The Eurasia Proceedings of
Educational & Social Sciences (EPESS)
ISSN: 2587-1730

The Eurasia Proceedings of Educational & Social Sciences (EPESS), 2014

Volume 1, Pages 68-72

ICEMST 2014: International Conference on Education in Mathematics, Science & Technology

PREDICTING STUDENTS' ACADEMIC PERFORMANCE USING ARTIFICIAL NEURAL NETWORK : A CASE STUDY FROM FACULTY OF ORGANIZATIONAL SCIENCES

Sonja ISLJAMOVIC
Faculty of Organizational Sciences, University of Belgrade
sonjaisljamovic@gmail.com

Milija SUKNOVIC
Faculty of Organizational Sciences, University of Belgrade
milijas@fon.bg.ac.rs

ABSTRACT: University students' retention and performance in higher education are important issues for educational institutions, educators, and students. Educational data mining is focused on developing models and methods for exploring data collected from educational environments in order to better understand and improve educational process. Analyzing and determining patterns among indicators of academic success (study grade point average) and their correlation to students' personal, high school, admission data can present be a good foundation in process to adapt and improve the curriculum of higher education institutions, according to the students' characteristics. In this paper we use different artificial neural network algorithms in order to find the best suited model for prediction of students' success at the end of their studies. Additionally, we identified which factors had the crucial influence on overall students' success. Data were collected from the graduated students of Faculty of Organizational Sciences, University of Belgrade.

Keywords: Educational data mining, student success, artificial neural network, university education

INTRODUCTION

Academic education represents one of the key roles in the process of a country modernization and achieving competitive advantage. The process of education and development of human resources is becoming increasingly important concept, by which it can be determinant success in business and in society as a whole unit. It can be said that the new social elite is coming from academic circles with the aim to stabilize and improve the overall state of the economy and society.

In this research, artificial neural networks are used in order to make a system for early prediction of the success of students based on their success on the exams on the first year of studies. Early prediction of students' success could help teachers to identify students that have potential for advanced courses and also students who need the additional education in order to improve their knowledge. On the other side, developed system could be useful for students, so that they could estimate their future success in their studies on the basis of their learning habits, their work and grades. This could help them to identify on time whether they need an additional effort in order to achieve the success they want.

The paper has following structure: Section 2 gives a review of the state-of-the-art for prediction of students' success with artificial neural networks. In Section 3, the structure and the function of artificial neural networks, as well as the efficient methodologies for prediction are explained. In Section 4, we present the experimental results of this research. Evolved models in this research are evaluated on the real data about students and their success on the Faculty of Organizational Sciences in Belgrade. In section 5, results are discussed and the directives for further researches are given.

- Selection and peer-review under responsibility of the Organizing Committee of the conference

*Corresponding author: Sonja ISLJAMOVIC- sonjaisljamovic@gmail.com

SOLUTIONS FOR PREDICTING THE SUCCESS OF THE STUDENTS

The ability to predict the students' academic performance is very important in institution educational system, as well for faculty, university and educators, as well as for students. Data mining in the field of education (Educational Data Mining - EDM), as a new field of research, has developed in the last decade as a special area of application techniques and tools for detecting regularities and correlations in the data (data mining), with the aim of analyzing the unique data types that appear in educational system for solving various problems of educational and instructional improvement process (Romero & Ventura, 2007; Romero & Ventura, 2010). EDM is engaged in development, research and application of methods to detect regularities in the data in the database in the field of education, which would otherwise be difficult or almost impossible to analyze and determine the dependency patterns of behavior and learning among students, primarily because of the large amount of data (Romero et al., 2010).

Applications of data mining techniques in educational environments are more and more popular area of research and there is an increasing number of research papers published in the last years (Minaei-Bidgoli & Punch, 2003; Schumacher et al., 2010; Etchells et al.; 2006; Ayesha et al., 2010; Wook et al., 2009). Researches in educational data mining area focus on different aspects of educational process: students, teachers, teaching aids, organization of teaching etc. (Falakmasir & Habibi, 2010; Kumar & Chadha, 2011; Romero & Ventura, 2007; Guruler et al., 2010). One of the main goals and the basic application fields of educational data mining can be classified into the following categories: Prediction of students' success, Organization of teaching programs, Prediction of enrolment of students to the higher level of educational program, Discovering cheating during an on-line examination, Identification of abnormal/extreme values in the educational system, (Kumar & Chadha, 2011). In order to achieve these goals, different algorithms of data mining are used, such as: decision trees, artificial neural networks (ANNs), k-nearest neighbor, naive Bayes, support vector machines, cluster algorithms, (Ayesha et al., 2010).

Traditionally, academic researchers have used statistic models and methods in order to predict the success of the students. Today, there are many different approaches about classifying the students and predicting their grades Gonzalez and DesJardins (2002) tried to predict the success of the students, using the linear regression and neural network models. The results of their research point that the models of logistic regression do not predict behavior of students equally well as the artificial neural network models. Thomas and Hass (2001), compared the performance of the three different techniques of data mining for predicting students' behavior: neural networks, cluster algorithms and decision trees, where the model based on neural network gave the best results. Delgado (Delgado et al., 2006) used neural networks to predict success of the students on the exams, defined with binary classes (pass or fail). Wook (Wook et al., 2009), compared two data mining techniques, ANN and the combination of clustering and decision tree classification techniques for predicting and classifying students' academic performance, and based on that research they identified the patterns that influence or affect the student's academic performance. Guo (Guo, 2010) used neural network for establishing dynamic models for analyzing and predicting student course satisfaction.

In paper of (Stathacopoulou et al., 2007) neural networks were used to add learning and generalization abilities to the fuzzy model by encoding teachers' experience through supervised neural network learning. The study of (Wu et al., 2008) indicates that a properly trained ANN classification model can be a strong predictor for use in the learning disabilities diagnosis procedure, and furthermore, a well-trained ANN model can also be used to verify whether a learning disabilities diagnosis procedure is adequate. Artificial neural networks for prediction of the number of secondary-school pupils who will continue university studies are used in (Gerasimović et al., 2001). Isljamovic (Isljamovic et al., 2012) compared several classification algorithms in order to predict if a student would display excellent performance (i.e. highest grades) on a selected course in technology enhanced learning environment. ANNs were among the best algorithms by several evaluation criteria.

ARTIFICIAL NEURAL NETWORKS USAGE

Artificial neural networks (ANNs) are, according to their structure, function and data processing, similar to biological neural networks and represent a relatively good technique which solves the problem of classification and prediction. ANN presents a collection of mathematical models which can simulate some of the characteristics of biological neural systems and have similarities with adaptive human learning. They are made of a great number of connected neurons (processing elements), connected by their ties which contain permeable (weight) coefficients that are, according to their role, similar to synapses. The neurons are organized in three levels: input layer, one or more hidden layers and output layer. ANN process information like biological neural networks, with the possibility of remembering, learning and removing errors with high speed of getting the solution, so that the neural networks can be used for solving complex problems, such as classification and

prediction (Yeh, 1999). ANNs were successfully used in variety of disciplines for modeling complex and real problems (Liao & Wen, 2007).

In order to find the best quality model, we used six different methods for building neural network models:

- *Quick*: a method that uses rules of thumb and characteristics of data in order to choose an adequate topology of network;
- *Dynamic:* a method that automatically creates initial topology and then adds or takes away hidden nodes during training;
- *Multiple*: a method that creates several networks with different topologies, where the exact number depends on quantity of training data. These networks are trained in parallel and the model with the smallest error is represented as the final model.
- *Prune*: a method that initially creates large network (with a lot of hidden layers and nodes) and then removes the weakest units from the input and hidden layers. This method is usually slow, but very often gives better results than the other methods.
- *RBFN (radial basis function network)*: uses a technique similar to k-means algorithm of clustering in order to separate data in accordance with output variable.
- *Exhaustive prune*: a method very similar to Prune method, that initially creates a large network (with a lot of different layers and nodes), and then removes the weakest units from the input and hidden layers. Parameters for this method are adjusted so that they ensure a very detailed search of possible model's space, in order to find the best possible model. This method is the slowest, but usually gives the best results.

In this research, software package Clementine (SPSS) was used for ANN developed as one of temporarily the most efficient and most effective software solutions which enable data mining, as well as quick development of predicting model, with a high quality. Data processing in Clementine is being done by using different kinds of nodes which are connected, forming a certain flow of data over which data mining and prediction are being done.

STUDENT SUCCESS PREDICTION – CASE STUDY

Research study performed on the Faculty of Organizational Sciences, which represents one of leading faculties, is in group of faculties of technology and engineering sciences of University of Belgrade. Faculty of Organizational Sciences is a higher education institution that deals with education, scientific research and consultancy through development of knowledge and skills in management, information systems and technology with the aim to enable future professionals to develop potentials of commerce and society. This research includes the first five generations of students, with 1787 graduated student, who have been registered in accordance to the principles of higher education implemented by the use of the Bologna Declaration, which had completed their studies by November, 2013. In this time frame more than 60% of enrolled students have graduated, achieving an overall Grade point average (GPA) of 8.63 on average time of study almost 5 years.

As an input data (predictors), 15 variables were used and represent data correlated to students' personal characteristics (students' gender), high school information (high school GPA and high school type), admission data (entrance examination points) and the first year examination grades (individual grades at 11 examinations of the first year of elementary studies). On the other side, as an output from neural network, we used a GPA at the end of their studies. Ratio between the training data and the testing data was 80:20. Model evaluation is performed on so far non-used data.

Performances of developed models are measured by Absolute average error, Standard deviation and Linear correlation. Results and performance of developed ANN are presented in Table 1. The best values for each criteria of comparison are shown in bold font.

Table 1. Results of developed ANNs (6 algorithms)

Algorithm	*Absolute average error*	*Standard Deviation*	*Linear correlation*
Quick	0.233	0.295	0.893
Dynamic	0.239	0.300	0.889
Multiple	0.234	0.294	0.892
Prune	0.238	0.297	0.892
RBFN	0.259	0.324	0.873
Exhaustive Prune	**0.231**	**0.293**	**0.894**

Comparative analysis of the results, on the test sample, showed that the developed ANN gives acceptable results. An absolute average error of prediction in all the networks is from 0.231 to 0.259 and the linear correlation coefficient is always over 87%. Exhaustive Prune stands out as the best algorithm by all criteria, giving the

smallest absolute average error (0.231, the smallest standard deviation (0.293), and at the same time, the largest linear correlation coefficient (89.4%). A developed multilayer neural network with Exhaustive Prune method consists of input layer with 15 variables, output layer with one variable and two hidden layers with 30 and 20 neurons. Results with best ANN indicate that in almost 90% system will correct predict final student GPA, which represent more than good precision in prediction, especially if we know that GPA will be round on one or two decimals, then standard deviation will have even less influence to final result.

With usage of ANN give as one more advantage compared to some other data mining methods and technique, which is represented in calculation of relative importance of the input variables for predicting an output variable. In order to identify the major influenced factors of students' success, we used this advantage of ANN. Importance of the attributes of the developed ANN model are given in Table 2. It can be seen that 6 firstly ranked attributes within the developed model are considered as the most important (relative value over 0.1).

By doing further analysis, we can conclude that grades on the exam in the first semester of studies have a great influence to overall success of studies, as well as data from high school (High school type and High school GPA), where both groups of variables achieved importance over 0.1. On the other side, as it can be seen from Table 2, gender of the students and entrance examination points are not of the greatest importance for predicting student's success. Also some grades from the first year exams, especially on subjects form second semester had small influence on it, such as Management, Economy and Psychology/Sociology.

Table 2. Relative importance of the input variables

Input variable	***Relative importance***
Mathematics 1	0.1223
Basics of information-communicational technologies	0.1103
High school type	0.1072
Basics of organization	0.1034
High school GPA	0.1024
Introduction into the informational systems	0.1004
Production systems	0.0602
Mathematics 2	0.0513
English language 1	0.0502
Entrance examination points	0.0426
English language 2	0.0411
Management	0.0402
Student gender	0.0343
Economy	0.0215
Psychology/Sociology	0.0126

After selection of 6 the most important input variables, we repeated process of predicting GPA only with that 6 variables, in order to examine potential performance of prediction model with less input variables. Results and performance of developed ANN are presented in Table 3, by using same performance measurements.

Table 3. Results of developed ANNs (6 algorithms with 6 input variables)

Algorithm	*Absolute average error*	*Standard Deviation*	*Linear correlation*
Quick	0.259	0.323	0.886
Dynamic	0.262	0.330	0.881
Multiple	0.255	0.319	0.889
Prune	0.259	0.323	0.886
RBFN	0.294	0.367	0.849
Exhaustive Prune	**0.253**	**0.317**	**0.890**

Developed models with reduces number of input also gave satisfactory results. Once more, Exhaustive Prune stands out as the best algorithm by all criteria, giving the smallest absolute average error (0.253, the smallest standard deviation (0.317), and at the same time, the largest linear correlation coefficient (89%). A developed multilayer neural network with Exhaustive Prune method consists of input layer with 6 variables, output layer with one variable and two hidden layers with 24 and 18 neurons. Comparing two developed ANN with Exhaustive Prune method it can be concluded that there exists no significance, so by occasion it can be also used model developed with 6 input variables.

CONCLUSION

In this paper, model for prediction of students' success is developed in order to help professors to identify students that have potential for learning advanced courses and also students who need the additional education in

order to improve their knowledge. Using information about students after their first year of studies as input variables, a developed model of multilayer neural network has the ability to predict the success of students at the end of their studies. Development of such model gives a chance of recognizing which aspects of educational plan and program should be improved in order to induce students to work harder and improve their knowledge in certain scientific branches. This system should also be useful to students, so that they could adopt their learning habits, their work and grades, so that they could achieve the overall success they want.

Future developments and research, firstly in terms of advance use of the concepts EDM, will include integration of a large number of input variables, such as those that are directly related to studies and socio-economic and demographic indicators, their comparative analysis, and secondly in way of model development for predicting study success based on variables that were analyzed in this paper.

REFERENCE

Minaei-Bidgoli, B. Punch,*"Using Genetic Algorithms for Data Mining Optimization in an Educational Web-based System"*, Genetic and Evolutionary Computation, vol. 2, pp. 2252–2263, 2003.

Romero, S. Ventura, *"Educational data mining: a review of the state-of-the-art",* IEEE Trans. Syst. Man Cybernet. C Appl. Rev., 40(6), pp. 601–618, 2011.

Romero, S. Ventura, P.G. Espejo, C. Hervás,*Educational Data Mining 2008*, The 1st International Conference on Educational Data Mining, 2008.

C. Romero, S. Ventura, S, *"Educational data mining: A survey from 1995 to 2005",* Expert Systems with Applications, 33(1), pp. 135-146, 2007.

H. Guruler, A. Istanbullu, M. Karahasan, *"A new student performance analysing system using knowledge discovery in higher educational databases".* Computers & Education, 55(1), pp. 247-254. Elsevier Ltd. 2010.

I.C. Yeh, *Application of artificial neural network model and implementation*. Taiwan: Scholars Books, 1999.

J. Thomas, M. Hass, *Data Mining in Higher Education : University Student Declaration of Major*, Information Systems, 2011.

J.M. Gonzalez, S.L. DesJardins,*"Artificial neural networks: A new approach to predicting application behaviour"*, Research in Higher Education, 43(2), pp. 235–258, 2002.

M. Gerasimovic, L. Stanojevic, U. Bugaric, Z. Miljkovic, A. Veljovic, *"Using Artificial Neural Networks for Predictive Modeling of Graduates' Professional Choice"*, The New Educational Review, 23, pp. 175-188, 2011.

M. Wook, Y.H. Yahaya, N. Wahab, M.R.M. Isa, N.F. Awang, H.Y. Seong,*"Predicting NDUM Student's Academic Performance Using Data Mining Techniques"*, The Second International Conference on Computer and Electrical Engineering, pp. 357-361, 2009.

M.H. Falakmasir, J. Habibi, J. (n.d.), *"Using Educational Data Mining Methods to Study the Impact of Virtual Classroom in e-Learning"*, www.educationaldatamining.org, pp. 241-248, 2010. Myller, J. Suhonen, E. Sutinen, "*Using Data Mining for Improving Web-Based Course Design"*, Proc. International Conference on Computers in Education, USA, Washington, pp. 959- 964, 2002.

P. Schumacher, A. Olinsky, J. Quinn, R. Smith,*"A Comparison of Logistic Regression, Neural Networks, and Classification Trees Predicting Success of Actuarial Students"*, Journal of Education for Business, 85(5), pp. 258-263, 2010.

R. Stathacopoulou, M. Grigoriadou, M. Samarakou, D. Mitropoulos, "*Monitoring students' actions and using teachers' expertise in implementing and evaluating the neural network-based fuzzy diagnostic model"*, Expert Systems with Applications, 32(4), pp. 955-975, 2007.

S. Ayesha, T. Mustafa, A.R. Sattar, M.I. Khan,*"Data Mining Model for Higher Education System",* Europen Journal of Scientific Research, Vol.43, No.1, pp. 24-29, 2010.

S. Išljamović, M. Vukićević, M. Suknović,*"Demographic influence on students' performance - case study of University of Belgrade"*, TTEM - Technics Technologies Education Management, 7 (2), pp. 645-666, ISSN 1840-1503, 2012.

S.H. Liao, C.H. Wen, *"Artificial neural networks classification and clustering of methodologies and applications – literature analysis form 1995 to 2005"*, Expert Systems with Applications, 32(1), pp. 1–11, 2007.

T. Etchells, A. Nebot, A. Vellido, P.J. Lisboa, F. Mugica, *"Learning What is Important: Feature Selection and Rule Extraction in a Virtual Course"*, in The 14th European Symposium on Artificial Neural Networks, ESANN, Bruges, Belgium, pp. 401–406, 2006.

T.K. Wu, S.C. Huang, Y.R. Meng, *"Evaluation of ANN and SVM classifiers as predictors to the diagnosis of students with learning disabilities*", Expert Systems with Applications, 34(3), pp. 1846-1856, 2008.

V. Kumar, A. Chadha, *"An Empirical Study of the Applications of Data Mining Techniques in Higher Education",* International Journal of Advanced Computer Science and Application, 2(3), pp. 80-84, 2011.

W.W. Guo, "*Incorporating statistical and neural network approaches for student course satisfaction analysis and prediction"*, Expert Systems with Applications, 37(4), pp. 3358-3365, 2010.