The Development of the Staking-Ensemble of Methods for Analyzing Academic Data

Indira Uvaliyeva, Zhenisgul Rakhmetullina, Olga Baklanova

East Kazakhstan Technical University, Department of Engineering mathematics, A. K. Protazanov Str. 69, 070004, Ust-Kamenogorsk, Kazakhsta
e-mail: {iuvalieva, zhrakhmetullina, obaklanova}@ektu.kz

György Györök

Óbuda University, Alba Regia Technical Faculty, Budai út 45, H-8000 Székesfehérvár, Hungary, e-mail: gyorok.gyorgy@amk.uni-obuda.hu

Abstract: Within the framework of this article, the Staking-ensemble of methods for analyzing academic statistics is proposed, which makes it possible to increase the effectiveness of implementing academic monitoring tasks. The relevance of the topic is the need to develop a distributed information and analytical system that integrates information resources and general principles of models and methods of monitoring the infrastructures of academic facilities based on education statistics. Education statistics is a system of indicators characterizing quantitative and qualitative changes taking place in the field of education which makes it possible to obtain information for each level of education on the number of educational institutions, the contingent of students, characteristics of the internal efficiency of the learning process, data on admission to educational institutions, graduation of specialists, quantitative and qualitative characteristics of the teaching staff, the state of the material and technical base of educational institutions. The object of the study is the system of formation of statistical data in education. The subject of the study are the approaches of combining intelligent methods of data analysis. The idea of the work is the use of modern methods of data processing in the implementation of academic monitoring in order to successfully solve the tasks of the state program for the development of education. The goal of the study is to develop an algorithm for Staking-ensemble methods for analyzing academic statistics to improve the effectiveness of academic monitoring tasks. The scientific novelty of the research is the staking-ensemble for analyzing education statistics, which aggregated the following 3 types of intellectual models: the Bayes algorithm, the decision tree algorithm, and the neural network. The practical importance of the research results lies in the applicability of the proposed Staking-ensemble algorithm for solving the problems of information and analytical support of management decision-making when tracking the business processes of monitoring, controlling and forecasting the state of distributed academic objects at various levels of training, management and functioning.

Keywords: academic statistics; analysis of indicators; Staking-ensemble algorithm; combination of intellectual methods

Introduction

Great interest in education in society and the need to make effective management decisions aimed at improving the quality of education require the use of prompt and reliable information on the state and development trends of the entire education system, its analysis and adequate interpretation. To solve this problem, it is necessary to use the processing algorithms of the system of indicators characterizing the quantitative and qualitative changes taking place in the field of education. In this study, it is proposed to apply a combination of the following algorithms for intelligent data processing: Bayes algorithm, decision tree algorithm and neural network. Education statistics cational statistics is a system of indicators characterizing quantitative and qualitative changes taking place in the field of education which makes it possible to obtain information for each level of education on the number of educational institutions, the contingent of students, characteristics of the internal efficiency of the learning process, data on admission to educational institutions, graduation of specialists, quantitative and qualitative characteristics of the teaching staff, the state of the material and technical base of educational institutions [1-4]. These data reflect the state, general assessment, trends and dynamics of development, the effectiveness of the education system, which are used to determine the effectiveness of its functioning and forecasts regarding the prospects for the development of its individual components and phenomena, as well as the basis for developing the necessary management decisions. In order to improve the efficiency of decision-making in the field of education, this study proposes a comprehensive application of methods of intellectual processing of educational statistics data.

In recent years, highly specialized data mining packages have appeared. For such packages, orientation to a narrow range of practical problems is often characteristic, and their algorithmic basis is a model using a neural network, solving trees, limited search, etc. Such developments are substantially limited in practical use. First, the approaches inherent in them are not universal with respect to the dimensions of the tasks, the type, complexity and structure of the data, the magnitude of the noise, the inconsistency of the data, and so on. Secondly, created and "tuned" to solve certain problems, they can be completely useless for others. Finally, many tasks of interest to a practical user are usually broader than the possibilities of a separate approach [5].

At the same time, the variety of algorithms for extracting knowledge (Data Mining) suggests that there is no one universal method for solving all problems [6-8]. In addition, the application of various analysis and modeling tools to the

same dataset may have different objectives: either to construct a simplified, transparent, easily interpretable model to the detriment of accuracy, or to build a more accurate but also more complex, and therefore less interpretive model.

Thus, one of the urgent tasks of the modern approach to data processing, including forecasting, is to find a compromise between indicators such as accuracy, complexity and interpretability. Most researchers prefer obtaining more accurate results since for the end users the concept of transparency is subjective. The accuracy of the results depends on the quality of the data source, the subject area and the data analysis method used.

Obtaining more accurate results is more important since in recent years there has been a significant increase in the interest in the accuracy of Data Mining models based on intelligent teaching methods, by combining the efforts of several methods and creating ensembles of predictor models, which allows improving the quality of solving analytical problems [9-12]. The training of the ensemble of models is understood as the training of the final set of basic classifiers, the results of forecasting which are then combined, and the forecast of the aggregated classifier is formed.

When forming an ensemble of models, it is necessary to solve three main tasks [13-15]:

- to choose a basic model;

- to determine the approach to the use of a training set;

- to choose a method of combining the results.

Due to the fact that the ensemble is an aggregated model consisting of separate basic models, two alternatives are possible in its formation:

- the ensemble is made up of basic models of the same type, for example, only from decision trees, only from neural networks, etc.;

- the ensemble is made up of models of different types - decision trees, neural networks, regression models, etc.[16,17].

On the other hand, when constructing an ensemble a training set is used, for the use of which there are two approaches:

- re-selection, i.e., several subsamples are extracted from the initial training set, each of which is used for training one of the ensemble models;

- the use of one training set for the training of all ensemble models.

1 Mathematical Description of the Ensemble Algorithm

Let X be the set of object descriptions, Y is the set of answers, and there is an unknown "target dependence" - the mapping f: $X \rightarrow Y$ whose values are known only for the objects of the final training sample (1):

$$(X, Y) = \{(x_1, y_1), \ldots, (x_{|(X,Y)|}, y_{|(X,Y)|})\} \quad (1)$$

It is required to construct an algorithm a: $X \rightarrow Y$ approximating the target dependence on the entire set X.

Along with the sets X and Y, we introduce an auxiliary set R, called the space of estimates. We consider algorithms having the form of a superposition (2):

$$a(x)=C(b(x)) \quad (2)$$

the function b: $X \rightarrow R$ is called an algorithmic operator, and the function C: $R \rightarrow Y$ is a decisive rule. Many classification algorithms have exactly this structure: first estimates of the object's belonging to classes are calculated, then the decision rule translates these estimates into the class number. The value of the estimate b (x) can be the probability that the object belongs to the class, the distance from the object to the separating surface, the degree of confidence in the classification, and so on.

The composition T of algorithms $a_t(x)=C(b_t(x))$, $t=1,\ldots,T$ is the superposition of the algorithmic operators $b_t : X \rightarrow R$, correcting the operation F: $RT \rightarrow R$ and the decision rule C: $R \rightarrow Y$ (3):

$$a(x) = C(F(b_1(x),..b_T(x))), x \in X \quad (3)$$

The algorithms a_t, and sometimes the operators b_t, are called basic algorithms.

Superpositions of the form $F(b_1, \ldots b_T)$ are mappings from X to R, that is, again, by algorithmic operators.

We denote the set of base classifiers as α. Using the terms formulated above, the idea of stacking is the use as the algorithmic operators b_t of the base classifiers $A \in \alpha$, and as the corrective operation F of some meta-classifier M.

In turn, for combining the results, there are several best-known corrective operations:

Simple Voting:

$$b(x) = F(b_1(x),\ldots b_t(x)) = \frac{1}{T}\sum_{t=1}^{T} b_t(x) \quad (4)$$

Weighted Voting (5):

$$b(x) = F(b_1(x),\ldots b_t(x)) = \sum_{t=1}^{T} w_t \times b_t(x), \quad \sum_{t=1}^{T} w_t = 1, w_t \geq 0; \tag{5}$$

Mixture of Experts [18] (6):

$$b(x) = F(b_1(x),\ldots b_t(x)) = \sum_{t=1}^{T} w_t(x) \times b_t(x), \quad \sum_{t=1}^{T} w_t(x) = 1, \forall x \in X. \tag{6}$$

It is obvious that Simple Voting is only a special case of Weighted Voting, and Weighted Voting is a special case of a Mixture of Experts. It is also worth noting that, because of the large number of degrees of freedom, the training of a Mixture of Experts takes much longer than other compositional algorithms, so their practical applicability is justified only in the case of a priori information on the competence functions $g_t(x)$, which are most often determined by:

- the sign of $f(x)$ (7):

$$w_t(x;\alpha;\beta) = \sigma(\alpha f(x) + \beta), \quad \alpha, \beta \in R; \tag{7}$$

- direction $\alpha \in R^n$ (8):

$$w_t(x;\alpha;\beta) = \sigma(x^{\mathrm{T}}\alpha + \beta), \quad \alpha \in R^n, \beta \in R; \tag{8}$$

- distance to $\alpha \in R^n$ (9):

$$w_t(x;\alpha;\beta) = \exp(-\beta\|x - \alpha\|^2), \quad \alpha \in R^n, \beta \in R; \qquad \text{.....(9)}$$

more complex methods (by means of estimating the density of data distribution, etc.). In the above formulas of functions, the competencies of the sigmoid are equal (10) [19]:

$$\sigma(z) = \frac{1}{1+e^{-z}}. \tag{10}$$

The use of model ensembles to solve various analysis problems opens up wide opportunities for improving the efficiency of Data Mining models. Therefore, in the past few years, active research has been carried out in this area, resulting in a large number of different methods and algorithms for the formation of ensembles. Among them, the most widely used methods are bagging, boosting and stacking [20].

2 Mathematical Description of the Stacking Algorithm

The Stacking algorithm is not based on a mathematical model. His idea is to use as the base models different classification algorithms, trained on the same data. Then, the meta-classifier is trained on the initial data, supplemented by the results of the prediction of the basic algorithms. Sometimes the meta-classifier uses not the results of the prediction of the basic algorithms, but the estimates of the distribution parameters obtained by them, for example, the probabilities of each class.

The generalized scheme for implementation of the stacking algorithm is shown in Figure 1. The idea of Stacking is that the meta-algorithm learns to distinguish which of the basic algorithms should be "trusted" on which areas of the input data.

The Stacking algorithm is trained using cross-validation: the data are randomly divided into n times, T1 and T2, containing each time (for example, for n = 10) 90% and 10% of the data, respectively. Basic algorithms are trained on data from T1, and then applied to data from T2. The data from T2 is combined with the forecasts of the base classifiers and trains the meta-classifier. Since after n partitions all available data will end up in some one of T1 and in some of T2, the meta-classifier will be trained on full data [21].

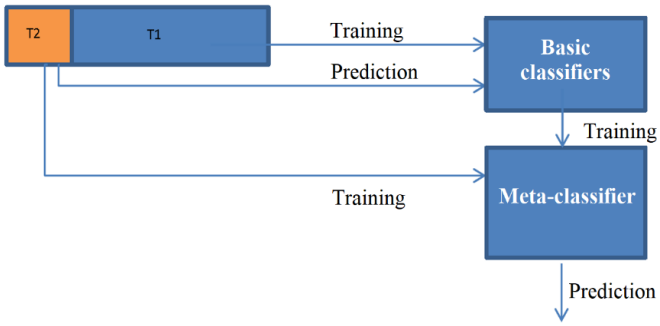

Figure 1
Generalized scheme for implementing the Stacking-ensemble algorithm

The statement of the classification problem is as follows.

Let X be the set of descriptions of objects, Y is a finite set of numbers (names, labels) of classes. There is an unknown "objective dependence" - the mapping f : $X \to Y$ whose values are known only on the objects of the final training sample $(X, Y) = \{(x1, y1), \ldots, (x|(X,Y)|, y|(X,Y)|)\}$.

It is required to construct an algorithm a a: $X \to Y$, capable of classifying an arbitrary object $x \in X$. The notation used in Table 1 is used.

Table 1
Notations and their descriptions

Notation	Description
(X0, Y0)	validation sample
A	basic classifier used to build meta tags
A.fit(X, Y)	function of training classifier A on (X, Y)
A.predict(X)	function predicting the target variable for X by the classifier A
M	some metaclassifier
MF(X, A)	meta-attribute obtained by classifier A for sample X
P	the final prediction of the stack for the validation sample
concatV(Xi , Xj)	concatenation operation Xi and Xj on columns
concatH(Xi , Xj)	concatenation operation Xi and Xj in rows

The idea of stacking is to train the meta-classifier M on (1) the original characteristics, the matrix X, and (2) on the predictions (meta-features) obtained with the help of the base classifiers. The meta-attributes obtained using the classifier A for the sample X will be denoted by MF(X, A).

Figure 2 shows the implementation of the stacking algorithm.

The simplest stacking algorithm for P sample prediction is based on dividing the training sample (X, Y) into 2 parts: (X1,Y1) and (X2,Y2).

```
A.fit(X1,Y1)
MF(X2;A) := A.predict(X2)
MF(X0;A) := A.predict(X0)
M.fit(concatV(X2;MF(X2;A)); Y2)
P := M.predict(concatV(X0;MF(X0;A)))
```

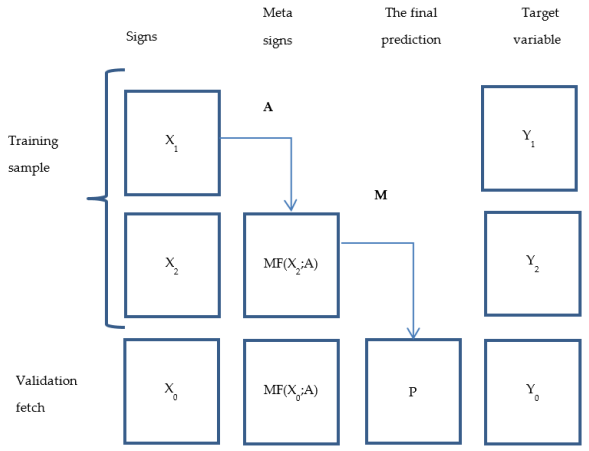

Figure 2
Schema implementation of the staking-ensemble algorithm

This algorithm is less effective because the meta-classifier is trained only on the second part of the sample P. To improve the efficiency of the prediction, you can apply this algorithm several times using different partitions and then average the predictions. The algorithm for this approach is as Fig. 3:

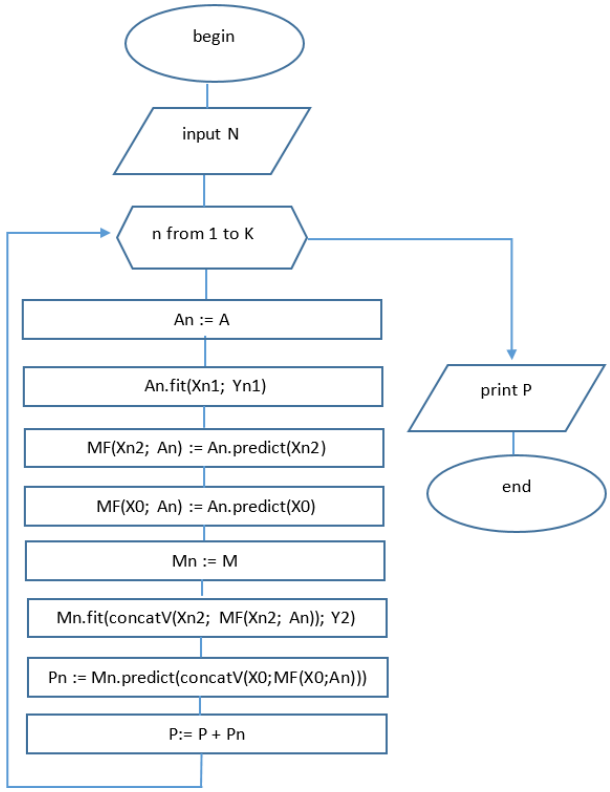

Figure 2
Algorithm of approach

Next, we represent the modification of the second algorithm, which allows us to add a new base classifier for learning set A. To do this, we need to get the meta-feature MF (X, A) for the entire training sample. The algorithm of this approach is as follows:

A1 := A
A1.fit(X1; Y1)
MF(X2;A1) := A1.predict(X2)
MF(X0;A1) := A1.predict(X0)
A2 := A
A2.fit(X2; Y2)
MF(X1;A2) := A2.predict(X1)
MF(X0;A2) := A2.predict(X0)

```
MF(X;A) := concatH(MF(X1;A2);MF(X2;A1))
MF(X0;A) := (MF(X0;A1) +MF(X0;A2)) / 2

M.fit(concatV(X;MF(X;A)); Y )
P := M.predict(concatV(X0;MF(X0;A)))
```

Thus, the modified Staking algorithm is presented, which allows the meta-classifier M to be used for learning the entire sample X. An accepted one is also a stacking representation in the form of a multilevel scheme where the features are denoted as "Level 0", meta-features obtained by training the base classifiers on the attributes as "Level 1", and so on.

3 Application of the Staking-Ensemble Algorithm for the Tasks of Academic Statistics

To construct an ensemble of models on the data of academic statistics, the Stacking algorithm was used, which aggregated the following 3 types of intellectual models: the Bayes algorithm (BC), the decision tree algorithm (DT), and the neural network (NN). The training was done on a single data set. The meta-level data used to train the meta-model will be the results of predicting models for each of the fields. To be able to be used as a meta-model of the BC, we include in the meta-level data of the class field a number for the BC, DT, and NN. Thus, the stacking-ensemble algorithm tries to train each classifier using the meta learning algorithm, which allows to find the best combination of outputs of base models. The structural scheme of the ensemble on the basis of stacking is presented in Figure 3.

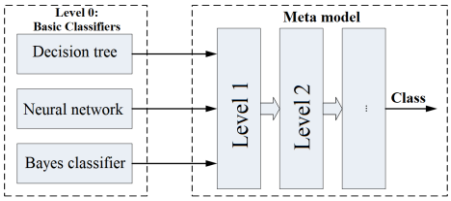

Figure 3
Block diagram of the ensemble on the basis of stacking

The base models form Level 0. At the input of the meta-model, also called the Level 1 model, the results are output from the outputs of the base models. The Level 1 example has as many attributes as there are Level 0 models, and the attribute values themselves are model outputs of the zero level. Then, based on the results obtained by the model of Level 1, a Level 2 model is constructed, etc., until some of the conditions for stopping the training are fulfilled.

Figure 4 shows the scheme of the algorithm for the ensemble of BC, DT, and NN models based on stacking.

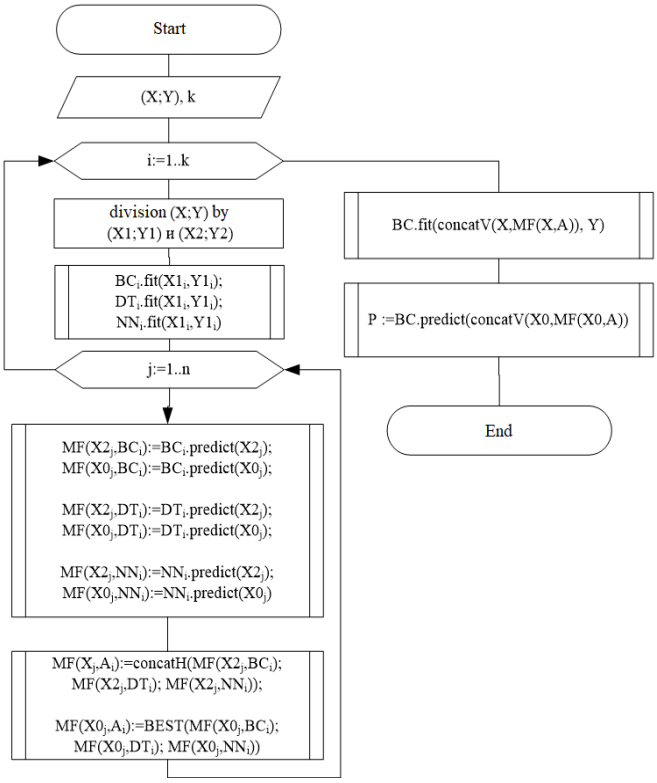

Figure 4
Diagram of an algorithm for the ensemble of BC, DR, and NC models based on stacking

The most common way to reduce the distortions caused by unsuccessful sampling is to repeat the entire learning and testing process several times with different random samples. At each iteration, a certain amount of data (for example, 2/3) is randomly selected for training, and the remaining ones are used for testing. For this, it is necessary to determine the number of iterations of the partitioning of the original data. Next, you specify the vector into which the data for learning the meta-model is collected. For each iteration, the following steps must be taken:

Step 1: Divide the original data into 2 parts: the first part is intended for learning the models, the second is for testing the model forecast. The proportion of the initial data division is chosen proceeding from the fact that as all the initial data after all the iterations are either in the 1st or 2nd part. To accomplish this, several analysis structures have been made in Analysis Services (the version number is added in the name). The sampling variability is specified through the HoldoutSeed property. It is different for different options.

Step 2: We teach the three initial models on the first part of the data.

Step 3: For each data line of the second part, you need: implement a forecast for each model; choose the most accurate of the forecast values; write a row with data and a prediction into a vector. If the vector already has this data line, then a more accurate prediction from the existing and obtained from Step 3 is recorded for it.

Step 4. After implementing all the iterations, the algorithm for constructing the meta-model is chosen. The meta-model is trained on the basis of the data collected in the vector.

To estimate the accuracy of the forecast for all iterations, a forecast accuracy chart, a classification matrix, and a cross-checking were constructed.

In the forecast accuracy chart, the forecasts of all models are compared. In this diagram, you can set the accuracy display for forecasts in general or for forecasts of a certain value. It displays a graphical representation of the accuracy change provided by the mining model. The X-axis of the diagram represents the percentage of the verification dataset used to compare forecasts. The Y-axis of the diagram represents the percentage of predicted values.

The classification matrix is created by sorting all the options into categories: whether the predicted value corresponds to the real one and whether the predicted value was true or false. These categories are sometimes also referred to as a false positive result, a true positive result, a false negative result and a true negative result. Then, all the variants in each category are recalculated, and the quantities obtained are output as a matrix.

During cross-validation, mining structures are split into cross-sections, after which cyclic learning and model checking for each data section is performed. To split the data, several sections are indicated, and each section, in turn, plays the role of verification data, while the remaining data are used to train the new model. Then, for each model a set of standard accuracy indicators is formed. Comparing the indicators of the models created for each section, you can get a good idea of how accurate the mining model is for the entire data set.

4 Implementation of the Staking-Ensemble Algorithm

To implement the experimental study of the Staking-ensemble algorithm of intellectual methods of data analysis, the following 2 objectives of academic monitoring at the higher education level were identified: 1) analysis of academic statistics data for forecasting GPA transfer points; 2) analysis of the data of academic statistics for forecasting the employment of a graduate of a university.

The study was conducted on the basis of the data of academic statistics at D. Serikbayev EKTU for 2017 - 2020 on baccalaureate (the daytime form on the basis of secondary education).

With a credit academic system, GPA (Grade Point Average) is used as the indicator of student's progress. It is a weighted average assessment of the level of the student's academic achievements in the chosen specialty, which is used to transfer the student to subsequent academic courses. Based on the Model Rules for the ongoing monitoring of academic performance, intermediate and final certification of students in higher academic institutions, approved by the #125 order of the Ministry of Education and Science of the Republic of Kazakhstan from March the 18th, 2020 and the Rules for the organization of the academic process on credit technology of education approved by the #152 order of the Ministry of Education and Science of the Republic of Kazakhstan dated 20.04.2011, in D.Serikbaev EKTU, the following transfer values of GPA for bachelor students of the daytime form of education were established on the basis of secondary education in the context of courses: for transfer to the second course - 1.5; for transfer to the third rate - 1.67; for transfer to the fourth course - 2.0; for the transfer to the fifth course - 2.1.

In Kazakhstan, the problem of employing graduates becomes more important every year. Currently, the results of numerous studies show that the current higher education system of training specialists largely does not meet the needs of the labor market. The problem of finding a job in the specialty after graduation, the problem of the first job is exacerbated every year. Even more urgent are these issues for the graduate of the university in the conditions of high socio-economic uncertainty and risks arising during economic downturns and crises. According to the purpose of the activity regulations is the creation of an effective system to facilitate the employment of graduates of the D. Serikbayev EKTU in accordance with the qualification and their adaptation to the labor market.

Thus, 2 objectives of academic monitoring implemented on the basis of the Staking-ensemble algorithm were defined to predict the following indicators of academic statistics: 1) GPA obtainment rate; 2) indicator of employment.

The Staking-ensemble algorithm assumes that data is taken for training and is divided randomly into 2 parts. The first part is used to teach the original models. Part 2 is used to predict data based on the original models and based on this forecast a meta-model is built. This operation is performed several times - so that all the initial data in different breakdowns are in the 2nd part.

To implement this algorithm, several mining structures were created in Analysis Services (the variant number was added in the name). The sampling variability is specified through the HoldoutSeed property. It is different for different options.

For each variant, the probability of occurrence of an event for each model in the ensemble is calculated. For example, for the task of analyzing the data of a transfer point set – the probability that the student will not gain a transfer point,

for the problem with the employment of graduates – the probability that the student will not be employed. From these probabilities we choose the most suitable: maximum – if the student does not score a transfer point / is not employed, and minimal – if otherwise. Based on the probabilities for each test line, a vector is formed that will be used to train the meta-model. To view the received data for learning the meta-model in MS SQL Server Management Studio, one of the following procedures is called:

- CALL SPMining.SPMining.SPortalMiningProc.GenerateTableGPA()
- CALL SPMining.SPMining.SPortalMiningProc.GenerateTableEmployment()

The following algorithm is used to create a metamodel:

```
CREATE MINING MODEL [Staking - Transferable GPA score has been scored]
(
  [ID] LONG KEY,
  [Basic Education] TEXT DISCRETE,
  [Entrance score] LONG DISCRETE,
  [Group of Specialties] TEXT DISCRETE,
  [Type of Financing] TEXT DISCRETE,
  [Year of study] LONG DISCRETE,
  [Total GPA Score For The Previous Course] LONG DISCRETIZED,
  [Specialization] TEXT DISCRETE,
  [Form of Training] TEXT DISCRETE,
  [The Number Of Undeveloped Disciplines For The Previous Course] LONG DISCRETIZED,
  [Language of Instruction] TEXT DISCRETE,
  [Class] TEXT DISCRETE PREDICT
)
USING Microsoft_Naive_Bayes(MINIMUM_DEPENDENCY_PROBABILITY=0.01)
```

Next, this model needs to be trained. Classes are identified for training, which are allocated depending on the probability. Classes are specified by a string when learning a model:

<ClassName1> = <Probability1>; <ClassName2> = <Probability2>; ...

To build a meta-model, we will use the algorithms of BC, DT and NN. Figure 5 shows the subsystem of the data analysis of academic statistics based on the Staking-ensemble algorithm.

Using several models gives the following advantages:

- realize the prediction of the output indicator of academic statistics both in numerical form (NN) and in the form of classification (DT and BC), which determines the variety of possible methods of visualizing the results;

- consistency of the results for all models allows to make a conclusion about their reliability.

The results of the Staking-ensemble algorithm implemented in the data analysis subsystem are presented below.

The training based on the basic classifiers for analyzing the data of the first objective of academic monitoring showed the following results.

The window of the subsystem with learning outcomes of the base classifier based on the BC algorithm contains the following tabs for mapping the interaction between the predicted attributes and the input attributes for the options table: the dependency network; attribute profiles; attribute characteristics; comparison of attributes.

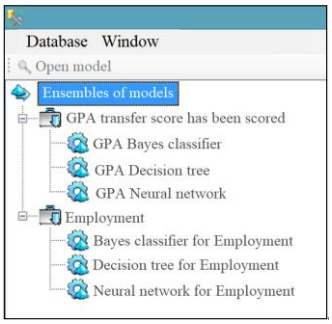

Figure 5

The main menu of the subsystem Staking-ensemble

Based on the data of academic statistics, the results of training the basic classifier on the basis of the BC algorithm for analyzing the data of the GPA set were obtained and are presented in the form of a network of dependencies in Figure 6.

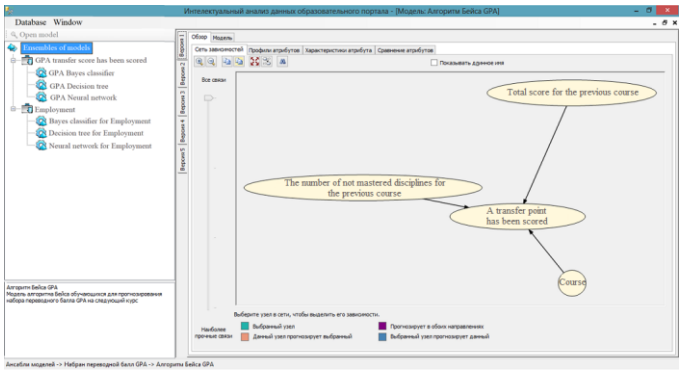

Figure 6

Subsystem window with learning base classifier results based on the BC algorithm for analysis of GPA data

The sub-system window of dependencies network tab displays the dependencies between the input indicators of the university's academic statistics and the predicted indicator that characterizes the GPA. The slider to the left of the viewer acts as a filter, tied to the strength of the dependencies. When you move the slider down, only the strongest ones are displayed. The symbols at the bottom of the subsystem window associate the color codes with the type of dependency on the

graph. As it can be seen in Figure 7, the indicator "Obtained GPA" is influenced by such indicators of academic statistics as the number of unfinished disciplines for a previous year, training, the total GPA score for the previous year of study. The basis of the second basic classifier was the decision tree algorithm, the results of which are shown in Figure 7a. This algorithm supports classification and regression, is used for predictive modeling of discrete and continuous attributes. As you can see from Figure 7b, the subsystem window contains the "Decision Tree" and "Dependency Network" tabs.

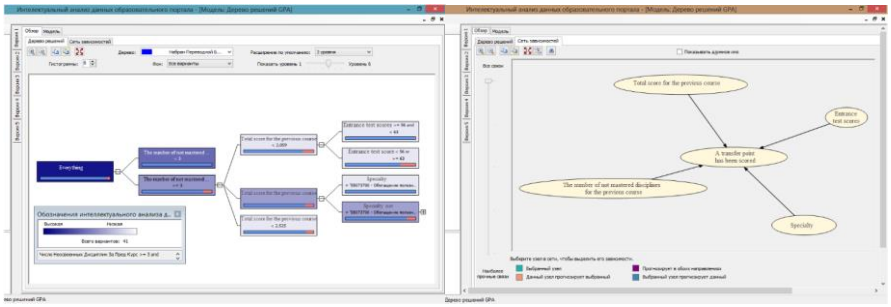

Figure 7
The window with the results of training the base classifier on the basis of the DR algorithm for analyzing the data of the GPA set

The results of training the basic classifier on the basis of the DT algorithm showed that the set of GPA has a dependence on the following indicators of the academic statistics of the university: the number of unfinished disciplines for the previous year, the GPA's total score for the previous year of study, specialty and the UNT score. Figure 8 shows the results of training the basic classifier based on the NN algorithm for analyzing data from the GPA set.

The meta-classifier Staking-ensemble was built on the basis of the BC algorithm. The results of the accuracy of the forecast implementation of Staking-ensemble based on the classification matrix.

The results of the meta-classifier training in the variation sample showed that the indicator of academic statistics "Obtained GPA" has a dependence on such indicators as the number of unfinished disciplines for the previous year, the GPA total score for the previous year, the Basic Education and the Course.

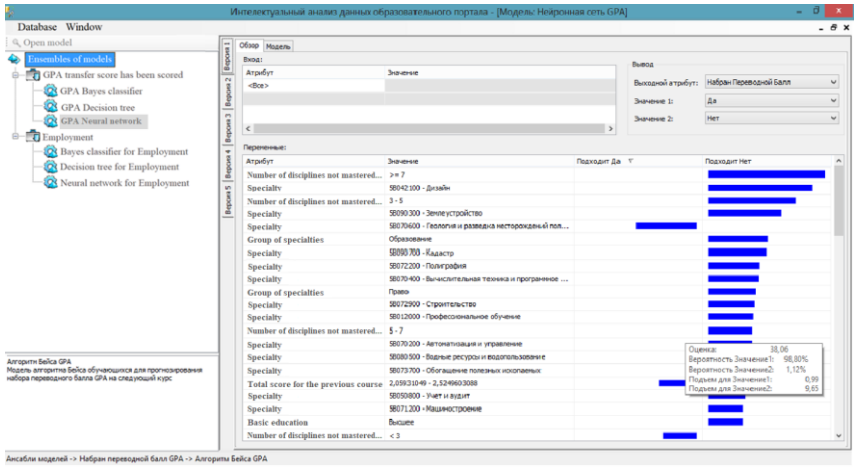

Figure 8

The window with the learning outcomes of the basic classifier based on the NN algorithm for analyzing the data of the GPA set

The windows with the results of the accuracy of the prediction of the implementation of the Staking-ensemble algorithm based on the accuracy chart is presented in Figure 9.

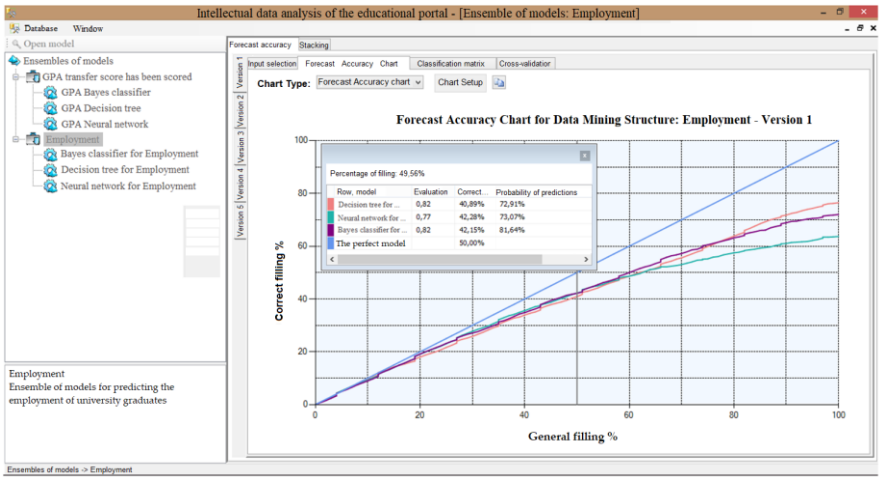

Figure 9

Staking-ensemble algorithm based on the accuracy chart

In this way the scientific novelty of the research is that for the first time the Staking-ensemble of methods for analyzing academic statistics is used to improve the effectiveness of academic monitoring processes. The practical importance of the research results lies in the applicability of the proposed Staking-ensemble

algorithm for solving the problems of information and analytical support of management decision-making when tracking the business processes of monitoring, controlling and forecasting the state of distributed academic objects at various levels of training, management and functioning.

Conclusion

To solve the analytical problems of academic statistics and to obtain more accurate results, a new approach is proposed that combines the several intellectual methods based on the creation of an ensemble of predictor models. To this end, a modified Staking-ensemble of intellectual methods for analyzing academic statistics has been developed, which allows the meta classifier to use the entire sample for learning and aggregates Bayesian algorithms, decision trees and neural networks.

To implement the experimental study of the Staking-ensemble algorithm of intelligent methods of data analysis, the indicators of academic statistics for forecasting the GPA transfer score and the employment of the graduate of the university were analyzed. The study was conducted on the basis of the data of academic statistics at D. Serikbayev EKTU for 2017-2020 on baccalaureate of the daytime form of education on the basis of secondary education.

References

[1] Cespón M. T., Lage J. M. D. Gamification, Online Learning and Motivation: A Quantitative and Qualitative Analysis in Higher Education //Contemporary Educational Technology. – 2022. – Т. 14. – №. 4. – С. ep381. https://doi.org/10.30935/cedtech/12297

[2] Larini M., Barthes A. Quantitative and statistical data in education: From data collection to data processing. – John Wiley & Sons, 2018

[3] Winkle-Wagner R., Lee-Johnson J., Gaskew A. (ed.). Critical theory and qualitative data analysis in education. – New York : Routledge, 2019

[4] Oharenko V., Kozachenko I. State Regulation Of The Education Sector Under The Conditions Of Decentralization //Baltic Journal of Economic Studies. – 2020. – Т. 6. – №. 3. – С. 99-106, doi: 10.30525/2256-0742/2020-6-3-99-106

[5] Romero C., Ventura S. Educational data mining and learning analytics: An updated survey //Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery – 2020 – Т. 10 – №. 3 – С. e1355 doi: 10.1002/widm.1355

[6] Zamri N. E. et al. Amazon Employees Resources Access Data Extraction via Clonal Selection Algorithm and Logic Mining Approach //Entropy – 2020 – Т. 22 – №. 6 – С. 596, DOI: 10.3390/e22060596

[7] Z. Rakhmetullina, R. Mukhamedova, R. Mukasheva and E. Aitmukhanbetova, "Mathematical Model for Clinical Decision Support

System Using Genetic Algorithm," 2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Istanbul, Turkey, 2020, pp. 1-5, doi: 10.1109/ISMSIT50672.2020.9255150

[8] Camacho D. et al. "The four dimensions of social network analysis: An overview of research methods, applications, and software tools" //Information Fusion. – 2020 – T. 63 – C. 88-120, DOI: 10.1016/j.inffus.2020.05.009

[9] L. Cai, R. Datta, J. Huang, S. Dong and M. Du, "Sleep Disorder Data Stream Classification Based on Classifiers Ensemble and Active Learning," 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), San Diego, CA, USA, 2019, pp. 1432-1435, doi: 10.1109/BIBM47256.2019.8983119

[10] F. Aziz, A. Lawi and E. Budiman, "Increasing Accuracy of Ensemble Logistics Regression Classifier by Estimating the Newton Raphson Parameter in Credit Scoring," 2019 5th International Conference on Computing Engineering and Design (ICCED), Singapore, 2019, pp. 1-4, doi: 10.1109/ICCED46541.2019.9161078

[11] K. Noh, J. Lim, S. Chung, G. Kim and H. Jeong, "Ensemble Classifier based on Decision-Fusion of Multiple Models for Speech Emotion Recognition," 2018 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, Korea (South), 2018, pp. 1246-1248, doi: 10.1109/ICTC.2018.8539502

[12] X. LI, T. SHI, P. LI and W. ZHOU, "Application of Bagging Ensemble Classifier based on Genetic Algorithm in the Text Classification of Railway Fault Hazards," 2019 2nd International Conference on Artificial Intelligence and Big Data (ICAIBD), Chengdu, China, 2019, pp. 286-290, doi: 10.1109/ICAIBD.2019.8836988

[13] S. Guo, H. He and X. Huang, "A Multi-Stage Self-Adaptive Classifier Ensemble Model With Application in Credit Scoring," in IEEE Access, Vol. 7, pp. 78549-78559, 2019, doi: 10.1109/ACCESS.2019.2922676

[14] S. P. Paramesh, C. Ramya and K. S. Shreedhara, "Classifying the Unstructured IT Service Desk Tickets Using Ensemble of Classifiers," 2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS), Bengaluru, India, 2018, pp. 221-227, doi: 10.1109/CSITSS.2018.8768734

[15] N. A. Nguyen-Thi, Q. Pham-Nhu, S. Luong-Ngoc and N. Vu-Thanh, "Using noise filtering techniques to ensemble classifiers for credit scoring," 2019 International Conference on Technologies and Applications of Artificial Intelligence (TAAI), Kaohsiung, Taiwan, 2019, pp. 1-6, doi: 10.1109/TAAI48200.2019.8959831

[16] Uvalieva, I; Turganbayev, E; Tarifa, F. Development of information system for monitoring of objects of education on the basis of intelligent technology: a case study of Kazakhstan//15th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA) Hammamet, TUNISIA, 2014, 909-914 (ISBN:978-1-4799-5907-5)WOS:000392734600124

[17] Mohammadi A., Shaverizade A. Ensemble deep learning for aspect-based sentiment analysis //International Journal of Nonlinear Analysis and Applications – 2021 – Т. 12 – С. 29-38, DOI: 10.22075/IJNAA.2021.4769

[18] Seni G., Elder J. F. Ensemble methods in data mining: improving accuracy through combining predictions //Synthesis lectures on data mining and knowledge discovery – 2010 – Т. 2 – №. 1 – С. 1-126, doi: 10.2200/s00240ed1v01y200912dmk002

[19] Dimić G. et al. An approach to educational data mining model accuracy improvement using histogram discretization and combining classifiers into an ensemble //Smart Education and e-Learning 2019 – Springer, Singapore, 2019 – С. 267-280, DOI: 10.1007/978-981-13-8260-4_25

[20] Lin J. L. et al. An ensemble method for inverse reinforcement learning //Information Sciences – 2020 – Т. 512 – С. 518-532, DOI: 10.1016/j.ins.2019.09.066