparing with actual frequency in test sample#5 (a) and test sample#6

(Fig. 4,a) also illustrates the assertion that "the best C&C quantifier does not necessarily have the best classification accuracy". So, the H3 classifier gave better predictions $\widehat{p_+}$, than H5 and H7 on the sample#5, but these classifiers had better accuracy of individual predictions (see table 1). On the contrary, in the case of (fig. 4,b) (sample#6), the best H3 classifier was the best C&C quantifier achieving quantification accuracy of Q3.

Conclusions. In introduction of this work, we have made an attempt to show that the decline in the quality of education in a number of countries is due to the lack of a really working hierarchical system monitoring the quality of education (MQE). The subsystem of the forecast of individual and collective performance of students is one of background elements of MQE. The major of goal of the presented study is the development of the methods of these forecasts using basic and ensemble classification algorithms. Based on the synthesis of previous works and analysis of experimental data from six samples, we show the following.

Individual forecast is advisable to carry out with the help of heterogeneous ensemble of supervised classifiers. It should be taken into account that these classifiers are trained and tested on unbalanced samples of educational data. In case of imbalanced dataset, only forecasts with accuracies above the base-line Aw= p+ the repetitiveness of the objects of the most represented class can be considered as useful predictions. Given this observation, the quality of the projections of heterogeneous ensembles is unsurpassed.

But, the main unsurpassed property of heterogeneous ensembles is the ability to obtain probabilistic individual forecasts, and using them as a basis, to create a new type of ensemble quantifier, which adequately solves the problem of prediction of collective performance of students.

More new research needs to be done to show the benefits of using heterogeneous ensembles in educational data mining.

REFERENCES

1. OECD (2019), PISA 2018 Results (Vol. I): What Students Know and Can Do, PISA, OECD Publishing, Paris. Available at: https://doi.org/10.1787/5f07c754-en.

2. *Abdullaev S.M., Lenskaya O.Yu., Salal Ya.K.* Computer Systems of Individual Instruction: Background and Perspectives, Bulletin of SUSU. Series "Education. Pedagogical science», 2018, Vol. 10, No. 4, pp. 64-71. DOI: 10.14529/ped180408. (in Russ.).

3. *Abdullaev S.M, Salal Y.K.* State and Prospects of Secondary Education System in the Iraqi Republic: to The Question of Monitoring the Quality of Education, Continuing education: Materials of the International forum: V International scientific and practical conference, ed. by I.A. Voloshinoy, I.O. Kotlyarovoy, 2019, Vol. 1, pp.14-26.

4. *Noskov M.V., D'yachuk P.P., Dobronets B.S., Vaynshteyn Yu.V., Kytmanov A.A., i dr.* The evolution of education in the context of informatization. Krasnoyarsk, 2019, Sibirskiy federal'nyy universitet, 212 p. (in Russ.).

5. *Silkina N.S., Sokolinskiy L.B.* Structural-Hierarchical Didactic Model of E-learning, *Bulletin of SUSU. Series: Computational mathematics and computer science,* 2019, Vol. 8 (4), pp. 56-83. DOI:10.14529/cmse190405. (in Russ.)

6. *Abdullaev S.M., Salal Y.K.* Ensemble Classification and Quntification: to Individual and Collective Student Performance Forecast, *Continuing education: international forum November 22, 2019: Materials of the fifth international scientific and practical conference " University of the XXI century in the system of continuing education»,* ed.by I.A. Voloshinoy, I.O. Kotlyarovoy, 2019, Vol.1, pp. 3-13. (in Russ.).

7. *Abdullaev S.M, Salal Y.K.* Economic deterministic ensemble classifiers with probabilistic output using for robust quantification: study of unbalanced educational datasets, *1st International Scientific and Practical Conference on Digital Economy (ISCDE 2019), Advances in Economics, Business and Management Research*, 2019, Vol. 105, pp. 658-665. DOI:10.2991/iscde-19.2019.128.

8. *Nesterov S.A., Smolina E.M.* The Assessment of the Results of a Massive Open Online Course Using Data Mining Methods, *Information and Telecommunication Technology in Education,* 2020, Vol. 13, No, 1, pp. 65-78. DOI: 10.18721/JCSTCS.13106.

9. *Mukesh Kumar, Salal Y.K.* Systematic Review of Predicting Student's Performance in Academics, *International Journal of Engineering and Advanced Technology (IJEAT)*, 2019, Vol. 8, No. 3, pp. 54-61.

10. WEKA The workbench for machine learning (available at https://www.cs.waikato.ac.nz/ml/weka/).

11. *Sadiq Hussain, Zahraa Fadhil Mushin, Yass Khudheir Salal, Paraskevi Theodorou, Fikriye Kutoglu, Hazarika G.C.* Prediction Model on Student Performance based on Internal Assessment using Deep Learning, 2019, Vol. 14(8), pp. 4-22. DOI: 10.3991/ijet.v14i08.10001.

12. *Mudasir Ashraf Bhat, Salal Y.K., Abdullaev S.M.* Educational data mining using base (individual) and ensemble learning approaches to predict the performance of students. MIND 2019: 1st International Conference on Machine Learning, Image Processing, Network Security and Data Sciences. Submitted for publication.

13. *Salal Y.K., Abdullaev S.M.* Optimization of Classifiers Ensemble Construction: Case Study of Educational Data Mining, *Bulletin of the South Ural State University. Ser. Computer Technologies, Automatic Control, Radio Electronics*, 2019, Vol. 19, No. 4, pp. 139-143. DOI: 10.14529/ctcr190414

14. *Salal Y.K., Abdullaev S.M., Mukesh Kumar*. Educational Data Mining: Student Performance Prediction in Academic, *Inter. Journal of Engineering and Advanced Technology (IJEAT),* April 2019, Vol. 8 (4) ,pp. 54-59.

15. *Salal Y.K., Abdullaev S.M.* Educational data mining using base and ensemble Learning approaches to predict student's performance. *Informatizaciya-i-Svyaz*, 2019, No. 5, pp.140-143.

16. *Chawla N.V.* Data mining for imbalanced datasets: An overview, Data Mining and Knowledge Discovery Handbook, Springer. Boston, MA, 2010, pp. 875-886. DOI: 10.1007/978-0-387-09823-4_455.

17. *Galar, M., Fernandez, A., Barrenechea, E., Bustince, H. and Herrera, F.* X A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches, *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews),* 2011, Vol. 42 (4), pp. 463-484. DOI: 10.1109/TSMCC.2011.2161285.

18. *Hartshorne J.K. and Germine L.T.* When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life span, *Psychological science*, 2015, Vol. 26 (4), pp. 433-443. DOI:10.1177/0956797614567339.

19. *Abdullaev S.M., Lenskaya O.Yu., Salal Ya.K.* Computer Systems of Individual Instruction: Features of Student Model, *Proceedings of the IV international scientific and practical conference, October 11-12, 2018, Chelyabinsk*, pp. 7-14. (in Russ.).

20. *González P., Castaño A., Chawla N.V., Coz J.J.D.* A review on quantification learning, *ACM Computing Surveys (CSUR)*, 2017, Vol. 50, No. 5, pp. 1-40. DOI: 10.1145/311780.

21. *Forman G.* Quantifying counts and costs via classification, *Data Mining and Knowledge Discov*, 2008, Vol. 17, No. 2, pp. 164-206. DOI: 10.1007/s10618-008-0097-y.

Статью рекомендовала к опубликованию к.т.н. О.Ю. Ленская.

Салал Ясс Кхудейр – Южно-Уральский государственный университет (национальный исследовательский университет); e-mail: Yasskhudheirsalal@gmail.com; г. Челябинск, просп. Ленина, 76; кафедра системного программирования; аспирант.

Абдуллаев Санжар Муталович – e-mail: abdullaevsm@susu.ru; кафедра системного программирования; д. геогр. н.; профессор.

Salal Yass Khudheir – South Ural State University; e-mail: Yasskhudheirsalal@gmail.com; 76, Lenin prospect, Chelyabinsk, 454080, Russia; the department of system programming; postgraduate student.

Abdullaev Sanjar Mutalovich – e-mail: abdullaevsm@susu.ru; the department of system programming; dr. of geogr. sc.; professor.

П.А. Чуб, Д.Н. Цветкова, Н.В. Болдырихин, Д.А. Короченцев

ОЦЕНКА ЗАЩИЩЕННОСТИ ПОМЕЩЕНИЯ ОТ УТЕЧЕК РЕЧЕВОЙ ИНФОРМАЦИИ В УСЛОВИЯХ ВОЗДЕЙСТВИЯ ШУМОВ

Рассматриваются особенности применения методик аттестации защищенных помещений. Такие методики разрабатываются и регламентируются Федеральной службой по техническому и экспортному контролю. Необходимость аттестации связана с наличием технических каналов утечки информации, по которым злоумышленником могут быть получены сведения, составляющие коммерческую, государственную или иную тайну. Наличие технических каналов утечки информации обусловлено физическими процессами, связанными с особенностями распространения акустических, электромагнитных и оптических