Improving University Operations with Data Mining: Predicting Student Performance

Mladen Dragičević, Mirjana Pejić Bach, Vanja Šimičević

***Abstract*—**The purpose of this paper is to develop models that would enable predicting student success. These models could improve allocation of students among colleges and optimize the newly introduced model of government subsidies for higher education. For the purpose of collecting data, an anonymous survey was carried out in the last year of undergraduate degree student population using random sampling method. Decision trees were created of which two have been chosen that were most successful in predicting student success based on two criteria: Grade Point Average (GPA) and time that a student needs to finish the undergraduate program (time-to-degree). Decision trees have been shown as a good method of classification student success and they could be even more improved by increasing survey sample and developing specialized decision trees for each type of college. These types of methods have a big potential for use in decision support systems.

***Keywords*—**Data mining, knowledge discovery in databases, prediction models, student success.

I. Introduction

CURRENTLY in Republic of Croatia there is an ongoing harmonization process in accordance with the Bologna declaration and reform of the higher education system and increase in student performance is one of the goals of the Bologna reform. A model for predicting student performance would provide useful information for proposing policies that could be implemented in the educational process and environment [1], [2].

There are a number of factors that are closely related to studying which influence students' success, like lecture attendance, passing the course by attending preliminary exams or regular end of term exams, student responsibility, time spent studying for an exam etc. Since there are several studies that focus these factors in their research, in this study focus is shifted more on demographic factors of students. In previous research it has been indicated that demographic factors have significant influence on student success and therefore they should be included in constructing models for predicting student success.

Term knowledge discovery in databases or data mining has been introduced relatively recently, in the nineties, but the scope that it covers has a much longer history [3]. Roughly speaking, scope that data mining covers include is statistics, artificial intelligence and machine learning. Statistics is the basis of the most technologies that are used in the process of knowledge discovery in databases. The purpose of statistics in the process is to study the data itself and correlations between the data. On the other side, artificial intelligence is based on heuristics and it represents an attempt to approach statistical problems similar to the human way of thinking [4]. Machine learning has a bit of both approaches and therefore it can be considered as a link between these two concepts. Data mining is usually carried out on larger quantities of data and extraction of new knowledge is usually done from databases, but it is not a rare concept to conduct data mining on data collected through surveys and by other methods. The process of knowledge discovery from databases is usually done in five steps: (1) Understanding problem, (2) Understanding data, (3) preparing data, (4) Modeling data, (5) Evaluation of the model. Knowledge discovery in databases is used in a number of applications for predicting students' success.

There are number of research that used intelligent methods for predicting students' success. Hardgrave and Wilson [5] compare neural networks with traditional statistical methods for the purpose of predicting students' success in the graduate study. In the follow-up of the research [32] use more additional models like linear regression, logistic regression and discriminant analysis. Naik and Ragothaman [33] use logit and probit models and compare them with neural networks for predicting students' success. Zaidah and Daliela [6] compare neural networks, linear regression and decision trees, and Oladokun, Adebanjo and Charles-Owaba [7] use multi-layer perception network. Matković, Tomić and Vehovec [8] have published a paper in which they analyze the efficiency of the process of higher education on a random sample of freshly graduated students. As a result of the research they concluded that the chances of successfully graduating are in correlation to the socioeconomic status of students. Zekić-Sušac, Frajman-Jakšić and Drvenkar [9] described models they created for predicting student success using neural network algorithms and classification decision trees. In the paper they also analyze factors that influence student success. Models were created based on demographic data of students, behavior and attitude toward studying. Success was measured by using GPA. Shaw, Marini and Mattern [10] used hierarchical linear modeling in order to predict first-year grade point average by using various variables of Advanced Placement exam. Another example of multilevel modeling is paper by Rienties and Tempelaar [11]. They showed that academic adjustment is

M. Dragičević is with Zagrebačka banka (part of Unicredit Group), Trg bana Josipa Jelačića 10, 10000 Zagreb (e-mail: mladen.dragicevic@unicreditgroup.zaba.hr).

M. Pejić Bach is with Faculty of Economics & Business Zagreb, University of Zagreb, Trg J.F. Kennedyja 6, 10000 Zagreb (e-mail: mpejic@efzg.hr).

V. Šimičević is with Hrvatski studiji, Borongajska cesta 83d, Zagreb (e-mail: vsimicevic@hrstud.hr).

the primary predictor for academic success. Allen et al. [12] also used multilevel models to predict student achievement from observed teacher interactions with students. Carnahan B., Meyer G., Kuntz L. A. [13] compared statistical and machine learning classifiers for predicting whether or not a student truck driver would pass commercial driver license examination and found much higher rate of predictive accuracy by using machine learning algorithms. Another interesting approach for detection models of college withdrawal was proposed by Pleskac, T. J. et al. [14] whose proposal is to use cognitive modeling based on signal theory. Marquez-Vera et al. [15] proposed usage of genetic programming in order to predict student failure. There are also examples of prediction models in elementary and secondary schools like Chiu and Webb [16] who used decision trees (feature based modeling) in order to improve prediction of elementary subtraction skills of nine- to ten-year-old primary school students; while Luo, Aye, Hogan et al. [17] used structural equations to analyze influence of parental behavior and control on learning outcomes of students in secondary schools. There are also studies that indicated that achievement motivation, self-perceived academic achievement, and sex significantly contribute to the final secondary school success [18]. Study conducted by Levpuscek, Zupancic, and Socan [19] showed important role of individual differences (both in ability and personality), as well as student perceptions of parent and teacher, play in the students' performance in mathematics.

The purposes of this paper are to find models that predict student success based on data set collected on the sample of students of Faculty of Economics and Business in Zagreb and Faculty of Political Science in Zagreb and to demonstrate methods of knowledge discovery from databases on an actual problem, like allocation of students among colleges and optimization of the newly introduced model of government subsidies for higher education.

II. Determinants of Student Performance

A. Higher Education System in Croatia

The system of higher education that is currently in Croatia is the Bologna system of education. Croatia, being part of the European Higher Education Area and future member of the European Union, accepted this system of higher education which is currently active in the countries members of the European Union. Croatia joined the Bologna process in 2001, after signing the Bologna Declaration on the Ministerial Conference on the Bologna Process in Prague. The key main idea of the Bologna process is forming transparent and comparable system of education between member countries and to increase mobility of students by harmonization of university programs.

The idea of creating common European Higher Education Area was initiated for the first time in 1988, in the Magna Charta Universitatum which was signed by 430 university rectors. In 1998 the so called Sorbonne Declaration was signed (also known as Joint declaration on harmonization of the architecture of the European higher education system) and after this, Joint declaration of the European Ministers of Education was signed which title is European Higher Education Area, or better known as The Bologna Declaration. This document was signed by 29 countries agreeing that in the period from the year 2000 to year 2010 they will undertake concrete measures with the goal of forming common European Higher Education Area.

A decision was made that countries should accept the system in which degrees of education are easily recognizable and comparable and that with every diploma there should be supplement in which finished the degree of education is explained. Furthermore, a two-cycle system was suggested: undergraduate and graduate program, where student can attend a graduate program after finishing undergraduate program. The first cycle, undergraduate program, must last minimum 3 years and education acquired from this program must be adequate to prepare individuals for employment in the European labor market. By finishing second phase of education (graduate program) students obtain the Master degree. The new credit system was also suggested: ECTS (European Credit Transfer System) which is a standard for comparing the study attainment and performance of students. It enables mobility of students between universities and encourages the lifelong learning system of education. It also enables the mobility of professors and administrative personnel as time spent on research and teaching in other European countries is acknowledged. The cooperation of European universities and maintaining the quality of the program is strongly encouraged in the Bologna declaration. Also encouraged is development of criteria and methodologies which should be used to evaluate quality of higher education system.

In Croatia, problem of adopting the Bologna system was additionally more complicated as a system that was actually in Croatia at that moment was significantly different than in other countries in Europe. The university was not integrated, so every collage existed as an independent legal entity (whit its own plan of development). Only after the Act on Scientific Activity and Higher Education came into force in the summer of 2003 and after a number of amendments in 2004, Croatian universities begin restructuring in order to be more similar to European. Croatia had to implement parallel the Bologna process and organizational reform of universities.

In Croatia there are two types of studies: university study and professional study (this is shown in the scheme below). The university study consists of 3 phases or cycles, altogether regular duration of 8 years. A first degree is an undergraduate university study program which last 3 to 4 years and is weighted in the amount of 180 to 240 ECTS. Upon finishing the undergraduate university study program Bachelor's degree of the profession is obtained. Duration of the university graduate study program is one to two years - depending of how long university undergraduate program lasted. University graduate program is weighted in the amount of 60 to 120 ECTS and upon finishing one obtains Master's or Doctor of profession degree. Together undergraduate and graduate program last 5 years and student has to accumulate minimal of 300 ECTS in order to successfully finish, or 60 ECTS per year. In the third phase, there are two possibilities: postgraduate university study or postgraduate specialist study. The postgraduate university study is programmed whit duration of 3 years and upon finishing a PhD degree is obtained. University specifies the number of ECTS points needed to finish this program. Postgraduate specialist study program duration is one to two years and after finishing this program one obtains the title of specialist in a certain scientific area.

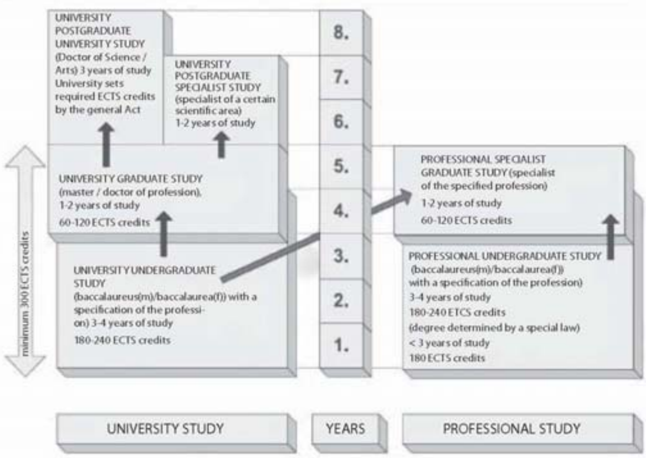

Fig. 1 Education concept in Republic of Croatia

A professional study program consists of two phases. First phase duration is 3 to 4 years and upon finishing, a Bachelor's degree is obtained. The first phase has 180 to 240 ECTS. If the program is shorter than 3 years, it has less than 180 ECTS and title that it is obtained upon finishing is determined by a special law. The second phase is professional specialist graduate study program in duration of one to two years and has 60 to 120 ECTS points. Upon finishing this program, one obtains title specialist of the specified profession. It should be also noted that it is possible to enroll in this program after finishing undergraduate university study program.

B. Measuring Student Success in Higher Education System in the Republic of Croatia

Student success can be measured in different ways; depending on individual perception what we consider to be successful. The official measure of success on universities is the grade point average (GPA) of all courses student has passed and accumulated ECTS points.

The GPA is calculated as the arithmetic mean of all past courses. In Croatia, this system is common at all levels of education. Grading scale is an ordinal scale and consists of five grades, where 5 indicate excellent and it is considered as top grade. To pass, the student needs to obtain at least 2 (means enough). Grade 1 means student did not satisfy the criteria to pass.

After Croatia joined Bologna process, new measure was introduced in the Croatia educational system: ECTS. As it was stated earlier, ECTS points were introduced in order to increase mobility and transparency. ECTS requirements for every study program are known and the number of accumulated ECTS points are often criteria for students did they satisfy to pass the semester. If the student did not accumulate enough ECTS points, he or she cannot continue into the next semester until enough ECTS points are accumulated which has influence on the time that a student needs to finish the program. Transparency of study programs by using ECTS points is achieved by establishing a methodology how much ECTS points certain course is worth.

Credit points of ECTS system are values (points) associated with every course or module and reflect work load of students that is required in order to finish the course or module [20]. Total sum of ECTS points per academic year is 60 ECTS, or 30 ECTS per semester. The student receives the full amount of associated ECTS points upon passing the course. One ECTS point is worth approximately 25 to 30 working hours. This number was derived from the weekly workload which is in most countries 37 to 50 work hours per week.

After joining Bologna process, another measure of success was introduced, as a combination of GPA and ECTS. It is called weighted GPA and it represents GPA weighted with ECTS. It is obtained as a sum of products of GPA and ECTS divided by the sum of ECTS.

III. Methodology

For the purpose of creating models, data were collected through an anonymous survey with random sampling method. Models for predicting were created with the help of an environment for data mining, text mining, machine learning, predictive and business analytics called RapidMiner (with usage of algorithms from WEKA environment) and integrated workspace and language R.

A survey was carried out on the sample of last year undergraduate program students, and 119 students of Faculty of Economics and Business in Zagreb and 83 students of the Faculty of Political Science in Zagreb participated in the survey. These two colleges were chosen because they are currently two largest colleges with a four year long undergraduate program (four years of undergraduate program plus one year of master program). Part of the collected data set was used to create models (80%), while another part was used as validation data set (20%) on which we tested the performance of algorithms.

In order to predict student success, it is necessary to determine by which criteria student success is measured. Distinguishing successful from unsuccessful students is somewhat subjective assessment, but there are two important measures that can be set as criteria by which we can distinguish successful from unsuccessful students: Grade Point Average (GPA) and time that a student needs to finish the undergraduate program (time-to-degree). In this paper ECTS points are not considered as a third measure by which student success can be measured exactly since it is a relatively new concept in education system of Croatia.

Questionnaire used in the survey was consisted of 18 questions that are referring to: demographic factors (gender, age, mother education), characteristics of studying (model of tuition fee, length of studying, parallel studying on another college, did the student finished already some college program, GPA) and situational criteria (working while studying, location of permanent residence, location of temporary residence, who is financing expenses and accommodation for the time of studying). Questionnaire ends with a series of questions that are related to their parents' employment, parents' income and parents' level of degree.

In order to develop models first it is required to pre-process the data. For certain algorithms specialized procedure of pre-processing is required, but some of the procedures are common to all of them. After the survey was carried out, missing values were found in 22 questionnaires and they were replaced by average values. Also, some of the variables were transferred from numeric to categorical.

Decision trees were used as an algorithm for predicting students' success. Decision trees are structures in the shape of an upside-down tree that represent groups of decisions from which rules for classification of data are created [21]. Developing a process of the decision tree is consisted of two parts: tree building and tree pruning. Tree is constructed on top to bottom principle where on top the root node is placed that branches out and in every next node it branches out further by certain criteria. Branching continues until certain criterion is met. Last node is called a leaf. This is the node that does not branch any further. In every node there is an attribute (independent variable) and in every leaf there is a dependent variable that is the goal to predict in a model. Based on this dependent variable model classifies data set. Created decision trees are to be read from root node to the leaves and this is how rules are created. Decision tree pruning is a process in which we remove unnecessary nodes, but in a way that pruned tree does not have an increased classification error on the data set (in setting boundaries), comparing to the unpruned decision tree.

IV. Results

Decision tree methodology is based on the algorithms that select only those attributes (based on different measures, e.g. entropy) that significantly contribute to successful classification. Two decision trees were computed using C4.5 algorithm and two dependent variables: Time-to-Degree criterion and Grade Point Average. Table I contains the variables used in modeling decision trees. In the first column there are all of the variables used. In second column there are variables statistically significant for the decision tree with time-to-degree dependent variable. In third column there are variables statistically significant for the decision tree with Grade Point Average. It is obvious that most of the variables in the dataset were not used in the model, such as: Model of tuition fee, Length of study, A parallel study on another college, Did the student finished already some college program, Location of permanent residence, Location of temporary residence, Parents employment, and Parents income.

TABLE I
SUMMARIZED VIEW OF CREATED DECISION TREES

Variable group	Variables used for modeling	Dependent variables	
		time-to-degree	Grade Point Average
Demographic	Gender	✓ (female)	not used
	Age	✓ (younger)	not used
Characteristics of study	Model of tuition fee	not used	not used
	Length of study	not used	not used
	A parallel study on another college	not used	not used
	Did the student finished already some college program	not used	not used
Situational criteria	Working while studying	✓ (yes)	✓ (yes)
	Location of permanent residence	not used	not used
	Location of temporary residence	not used	not used
	Dormitory or parents	✓ (dormitory)	
	Who is financing expenses and accommodation for the time of studying	✓ (student)	✓ (student)
Parent background	Parents employment	not used	not used
	Parents income	not used	not used
	Parents level of degree	✓ (mother) ✓ (father)	✓ (mother)

Table II presents a comparison of model classification performance. Models are created based on 80% data from total collected data set, while another 20% was used as a validation data set on which model evaluation was performed. Table II represents a comparison of model classification performance. A model that predicts student success based on the time-to-degree criterion correctly classified in 72.5% of the cases, while model that is based on the GPA was successful in 82.5%. We can conclude that model which predicts student success concerning the GPA is slightly better than the time-to-degree based model. Based on the results in the Table II, we can conclude that model developed with decision trees using C4.5 algorithms predict students' performed with great success.

TABLE II
COMPARISON OF MODEL CLASSIFICATION PERFORMANCE

	Decision tree 1 – Time-to-degree criterion	Decision tree 2 – GPA criterion
Correctly classified instances	72.5%	82.5%
Incorrectly classified instances	27.5%	17.5%

V. Discussion

Table III contains a comparison of our research and other selected research based on dependent variables, algorithms used and success of prediction. Of selected research our research was the only one that used time-to-degree as the dependent variable. All of the other presented research used Grade Point Average as dependent variables. Best prediction strength was attained by using neural networks for modeling determinants of students' success [22], [7]. However, other research did not prove the outperformance of neural networks over other methods [9]. In our research, usage of decision tree outperforms neural networks of other authors work with the predictive success of 82.5% if using Grade Point Average as the dependent variable.

TABLE III
COMPARISON WITH OTHER RESEARCH

Research	Dependent variables	Independent variables	Algorithms used	Success of prediction
Our research	time-to-degree; Grade Point Average	Gender, age, duration of enrolment in the study program, study program financing, parallel studying, finished other program, work, residence, cost financing, accommodation, parent education, parent income, parental employment status	Decision trees, C4.5 algorithm	75%; 82.5%
[22]	Grade Point Average	Campus location, citizen status, gender, race, GPA (4UGPA), GPA (2UGPA), GMAT, age, faculty, completed the program	Neural networks	89.13%
[6]	Grade Point Average	IT knowledge, finished school, programming knowledge, the financial situation of the parents	Neural networks, decision trees, regression	89.5% (neural networks)
[7]	Grade Point Average	UME result, GPA in courses mathematics, English, chemistry and physics, advanced mathematics, student age, time since finished high school, parent education, location of high school, finished high school type, location of faculty and residence, student gender	Neural networks	74%
[9]	Grade Point Average	Gender, scholarship, time spent learning, study materials, midterm exams, lectures, practicum, GPA importance for students	Neural networks, decision trees	66.26%; 79.5%

[a] Comparison of our research and other selected research based on dependent variables, independent variables significant for decision tress, algorithms used and success of prediction

VI. CONCLUSION

Republic of Croatia is currently in the process of many changes concerning higher educational system. Bologna process is being conducted and new system of scholarship is being adopted with the intention of providing equal conditions to all students. Also, there is a problem of unequal distribution of students among colleges which has influence that some sectors have surplus of labor, while others have a shortage.

In the paper we described the process of knowledge discovery from databases and demonstrated this process for a practical example of a current actual problem. Two models were successfully created which predict student success based on GPA criterion and time student needs to finish the undergraduate program (time-to-degree) criterion.

By increasing the sample size and including more colleges in the survey, it would be possible to increase performance of the models. Except for demographic factors, it would be possible to include all other factors that were recognized in earlier studies as factors that can influence student success, like time students spend preparing for exams, passing course by attending preliminary exams or regular end of term exams, lecture attendance, GPA in high school, type of high school etc. Therefore, it would be possible to create specialized models for every type of collage depending on the program characteristics of these colleges and as shown by Cristobal, et al. [23] and Hsu, Wu and Lin [24] another possibility of application is in the increasingly growing scope of e-learning. Models like this could be used as a part of decision support system to help young people chose which college to enroll, predict the volume of budget for subsidies of higher education system in the years to come, thus increasing the number of students that finish the undergraduate university program instandard time-to-degree or in context-prediction performance as shown by Lee, S., Lee, K.C. [25]. Future studies should also include mobility as a predictive factor of student performance [26].

Models like this could be used to optimize student enrolment and to increase the percentage of students that finish programs in time and with better GPA. Models can be additionally improved by introducing some other influential variables. In earlier studies [27] it has been shown that there is correlation between test anxiety of students on one side and parental attitudes toward academic studies, students' self-reported attitudes toward academic studies and parental involvement in academic studies on other side so this aspect could be further implemented into model. Another strong indicator of student success that should be implemented into model could be high school graduation score, high school GPA [28] and college admission exam score of certain subjects [29]. Apart from cognitive ability, also emotional intelligence should be accounted into prediction of academic performance [30]. There is also possibility to develop models which purpose would be to monitor student success. Similar to the study conducted by Fischbach et al. [31] who used Programme for International Student Assessment in order to predict class repetition and subject-specific grades, student exam results from earlier years of college enrolment could be used to predict next year student success and establish student monitoring models.

As third criteria of successful studying ECTS points could be taken into consideration, but only when the ECTS system of grading has been standardized and harmonized among colleges. Colleges could also take into account students' personal motivation and perception of courses [32] in order to adequately estimate ECTS points for every course.

REFERENCES

[1] A. Bertoncelj, D. Kovač, and R. Bertoncelj,"Success factors and competencies in organisational evolution," *Kybernetes*, vol. 38, no. 9, pp. 1508-1517, 2009.

[2] T. Hernaus, M. Pejić Bach, and V. Bosilj Vukšić,"Influence of strategic approach to BPM on financial and non-financial performance,"*Baltic Journal of Management*, vol 7, no. 4, pp. 376-396, 2012.

[3] M. Berry, and G. Linoff, *Mastering Data Mining: The Art and Science of Customer Relationship Management,*New York: John Wiley & Sons, 2000.

[4] S. Sumathi, S.N. Sivanandam, *Introduction to Data Mining And Its Applications*, Netherlands: Springer, 2006.

[5] B.C. Hardgrave, R.L. Wilson, and K.A. Walstrom,"Predicting graduate student success: A comparison of neural networks and traditional techniques," *Computers & Operations Research*, vol.21, no.3, pp. 249-263, 1994.

[6] I. Zaidah, and R. Daliela,"Predicting students' academic performance, comparing artifcial neural network, decision tree and linear regression,"in *21st Annual SAS Malaysia Forum Proceedings*, Kuala Lumpur, 2007, pp. 1-6.

[7] V.O. Oladokun, A.T. Adebanjo, and O.E. Charles-Owaba,"Predicting Students' Academic Performance using Artifcial Neural Network: A Case Study of an Engineering Course,"*The Pacifc Journal of Science and Technology*, vol. 9,no. 1, pp. 72-79, 2008.

[8] T. Matković, I. Tomić, and M. Vehovec," Efikasnostnasuprotdostupnosti? O povezanostitroškovaiishodastudiranja u Hrvatskoj,"*Revijazasocijalnupolitiku*, vol. 17, pp. 2215-237, 2012.

[9] M. Zekić-Sušac, A. Frajman-Jakšić, and N. Drvenkar," Neuronskemrežeistablaodlučivanjazapredviđanjeuspješnostistudiranja," *Ekonomskivjesnik*, no.2, pp. 314.-327, 2009.

[10] E.J. Shaw, J.P. Marini, and K.D. Mattern,"Exploring the Utility of Advanced Placement Participation and Performance in College Admission Decisions,"*Educational and Psychological Measurement*,vol. 73, no. 2, pp. 229-253, 2013.

[11] B. Rienties, and D. Tempelaar,"The role of cultural dimensions of international and Dutch students on academic and social integration and academic performance in the Netherlands,"*International Journal of Intercultural Relations*,vol. 37, no. 2, pp. 188-201, 2013.

[12] J. Allen, A. Gregory, and A. Mikami,"Observations of Effective Teacher-Student Interactions in Secondary School Classrooms: Predicting Student Achievement With the Classroom Assessment Scoring System-Secondary,"*School Psychology Review*,vol. 42 no. 1, pp. 76-98, 2013.

[13] B. Carnahan, G. Meyer,and L.A. Kuntz,"Comparing statistical and machine learning classifiers: Alternatives for predictive modeling in human factors research,"*Human Factors,*vol. 45 no. 3, pp. 408-423, 2003.

[14] T. J. Pleskac, J. Keeney, and S.M. Merritt,"A detection model of college withdrawal," *Organizational Behavior and Human Decision Processes,* vol. 115,no. 1, pp. 85-98, 2011.

[15] C. Marquez-Vera, A. Cano, and C. Romero,"Predicting student failure at school using genetic programming and different data mining approaches with high dimensional and imbalanced data,"*Applied Intelligence*,vol. 38 no. 3, pp. 315-330, 2013.

[16] B.C. Chiu, and G.I. Web,"Using decision trees for agent modeling: Improving prediction performance,"*User Modeling and User-Adapted Interaction*,vol. 8 no. 1-2, pp. 131-152, 1998.

[17] W. Luo, K.M. Aye, and D. Hogan,"Parenting behaviors and learning of Singapore students: The mediational role of achievement goals," *Motivation and Emotion*,vol. 37,no. 2, pp. 274-285, 2013.

[18] F. Fischer, J. Schult, and B. Hell,"Sex differences in secondary school success: Why female students perform better,"*European Journal of Psychology of Education*,vol. 28,no. 2., pp. 529-543, 2013.

[19] M.P. Levpuscek, M. Zupancic, and G. Socan, "Predicting Achievement in Mathematics in Adolescent Students: The Role of Individual and Social Factors,"*Journal of Early Adolescence,*vol. 33,no. 4, pp. 523-551, 2013.

[20] University of Zageb,"*Prvikoraci u Bolonjskomprocesu, Sveučilište u Zagrebu*," http://www.unizg.hr/fileadmin/rektorat/dokumenti/bologna/Bologna.pdf, 2005.

[21] H. Hand, H. Mannila, and P. Smyth,*Principles of Data Mining*, Cambridge: MIT Press, 2001.

[22] N. Bijayananda, and R. Srinivasan,"Using neural networks to predict MBA student success," *College Student Journal*, vol. 38, no. 1, pp. 143-150, 2004.

[23] E. Cristobal, G. Pedro, and A. Zafra, "Web usage mining for predicting final marks of students that use Moodle courses,"*Computer Applications in Engineering Education*,vol. 21,no. 1, pp. 135-146, 2013.

[24] C. Hsu, S.P.J . Wu, and B. Lin,"The Discovery of Network Structure for E-Learning Participation Prediction: An Integrated Bayesian Networks Approach," Journal of Internet Technologyvol. 14,no. 2, pp. 251-263, 2013.

[25] S. Lee, and K.C. Lee,"Context-prediction performance by a dynamic Bayesian network: Emphasis on location prediction in ubiquitous decision support environment," *Expert Systems With Applications,*vol. 39,no. 5, pp. 4908-4914, 2012.

[26] J. Bošnjović, and V. Trivun,"Academic Mobility in the Western Balkans," *Business Systems Research*, vol. 4, no. 1, pp. 76-86, 2013.

[27] E. Shadach, and O. Ganor-Miller,"The role of perceived parental over-involvement in student test anxiety,"*European Journal of Psychology of Education*, vol. 28,no. 2, pp. 585-596, 2013.

[28] M .Komarraju, A. Ramsey, and V. Rinella,"Cognitive and non-cognitive predictors of college readiness and performance: Role of academic discipline," *Learning and Individual Differences,* vol. 24, pp. 103-109, 2013.

[29] E.P. Bettinger, B.J. Evans, and D.G. Pope,"Improving College Performance and Retention the Easy Way: Unpacking the ACT Exam," *American Economic Journal-Economic Policy,*vol. 5 no. 2, pp. 26-52, 2013.

[30] M. Sanchez-Ruiz, S. Mavroveli, and J. Poullis,"Trait emotional intelligence and its links to university performance: An examination," *Personality and Individual Differences,*vol. 54,no. 5, pp. 658-662, 2013.

[31] A. Fischbach, U. Keller, and F. Preckel, "PISA proficiency scores predict educational outcomes,"*Learning and Individual Differences,*vol. 24, pp. 63-72, 2013.

[32] A. Lizzio, and K. Wilson,"First-year students' appraisal of assessment tasks: implications for efficacy, engagement and performance," *Assessment & Evaluation in Higher Education,*vol. 38,no. 4, pp. 389-406, 2013.

[33] Naik, B., andRagothaman, S. "Using Neural Networks to Predict MBA Student Success," *College Student Journal*, vol. 38, no. 1, pp. 143-149, 2004.