Targeting additional effort for students’ success improvement: The highest effect group selection method

Dušan Munđar* 1 and Damira Keček 2
1 Faculty of Organization and Informatics, University of Zagreb, Varaždin, Croatia
2 University North, Varaždin, Croatia

Abstract. Goal of our research is to present a method for selection of a group of students. The selected group of students is supposed to receive additional teaching attention in order to improve their performance in the course. Guiding line used, for the group selection, is: value of an action is proportional to the benefit that produces to a customer. In our case customers are students. The selection method is based on multinomial logistic regression, Poisson-type discrete variables modelling the number of points achieved in a term exam, and transition probability matrix. We use data of a study progress during monitoring program in an undergraduate mathematical course. Demographic data and other attributes about previous performance were not included in the analyses. In the first part of the paper we present methodology, while in the second part we introduce data used for demonstration of the proposed method. At the end of the paper, an individual approach for the final selection is proposed. Criteria for selection is clear: increased probabilities of obtaining desired final grade for a student. Weighing of criteria is subjective depending on the goals of the decision makers.

Keywords: learning analytics, predictive analytics, students’ success, multinomial logistic regression, Poisson distribution

1. Introduction

Development of information and communication technologies provides collection of large amounts of data about the students learning process and their study progress. Learning analytics, as a discipline whose purpose is to analyze and improve learning process, accordingly gains additional development possibilities.

*Corresponding author.
Learning analytics does not involve only the measurement and data collection by some institution, but also the understanding of these data, analysing and creating reports based on the data in order to understand and optimize the learning and the environment in which learning occurs (Society for Learning Analytics Research, 2011), (U.S. Department of Education, Office of Educational Technology, 2012). Significant challenges (Ferguson, 2012) that Learning Analytics must deal with are systematized in the following: to consolidate experience from the learning sciences, to work with a wider range of datasets, to engage with learner perspectives and to develop a set of ethical guidelines directions. To overcome the gap in knowledge exchange and conversation between researchers, vendors and experts, Siemens (Siemens, 2012) pointed out some improvements inside Learning Analytics, which involves new tools, techniques and people development, addressing concerns connected with the data openness and ethics, expanding and transitioning the goal of analytics activity and enhancing connections to related areas. In order to determine how the Learning Analytics tools are successful and what is their impact on learning, Scheffel et al. (Scheffel et al., 2014) proposed a five-dimensional framework of quality indicators to standardize the evaluation of the Learning Analytics tools. The proposed framework consists of five criteria and quality indicators as follows: Objectives, Learning Support, Learning Measures and Output, Data Aspects and Organisational Aspects.

Reyes (Reyes, 2015) concluded that Learning Analytics will provide students to take the initiative in learning process advance by using the data about the factors of student’s success, about allocation of resources and the effectiveness of teaching which will finally contribute to the learning effectiveness. Authors Mor et al. (Mor et al., 2015) showed that the synergistic effects of Learning Analytics and two approaches, Learning design and Teacher inquiry, can improve the quality of teaching and learning. In this context, learning design contributes with its semantic structure for analytics and teacher inquiry with defining profound questions to analyse. As concluded in (Gašević, 2015), Learning Analytics should be developing, building and connecting with existing knowledge about learning and teaching. It is important to inform a wider audience with the possibilities that Learning Analytics provide. Various activities, such as projects (e.g. Learning Analytics Community Exchange, 2017), conferences (e.g. Learning Analytics and Knowledge, 2017) and journals (e.g. Journal of Learning Analytics, 2017) are held to accomplish the goal of providing the information. It can be concluded that Learning Analytics nowadays receives importance in academic as well as in professional circles.

This paper presents the results of estimating probability of obtaining the final grade on mathematical course based on points achieved at the two term exams as well as the impact of the additional points obtained in the first term exam on the final grade. Two research questions are: “Which student should be chosen to effect on to increase probability of passing the course?” and “Which student should be chosen to increase final grade on the course?” Student at risk is student who needs a little assistance from professor for passing the course or for increasing the final grade. In the paper identification of students at risk is presented. The developed
methodology for the detection of these students is based on multinomial logistic regression. Computation is done in R Studio, a user interface for R programming language.

2. Literature review

This section provides an overview of relevant research related to the use of Learning Analytics in order to improve students’ success as well as predict students at risk, i.e. students who are at greater risk of failing the course or dropping out of the course.

By using visualization, decision trees, class association rules and clustering approaches authors Kotsiantis et al. (Kotsiantis et al., 2013) analyzed students’ perceptions of Moodle and students’ interaction with Moodle. As the collected data combined with Learning Analytics approaches can provide efficient information about the educational process, in this study the analysis showed that failure in the course was connected with negative students’ attitudes and perceptions according to Moodle, while the increased use of Moodle led to excellent grades. By using the logistic regression, the authors Barber and Sharkey (Barber, 2012) created a model which identifies students who are at risk that will not succeed in enrolled course. The probability that a student will fail the course is calculated based on data obtained from learning management system and student system. Knowing such indicators can help students succeed in the course by beforehand intervention and, if necessary, by providing additional services. Munđar and Erjavec (Munđar, 2015), by using logistic model, assessed the probability of passing the exam in mathematics based on the students’ results on Matura state exam and high school grades. Identification of such data is important not only for successfully passing the mathematical courses, but also for the study success at whole. Similarly, a model for predicting mathematics course passing rates was developed by applying classification trees and neural networks (Keček, 2012). Regarding the accuracy of classification, which students have passed the course and which have not, classification trees have proven to be a better method in predicting mathematics course passing rates than neural networks. Prediction of the study success was done in (Munđar et al., 2015) by using multiple regression. Study success was measured by number of ECTS credits which student achieved one year after enrolment at the faculty, while enrolment points achieved based on the results of State Matura exams and points achieved based on high school grades were considered as predictors for that study success. The results of that research showed that the high school grades in general are better predictor in achieving more ECTS credits than the State Matura exam. In order to improve retention, completion and graduation rates authors Essa and Ayad (Essa, 2012) developed a Student Success System that enables, through the machine intelligence and statistical techniques, monitoring student with the aim of improving its study success. The System can identify a student who is at risk, understand why the student is endangered, plan interventions to alleviate this risk and can provide feedbacks about effectiveness of undertaken interventions.
3. Methodology

Based on data, multinomial logistic regression and modelling with Poisson distribution type of random variables as the basic methods, we are proposing the following method for detection of targeting group. Strategy for detection students at risk is based on transition rates from 1st term exam to 2nd term exam and from achieved points at 2nd term to the final grade. The emphasis is placed on points increase at the 2nd term exam whose shift results in the greatest change in success probability. Our method has three main parts. The first part is presentation of multinomial logistic regression for final grade prediction. The second part is presentation of a model for transition probabilities which are used to predict success of a student depending of success in the first part of the course. The third, presentation of a method for valuation of additional effort by estimation of the effect it produces.

3.1. Grade prediction

Multinomial logistic regression is a method developed to estimate the probabilities of the categorical response variable, which has more than two possible outcomes. Generally, for \( l \) possible outcomes \( \{1, 2, \ldots, l\} \) of the response variable \( Y \), where outcome 1 is chosen as a standard category, multinomial probabilities \( P[Y = k] = p_k, k = 1, \ldots, l \) are parameterized as

\[
p_1 = P[Y = 1] = \frac{1}{1 + \sum_{i=2}^{l} e^{\alpha_i + x\beta_i}} \tag{1}
\]

and

\[
p_k = P[Y = k] = \frac{e^{\alpha_k + x\beta_k}}{1 + \sum_{i=2}^{l} e^{\alpha_i + x\beta_i}}, \tag{2}
\]

for \( k = 2, \ldots, l \). Those probabilities sum to 1 and any other outcome, instead of outcome 1, could be chosen as a standard category. (Ledolter, 2013).

In our paper, we use multinomial logistic regression in two occasions. Firstly, we predict grade based on success on the 1st term exam. In the second part of the analysis, we predict grade by using points achieved on the first two term exams.

3.2. Transition rates modelling

Prediction of number of points achieved on the 2nd term exam is estimated by using Poisson distribution variables. Expected value of those variables depends on the number of points achieved on the 1st term exam. Transition matrix, \( T = [t_{ij}] \), of type \((M + 1) \times (M + 1)\), contains transition probabilities from the 1st term exam score, i.e. number of points achieved, to score achieved on the 2nd term exam. \( M \) is maximal number of points that a student can achieve in a term exam (same for all term exams in our case).

\[
t_{x_1+1, x_2+1} = P(X_2 = x_2 \mid X_1 = x_1), \text{ where } x_1, x_2 \in \{0, \ldots, M\} \tag{3}
\]
As mentioned before, we assume that the number of points achieved for a student at a term exam follows Poisson distribution. Probability of scoring points are given by

\[
P(X_2 = x_2) = \frac{(\lambda x_1)^{x_2} \exp(-\lambda x_1)}{x_2!}
\]  

(4)

Expected value of number of points achieved, \( \lambda x_1 \), depends on number of points \( x_1 \) achieved at the first term exam.

\[
\lambda x_1 = \lambda (\beta_0 + \beta_1 x_1 + \beta_2 x_1^2)
\]  

(5)

The parameters \( \lambda \), \( \beta_0 \), \( \beta_1 \), and \( \beta_2 \) are estimated using maximum likelihood estimation method, i.e. by determining value of parameters that minimize the value cost function \( C \)

\[
C(\beta_0, \beta_1, \beta_2) = - \sum_{(x_1, x_2) \in S} \ln (p(X_2 = x_2 \mid X_1 = x_1))
\]  

(6)

Summation goes for each observation, combination of points achieved \((x_1, x_2)\) at the 1st and the 2nd term exam for each student, in an academic year which is denoted by \( S \).

3.3. Effect measurement

In order to compute estimated effect of addition teaching effort we need to estimate parameters of another multivariate model, i.e. we estimate probability of a student receiving the grade \( g \), when we know his/her score at the first two term exams:

\[
P(G = g \mid X_2 = x_2, X_1 = x_1).
\]  

(7)

In our case, we do that by modelling probabilities of each grade with multinomial logistic regression

\[
P[Y = 4] = \frac{1}{1 + \sum_{i=1}^{3} e^{\alpha_i + x_2 \beta_i}}; \quad P[Y = k] = \frac{e^{\alpha_k + x_2 \beta_k}}{1 + \sum_{i=1}^{3} e^{\alpha_i + x_2 \beta_i}}, \quad k = 1, 2, 3
\]  

(8)

where \( x = (x_1, x_2) \) are results of the 1st and the 2nd term exam.

Initial probability that student achieve grade \( g \), knowing his/her result at the 1st term exam, i.e. knowing number of points \( x_1 \) achieved at the 1st term exam, is given by:

\[
P_0(G = g X_1 = x_1) = \sum_{x_2=0}^{M} P(G = g \mid X_2 = x_2, X_1 = x_1) P(X_2 = x_2 \mid X_1 = x_1).
\]  

(9)
Probability after treatment, i.e. probability that student who achieved $x_1$ points at the 1st term exam will achieve grade $g$, after receiving treatment of additional teaching action equivalent to raise of $m$ points, is:

$$P_{+m}(G=g|X_1=x_1) = \sum_{x_2=0}^{M} P(G=g|X_2=\min(x_2+m, M), X_1=x_1) P(X_2=x_2|X_1=x_1)$$

(10)

Difference in probabilities in our focused measure is then:

$$\Delta_{+m}(G = gX_1 = x_1) = P_{+m} - P_0$$

(11)

If decision maker do not value increase in each grade equally, it is possible to take the opinion in valuation by weighing increased probabilities, for example by formula:

$$V(x_1) = \sum_{g=2}^{4} w(G = g) \Delta_{+m}(G \geq gX_1 = x_1).$$

(12)

4. Data

In this paper, we conducted the analysis for the course Financial Mathematics (FM) performed at the study program Information and Business Systems (IBS) at the Faculty of Organization and Informatics, University of Zagreb. In this analysis, 158 full-time students of academic year 2015/2016 and 144 full-time students of academic year 2016/2017 were involved, while part-time students were not included in the analysis.

Elements of students’ work monitoring in this course included: Homework assignments (HW), Short examinations (SE), Project (PRO), Term exam 1 (TE1), Term exam 2 (TE2) and Term exam 3 (TE3). Students can achieve maximum of 20 points at each Term exam as well as at the Project, maximum of 10 points at Homework assignment and maximum of 10 points at Short examinations. We based this analysis on the term exams scores. Term exams scores are the most significant variables for students’ success prediction (when other attributes about students, except their performance on course, are not available).

Table 1 presents descriptive statistics of variables Homework assignment (HW), Short examinations (SE), Project (PRO), 1st term exam (TE1), 2nd term exam (TE2) and 3rd term exam (TE3) for academic year 2015/2016 and for academic year 2016/2017. Total number of points (TOTAL) is equal to the sum of points achieved through HE, SE, PRO, TE1, TE2 and TE3. It can be noted that all of the observed measures of location, mean, median, first and third quartile, for most variables in academic year 2016/2017 had a bit higher values than in academic year 2015/2016. Values of variables HW and SE in two observed academic years were almost equal. The highest total number of points achieved in academic year 2015/2016, out of the maximum 100, was 90 and in academic year 2016/2017 was
95 points. It can also be noted that all measures of location for variable TOTAL were very similar in academic year 2016/2017 and academic year 2015/2016.

Table 1. Descriptive statistics of students’ points achieved during the course.

<table>
<thead>
<tr>
<th>Financial mathematics (FM)</th>
<th>Academic Year</th>
<th>Academic Year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2015/2016</td>
<td>2016/2017</td>
</tr>
<tr>
<td></td>
<td>1st Q.</td>
<td>Median</td>
</tr>
<tr>
<td>Home work assignment (HW)</td>
<td>5.00</td>
<td>6.00</td>
</tr>
<tr>
<td>Short examination (SE)</td>
<td>3.87</td>
<td>4.78</td>
</tr>
<tr>
<td>Project (PRO)</td>
<td>12.00</td>
<td>14.00</td>
</tr>
<tr>
<td>Term exam 1 (TE1)</td>
<td>5.63</td>
<td>9.00</td>
</tr>
<tr>
<td>Term exam 2 (TE2)</td>
<td>7.50</td>
<td>11.00</td>
</tr>
<tr>
<td>Term exam 3 (TE3)</td>
<td>7.00</td>
<td>10.00</td>
</tr>
<tr>
<td>Total number of points (TOTAL)</td>
<td>41.00</td>
<td>54.00</td>
</tr>
</tbody>
</table>

Figure 1. Graphical representation of points and grades dependency (ac. year 2016/2017).

Outcomes of the categorical response variable named GRADE are: 0 – student did not accomplish their minimum commitments related to the course, 1 – student accomplished their minimum commitments related to the course, but didn’t pass the course, while categories 2, 3, 4 and 5 represents grades according to the Croatian national grading scale. Grade 2 is equivalent to sufficient, 3 is equivalent to good, 4 to very good and 5 to excellent. Grades are assigned according to the previously agreed scoring scale rating and according to the achieved total number of points. According this scale, in academic year 2015/2016 around 68 percent of students passed the course and two students got excellent grade. In academic year
2016/2017 the results are even a bit better. Around 71 percent of students passed the course, and two of those were given excellent grade. In both academic years, the largest number of students got grade 2.

For better understanding of above results, Figure 1. shows a graphical representation of the results through box plots, where each of the box plot shows how the variable GRADE depends on the points achieved on the individual element of students’ work monitoring.

5. Results

5.1. Grade prediction: Case of Financial Mathematics course in academic year 2016/2017

For further data analysis, outcomes 0 and 1 of the variable GRADE were merged, as well as outcomes 4 and 5, since grades categories 0 and 5 included only few students. Thus, the analysis continued with four outcomes of variable GRADE: 1, 2, 3 and 4. For those outcomes of the variable GRADE, where outcome 4 represents a standard category, multinomial probabilities are as follows:

\[
P(Y = 4) = \frac{1}{1 + \sum_{i=1}^{3} e^{\alpha_i + x \beta_i}} = \frac{1}{1 + e^{\alpha_1 + x \beta_1} + e^{\alpha_2 + x \beta_2} + e^{\alpha_3 + x \beta_3}}
\]

\[
P(Y = k) = \frac{e^{\alpha_k + x \beta_k}}{1 + \sum_{i=1}^{3} e^{\alpha_i + x \beta_i}} = \frac{e^{\alpha_k + x \beta_k}}{1 + e^{\alpha_1 + x \beta_1} + e^{\alpha_2 + x \beta_2} + e^{\alpha_3 + x \beta_3}}, \quad k = 1, 2, 3.
\]

Prediction of students’ success in this analysis is based on the accomplishments at course assignments. Other variables, such as social variables e.g. gender, success at previous courses were not taken in consideration since the goal of the paper is introduction of the methodology for success assessment. Multinomial logistic regression, shortly explained in the Methodology part of this paper, can be used for estimation of probability for obtaining final grade.

| Coefficients | Estimate | Std. Error | Z value | Pr(<|z|) |
|--------------|----------|------------|---------|----------|
| $\alpha_1$  | 16.162   | 3.970      | 4.078   | 4.54e−05 |
| $\alpha_2$  | 14.412   | 3.947      | 3.651   | 0.000261 |
| $\alpha_3$  | 12.080   | 3.892      | 3.104   | 0.001909 |
| $\beta_1$   | −1.209   | 0.253      | −4.786  | 1.7e−06  |
| $\beta_2$   | −0.912   | 0.244      | −3.741  | 0.000183 |
| $\beta_3$   | −0.699   | 0.246      | −2.967  | 0.003010 |
Targeting additional effort for students’ success improvement: The highest effect group... 9

Figure 2 presents estimated probabilities calculated using only results of the 1st term exam, for academic year 2016/2017. Grade 0 and grade 1 are joint in one class called Grade 1. Similar approach is used for Grade 4 and Grade 5 since we have small number of students receiving Grade 4 (3 students) or Grade 5 (2 students).

![Figure 2. Estimated probabilities of final grade based on the 1st term exam: Case Financial Mathematics (ac. year. 2016/2017).](image)

5.2. Transition probabilities: Case of Financial Mathematics course in academic year 2016/2017

Transition matrix, $T = [t_{ij}]$, represents matrix of transition probabilities from 1st term exam score of points achieved to scoring point achieved at the 2nd term exam. Using methodology for assessment of parameters $\beta_0$, $\beta_1$, and $\beta_2$, we determined values of the transition probabilities. Table 3 presents submatrix of matrix $T$. Rows and columns names are number of points achieved at the 1st term exam, $x_1$, and the 2nd term exam, $x_2$, of values of 0, 5, 10, 15 and 20.

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.03%</td>
<td>...</td>
<td>8.67%</td>
<td>...</td>
<td>10.27%</td>
</tr>
<tr>
<td>5</td>
<td>0.01%</td>
<td>...</td>
<td>5.87%</td>
<td>...</td>
<td>11.95%</td>
</tr>
<tr>
<td>10</td>
<td>0.00%</td>
<td>...</td>
<td>2.78%</td>
<td>...</td>
<td>12.30%</td>
</tr>
<tr>
<td>15</td>
<td>0.00%</td>
<td>...</td>
<td>0.84%</td>
<td>...</td>
<td>9.16%</td>
</tr>
<tr>
<td>20</td>
<td>0.00%</td>
<td>...</td>
<td>0.15%</td>
<td>...</td>
<td>4.22%</td>
</tr>
</tbody>
</table>

Table 3. Estimated transition probabilities for points at the 1st and the 2nd term exam.

5.3. Effect measurement: Case of financial mathematics in academic year 2016/2017

In order to compute estimated effect of addition teaching effort we need to estimate parameters of another multivariate model, i.e. we estimate probability of a student receiving the grade g, when we know his/her score at the first two term exams:
\[ P(G = g \mid X_2 = x_2, X_1 = x_1) \]

So in our case of modelling by multinomial logistic regression:

\[ P[Y=4]=\frac{1}{1+\sum_{i=1}^{3} e^{\alpha_i+x_{ik}}}; \quad P[Y=k]=\frac{e^{\alpha_k+x_{ik}}}{1+\sum_{i=1}^{3} e^{\alpha_i+x_{ik}}}, \quad k = 1, 2, 3, \]

where are \( z = (x_1, x_2) \) results of the 1st and the 2nd term exam. Estimated values of coefficients with their standard error, and accompanied \( z \)-score and associated \( p \)-value are given in the Table 4.

Table 4. Parameters for multinomial logistic regression of grade probability based on results of the first two term exams.

| Coefficients | Estimate | Std. Error | Z value | Pr(<|z|) |
|--------------|----------|------------|---------|---------|
| \( \alpha_1 \) | 25.249 | 4.506 | 5.603 | 2.10e-08 |
| \( \alpha_2 \) | 19.570 | 4.290 | 4.561 | 5.08e-06 |
| \( \alpha_3 \) | 12.870 | 3.941 | 3.266 | 0.001091 |
| \( \beta_{11} \) | -1.036 | 0.281 | -3.685 | 0.000228 |
| \( \beta_{12} \) | -0.734 | 0.265 | -2.769 | 0.005614 |
| \( \beta_{13} \) | -0.503 | 0.256 | -1.965 | 0.049466 |
| \( \beta_{21} \) | -1.010 | 0.290 | -3.477 | 0.000508 |
| \( \beta_{22} \) | -0.597 | 0.269 | -2.221 | 0.026367 |
| \( \beta_{23} \) | -0.256 | 0.251 | -1.019 | 0.308090 |

For example, our model for Grade 4 (combines Grade 4 and 5) is:

\[ P[Y=4|X_2=x_2, X_1=x_1]=\frac{1}{1+e^{25.2-1.04x_1-0.50x_2+e^{19.6-0.73x_1-0.60x_2}+e^{12.9-0.50x_1-0.26x_2}}} \]

(17)

In special case, for students that achieved 15 points at the 1st and 15 points at the 2nd term exam, model gives probability of 16.33 % to receive Grade 4 or 5:

\[ P[Y = 4 \mid X_2 = 15, X_1 = 15] = 0.1633 \]

(18)

Additional effort (with estimated effect of \( m \) additional points at second term exam) given to a student, which received po\( x_1 \)nts at the first term, and would without treatment received at\( x_2 \) the second term exam, improves probability of the grade 4 or 5.

\[ \Delta P_{+m|x_1}=P[Y=4 \mid X_2=\min(x_2+m,M), X_1=x_1]-P[Y=4 \mid X_2=x_2, X_1=x_1] \]

(19)

\[ \Delta P_{+1|x_1}=P[Y=4 \mid X_2=16, X_1=15]-P[Y=4 \mid X_2=15, X_1=15]=0.2077 \]

(20)

In special case, for students that achieved 15 points at the 1st and would receive 15 points at the 2nd term exam effect of effort of 1 (\( m=1 \)) increases probability of receiving Grade 4 or 5 for 4.44 %, or to value of 20.77 %.
Targeting additional effort for students’ success improvement: The highest effect group...

If the effect of the effort is weighted with estimated probability of receiving a special number of points in the case of absence of additional effort.

\[
\Delta P_{+m} = \sum_{x_2=0}^{M} \Delta P_{+m|x_1} P(X_2 = x_2 \mid X_1 = x_1)
\]  

(21)

Table 5 presents increased probability of becoming equal of higher grade than initial estimated by multinomial logistic regression for each case of points at the 1st term exam. The table can be used to estimate effect of additional effort. For example, if we choose to give additional attention, equivalent to knowledge needed to earn one additional point, to a student who have 5 points at the 1st term exam, we increase probability of passing the course (receiving Grade 2 or higher) from 0.437 to value of 0.819.

Table 5. Table of probabilities and improved probabilities of final grade based of the 1st term exam.

<table>
<thead>
<tr>
<th>Points at the 1st term exam</th>
<th>Probability of the grade</th>
<th>Increased probability of the grade</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>2 or higher</td>
</tr>
<tr>
<td>0</td>
<td>0.137</td>
<td>0.148</td>
</tr>
<tr>
<td>1</td>
<td>0.172</td>
<td>0.189</td>
</tr>
<tr>
<td>2</td>
<td>0.213</td>
<td>0.237</td>
</tr>
<tr>
<td>3</td>
<td>0.256</td>
<td>0.296</td>
</tr>
<tr>
<td>4</td>
<td>0.311</td>
<td>0.363</td>
</tr>
<tr>
<td>5</td>
<td>0.361</td>
<td>0.437</td>
</tr>
<tr>
<td>6</td>
<td>0.413</td>
<td>0.517</td>
</tr>
<tr>
<td>7</td>
<td>0.6</td>
<td>0.598</td>
</tr>
<tr>
<td>8</td>
<td>0.488</td>
<td>0.678</td>
</tr>
<tr>
<td>9</td>
<td>0.504</td>
<td>0.751</td>
</tr>
<tr>
<td>10</td>
<td>0.503</td>
<td>0.816</td>
</tr>
<tr>
<td>11</td>
<td>0.482</td>
<td>0.869</td>
</tr>
<tr>
<td>12</td>
<td>0.445</td>
<td>0.911</td>
</tr>
<tr>
<td>13</td>
<td>0.394</td>
<td>0.943</td>
</tr>
<tr>
<td>14</td>
<td>0.335</td>
<td>0.965</td>
</tr>
<tr>
<td>15</td>
<td>0.272</td>
<td>0.979</td>
</tr>
<tr>
<td>16</td>
<td>0.210</td>
<td>0.988</td>
</tr>
<tr>
<td>17</td>
<td>0.153</td>
<td>0.994</td>
</tr>
<tr>
<td>18</td>
<td>0.105</td>
<td>0.997</td>
</tr>
<tr>
<td>19</td>
<td>0.066</td>
<td>0.999</td>
</tr>
<tr>
<td>20</td>
<td>0.039</td>
<td>0.999</td>
</tr>
</tbody>
</table>
6. Conclusions

Learning analytics is a discipline whose purpose is to analyze and improve learning process. Analysis of available data about the learning process gives additional development possibilities of the discipline. In our paper, we have introduced a simple probabilistic model for assessment of effect that additional teaching effort can produce for different subgroups of students. Subgroups of students are, in our example, divided considering only results in term exams. Analysts could take more variables or other variables in consideration when analyzing student progress data.

Presented method applied on data about progress on the Financial Mathematics course, an undergraduate mathematics, gave us assessment of the effect that teaching effort needed for a student to gain additional five points at the 2\text{nd} term exam. Example, presented in the Results part of this paper, for a student who has 5 points at 1\text{st} term exam assessed effect in increased probability of passing the course (receiving Grade 2 or higher) from 0.437 to value of 0.819. Similar conclusions can be deducted from the results for any student depending on the number of points achieved at the 1\text{st} term exam.

Teaching effect presented in term exam points increase ($m$) can be transformed to final grade probabilities, using presented method, for any teaching effect $m$. In special case, for students that achieved 15 points at the 1\text{st} term exam and would receive 15 points at the 2\text{nd} term exam effect of effort of 1 ($m = 1$) increases probability of receiving Grade 4 or 5 for 4.44 \%, or to value of 20.77 \%.

References


Contact addresses:
Dušan Mundar
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, HR-42000 Varaždin, Croatia
e-mail: dusan.mundjar@foi.hr

Damira Keček
University North
104. brigade 3, HR-42000 Varaždin, Croatia
e-mail: damira.kecek@unin.hr
Usmjeravanje dodatnog angažmana za poboljšanje uspjeha studenata: odabir grupe s najvećim učinkom

Dušan Munđar¹ i Damira Keček²

¹Fakultet organizacije i informatike, Sveučilište u Zagrebu, Varaždin, Hrvatska
²Sveučilište Sjever, Varaždin, Hrvatska


Ključne riječi: analitika učenja, prediktivna analitika, uspjeh studenata, multinomijalna logistička regresija, Poissonova distribucija