

Bilal Zorić, Alisa

Conference Paper

Predicting Students' Success Using Neural Networks

Provided in Cooperation with:

IRENET - Society for Advancing Innovation and Research in Economy, Zagreb

Suggested Citation: Bilal Zorić, Alisa (2019) : Predicting Students' Success Using Neural Networks, In: Proceedings of the ENTRENOVA - ENTERPRISE RESEARCH INNOVATION CONFERENCE, Rovinj, Croatia, 12-14 September 2019, IRENET - Society for Advancing Innovation and Research in Economy, Zagreb, pp. 58-66

This Version is available at:

<https://hdl.handle.net/10419/207664>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-nc/4.0/>

Predicting Students' Success Using Neural Networks

Alisa Bilal Zorić

University of Applied Sciences Baltazar Zaprešić, Croatia

Abstract

Fast technological changes and constant growth of knowledge in many areas have led to an increasing importance of different approach to education. Efficient education is the foundation of modern society and it has the most important role in preparing students for a very flexible labour market. Education is key for development and progress. The goal of this paper is to present a model for predicting students' success using Neural networks. The model is based on students' enrolment data that consisted of demographic and economic data and information about previous education. Students' efficacy is measured by grade point average in college, and students are divided into two groups: with grade point average below and above 3.5. This model can help educators to prepare students who are classified below average with additional classes to overcome the more difficult courses and, thus, reduce the percentage of students leaving the college because of insufficient prior knowledge.

Keywords: neural networks, educational data mining, student success

JEL classification: C45, I21

Introduction

The analysis of success at studies is crucial for higher education institutions because it affects the modification of the existing programs and the creation of new ones.

As a result of the lower prices of IT equipment, educational institutions have begun to collect large amounts of data about their students. This data remains unused unless it is further analysed. Data mining techniques may be of great help. There are many research studies on the use of data mining techniques in education. Even a new discipline was born, educational data mining. It deals with the study and development of new tools, models and algorithms for a large amount of data to find useful patterns and, thus, help understand students' behaviour and how they learn. We can say that educational data mining focuses on discovering new applications of data mining for solving problems regarding education. The application of educational data mining is wide. It is most often used to improve the process of studying (Mardikyan and Badur, 2011), to improve course organization (Rashan and Peiris, 2011), to support students in course selection (Kardan et al., 2013), to find problems leading to dropping out (Bhardwaj and Pal, 2012), and as a support for decision-making at student admissions (Yadav and Pal, 2012).

In this paper, we focused on the data made available to the University while enrolling new students. Based on these data we created a model for predicting students' academic performance using Allyuda Neurointelligence. When enrolling at the university, candidates are required to submit personal documents and fill in enrolment lists. Due to general computerization, this data is entered into computers but it is not analysed any further. By developing a data mining tool like Allyuda Neurointelligence, which does not require much knowledge about the operation of neural networks, it is possible to obtain useful information from this data. We wanted

to check if there are any indicators available on the enrollment form that can predict students' success. The main goal of the paper is to demonstrate the opportunities and benefits of using educational data mining and to show their application in the specific case.

A Review of similar research

There are two main types of research regarding educational data mining, one focuses on finding attributes that affect successful study, and others are related to finding a predictive model. Many studies have shown that the most important attribute for predicting students' performance is grade point average. A detailed review of research literature on EDM between 1995 and 2005 is provided by Romero and Ventura (2007). Period after 2005 is covered in A Review and Future Visions of Educational Data Mining by Baker and Yacef (2009). A systematical literature review on predicting student' performance by using data mining techniques is provided by Shahiri and Husain (2015). The main goal of their paper was to provide an overview on the data mining techniques that have been used to predict students performance and how the prediction algorithm can be used to identify the most important attributes in a students data.

Ramaswami and Bhaskaran (2010) have developed a predictive data mining model to identify the slow learners and have studied the influence of dominant factors on their academic performance using decision tree algorithm.

Bhardwaj and Pal (2012) used the decision tree method for the classification task of evaluating student's performance. Their findings could help in identifying the dropouts and students who need special attention and it could allow the teacher to provide special attention to those students (advising/counseling).

Osmanbegovic and Suljic (2012) compared different methods and techniques of data mining for the prediction of students' success. Their goal was to develop a model which can derive the conclusion on students' academic success measured by passing grade at the exam. They investigate which variables (socio-demographic variables, achieved results from high school and from the entrance exam, attitudes towards studying) may have an affect on students' success.

Kabakchieva (2012) used various data mining classification algorithms, including a rule learner, a decision tree classifier, a neural network and a Nearest Neighbour classifier to develop a data mining models for predicting student performance, based on their personal, pre- university and university-performance characteristics. The main goal of her research was to reveal the high potential of data mining applications for university management.

Mishra et al. (2014) used different classification techniques (J48 algorithm and Random Tree) to build performance prediction model based on different attributes (students' social integration, academic integration, and various emotional skills) because an early prediction of students' failure may help the management to increase success rate and student retention.

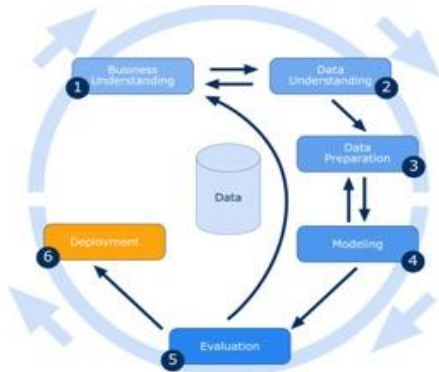
Methodology

Data mining process

Every data mining process is divided into six major steps as shown in Figure 1. First, there is Business Understanding in which business' needs and goals are transformed into a data mining problem. The second step is Data Understanding, which starts with initial data collected from available data sources. In this step, data load and data integration are done. The most important step is Data Preparation, and it usually takes

about 80% of the project time (Blake and Mangiameli, 2011). The final data set needs to be cleaned, formatted and constructed into a specific form. In the Modelling and Evaluation step, a modelling technique is selected and parameters are calibrated to optimal values. The final step is Deployment in which we present results of data mining (Oracle, 2019).

Figure 1
Data Mining Process



Source: BarnRaisers (N/A)

Data analysis

This research was done on the data collected from 76 students from the University of Applied Sciences Baltazar, Zaprešić. The table below presents the numbers of students distributed by average grade point in high school.

Table 1
Number of Students Divided by Average Grade Point in High School

| Average grade point range | Number of students | Percentage |
|---------------------------|--------------------|-------------|
| 4.5 do 5.0 | 12 | 16% |
| 4.0 do 4.5 | 31 | 41% |
| 3.5 do 4.0 | 27 | 35% |
| 3.0 do 3.5 | 6 | 8% |
| 2.5 do 3.0 | 0 | 0% |
| Total | 76 | 100% |

Source: Authors' work

We wanted to create a prediction model for students' academic performance, based on enrolment data. We designed a neural network using Alyuda NeuroIntelligence software package and we got a model by which we could determine student's academic performance: good, with average grade above 3.5, and bad, with average grade below 3.5. We used the following characteristics: status of study, gender, mother's education, father's education, average grade during high school, state matura level - Croatian language, mathematics, foreign language, state matura grade - Croatian language, mathematics, foreign language, average monthly income, accommodation during studies, scholarship and work during studies. Based on the status of studies, students were divided into two groups, full and part-time students. The variables which show mother's or father's education included these options: elementary school, secondary education, bachelor's, master's or doctor's degree. Since the academic year 2009/2010, in order to enrol in a college in Croatia, you have to pass a state matura exam. Mandatory subjects from which you have to

take state matura exams are: the Croatian language, mathematics, foreign language, and a candidate may choose the level they want to take: basic or high. Accordingly, we defined the variables of the Level of state matura. State matura grade can have standard values: pass, good, very good, excellent. We divided average monthly income into five categories: considerably below average, little below average, average, little above average and considerably above average. Variable "accommodation during studies" can have these values: with parents / relatives, payment of rent or dormitory. Possible answers to the question: "Would you work during study?" were: Yes, full time, Occasionally and No. We had problems with missing data, because in our dataset, we had students who graduated before 2009, and they did not have state matura exams.

Description and application of the chosen method

Neural network is a mathematical computing model based on the structure and functions of biological neural system. It is a nonlinear predictive model that learns through training and looks like biological neural networks in structure. (Han et al, 2011). Neuron is a basic unit of neural network. Every neuron has two parts: the activation and net function. The activation function is a mathematical formalism that is used to define the output behaviour of a neuron. This output is used as input for the next node and so on until a desired solution to the original problem is found. Activation function allows neural network learning complicated, non-linear mappings between inputs and response variables (Hinton et al., 2006). The net function determines how the network inputs are combined inside the neuron. (Schölkopf et al., 2002). Neural networks have three different layers: Input layer, Hidden Layers and Output Layer. Layers consist of nodes that are interconnected with an activation function. Patterns are presented to the network through the input layer, which communicates with one or more hidden layers where actual processing is performed through a weighted connection system. The hidden layers then connect to an output layer where the response is sent. Each input is sent to each neuron in the hidden layer and then each hidden layer's neuron's output is associated with each neuron in the next layer.

Neural network learns from processing many marked examples that are submitted during training. It uses this answer key to learn what characteristics of the input are needed to construct the correct output. After the neural network has processed a sufficient number of learning examples, it can start processing new, unrecognized samples. The larger the training set is, the more accurate the results are, because the program learns from experience.

Because of the generalized approach to problem solving, neural network has no limit to the areas of application, and its application grows every day. The most common uses of neural networks are: financial prediction, facial recognition, email spam filtering, pattern recognition, self-driving vehicle trajectory prediction, medical diagnosis, etc.

Results

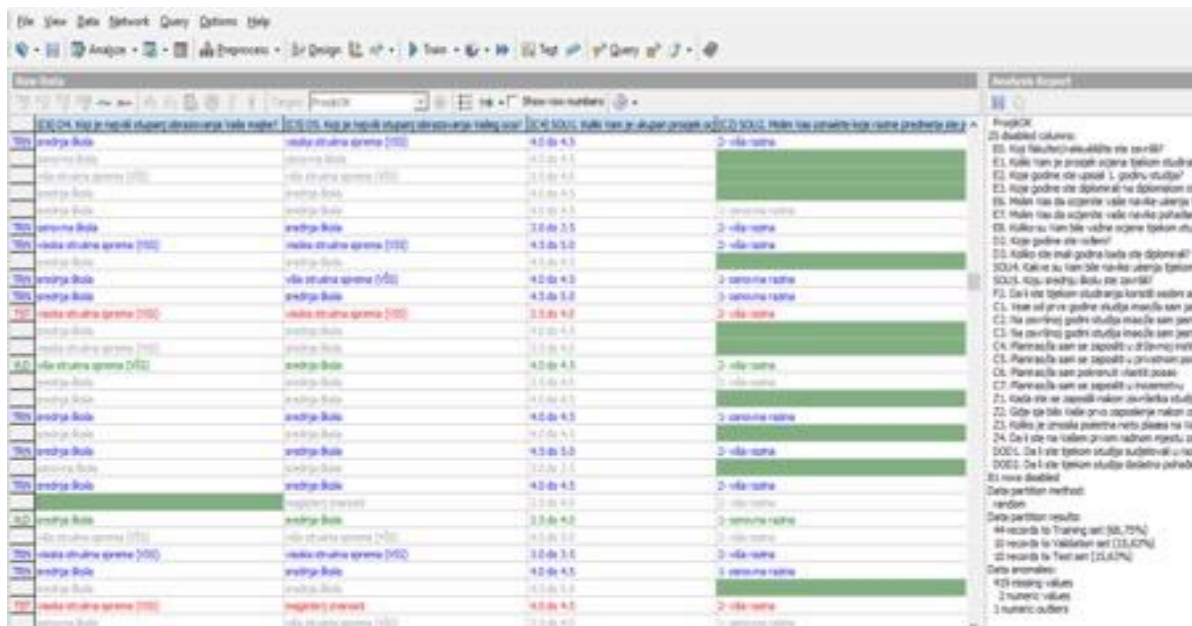
To create the model, we used Alyuda Neurointelligence software package and enrolment data acquired in the university enrolment process.

After selecting the database, the software goes through five phases: data analysis, data pre-processing, design, training and testing. In data analysis phase, we are determining the target characteristics that we want to calculate, defining characteristics we will use and characteristics which we will reject like name, surname and personal number. The target characteristic is `Projek_Ok` which has the value of

one if university average grade is higher than 3.5 and zero if university average grade is below 3.5.

Image below shows data analysis using Alyuda Neurointelligence software package. The software package randomly divides Data into three sets: training - blue (68.75 %), validation - green (15.63 %) and testing set - red (15.63 %).

Figure 2
Alyuda – Analysis

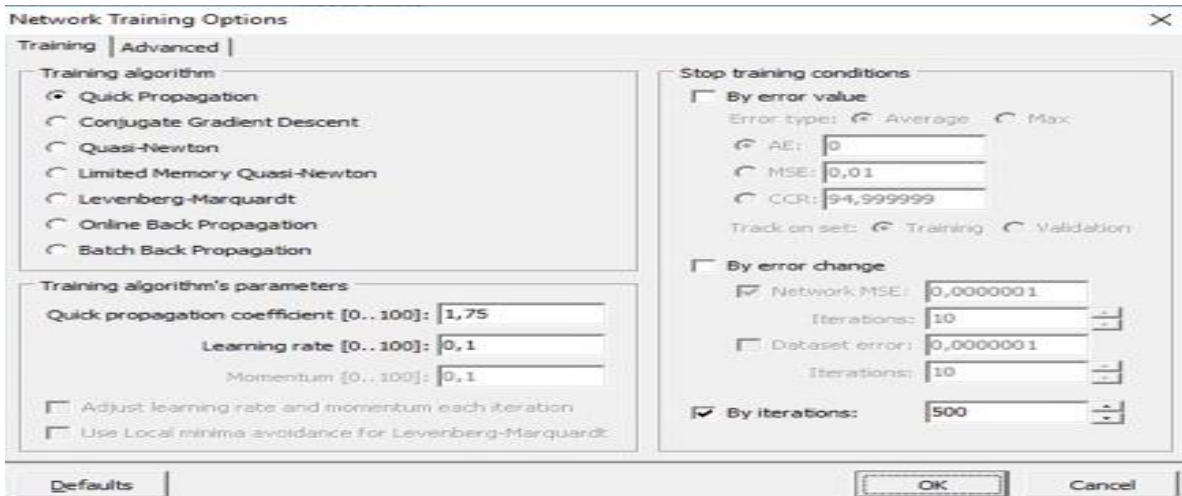


Source: Software package Alyuda Neurointelligence (Authors' work)

In pre-processing phase, the programme adds some columns if data is marked as Categorical, before pre-processing, there were 17 columns, and after, there were 54. In the network design phase, we are selecting a number of hidden layers. The program offers the best topology, which we can change. In our case, that is a neural network with one hidden layer with 28 neurons (Row Architecture in Figure 3).

After designing, there is training in which we can define different parameters like training algorithm and stop training conditions as shown in the Figure 2. The best training algorithm does not exist, the selection of algorithm depends on the characteristics of the problem we want to solve. Quick Propagation algorithm was used.

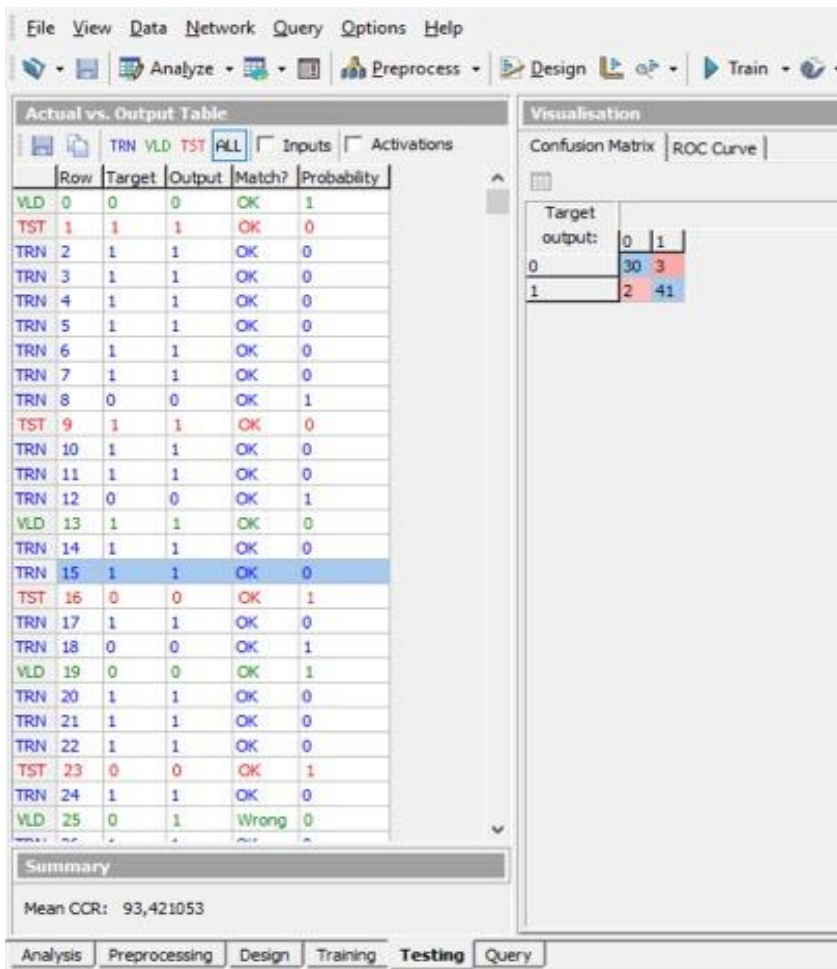
Figure 3
Alyuda – Network Training Options



Source: Software package Alyuda NeuroIntelligence (Authors' work)

After training the network, we get results as shown below.

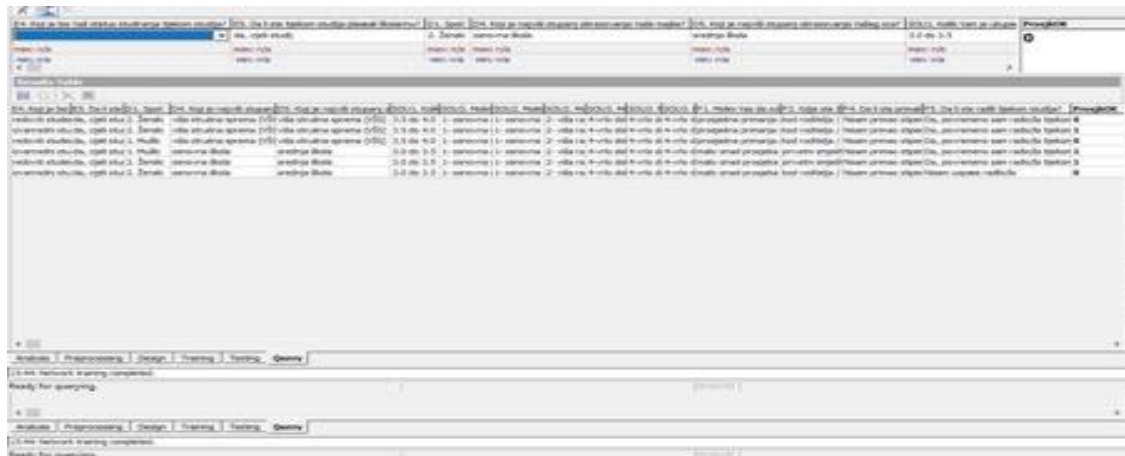
Figure 4
Alyuda – Network Training Results



Source: Software package Alyuda NeuroIntelligence (Authors' work)

In the end, we got a model using which we can check the likelihood of future students' academic performance by entering some parameters obtained during each student's enrolment process.

Figure 5
Alyuda – Query



Source: Software package Alyuda NeuroIntelligence (Authors' work)

We tried different neural networks topologies, and all gave similar results.

Discussion

Due to a large number of institutions of higher education in Croatia, the analysis of successful studies is important for attracting new students and the retention of the existing ones. If we discover the rules that trigger certain occurrences, such as rules that lead to students dropping out, in time, certain actions may be taken to prevent such occurrences. Based on performance analysis, some courses can be modified and new ones can be introduced. For example, if we notice that the majority of students in a certain course of study have problems with mastering a class, we can modify this class or divide it into two, which are easier to master.

Conclusion

The application of data mining in education is very wide and diverse. The prediction rate in our research was very high (93,42 %) due to insufficient data of lower-grade students (only 8%). The model created in this paper could be a good starting point for future researchers who could add the specifics of an institution to the model. With better data and larger quantities of it, the model could easily adapt to the needs of a particular higher education institution and, based on it, they could make decisions about student enrolment or the creation of study programmes. By analysing the data and adapting the model, as well as analysing the importance of the influence of individual variables on the performance at studies, the model could also be used as a support to decision-making in education, thus contributing to more successful studies and enhancing the quality of education in general. This research was made using a small database, applying only one method, neural networks. We could access other important information that could help universities get competitive advantage by using other methods such as association rule, segmentation, decision trees, clustering, outlier detection, etc. We wanted to show the simple usage of a complex method and to encourage others in similar research. Today, there are many great software packages

for data mining that do not require much pre-knowledge, and the results they provide are very useful.

References

1. Baker, R. S., Yacef, K. (2009), "The State of Educational Data Mining in 2009: A Review and Future Visions", *Journal of Educational Data Mining*, Vol. 1 No. 1, pp. 3-17.
2. BarnRaisers (N/A), available at: https://barnraisersllc.com/wp-content/uploads/2018/09/CRISP_en.png (03 April 2019).
3. Bhardwaj, B. K., Pal, S. (2012), "Data Mining: A prediction for performance improvement using classification", *International Journal of Computer Science and Information Security*, Vol. 9, No. 4, pp. 136-140.
4. Blake, R., Mangiameli, P. (2011), "The effects and interactions of data quality and problem complexity on classification", *Journal of Data and Information Quality (JDIQ)*, Vol. 2, No. 2.
5. Han, J., Pie, J., Kamber, M. (2011), *Data Mining Concepts and Techniques*, Elsevier, Burlington.
6. Hinton, G. E., Osindero, S., Teh, Y. W. (2006), "A fast learning algorithm for deep belief nets", *Neural computation*, Vol. 18, No. 7, pp. 1527-1554.
7. Kabakchieva, D. (2012), "Student performance prediction by using data mining classification algorithms", *International Journal of Computer Science and Management Research*. Vol. 1, No. 4, pp. 686-690.
8. Kardan, A. A., Sadeghi, H., Ghidary, S. S., Sani, M. R. F. (2013), "Prediction of student course selection in online higher education institutes using neural network", *Computers & Education*, Vol. 65, pp. 1-11.
9. Mardikyan, S., Badur, B. (2011), "Analyzing Teaching Performance of Instructors Using Data Mining Techniques", *Informatics in Education*, Vol. 10, No. 2, pp. 245-257.
10. Mishra, T., Kumar, D., Gupta, S. (2014, February), "Mining students' data for prediction performance", in *Proceedings of 2014 Fourth International Conference on Advanced Computing & Communication Technologies*, Rohtak, India, IEEE, pp. 255-262.
11. Oracle (2019), "Data Mining Concepts", available at: http://docs.oracle.com/cd/B28359_01/datamine.111/b28129/process.htm#DMCON002 (03 April 2019).
12. Osmanbegovic, E., Suljic, M. (2012), "Data mining approach for predicting student performance", *Economic Review: Journal of Economics and Business*, Vol. 10, No. 1, pp. 3-12.
13. Ramaswami, M., Bhaskaran, R. (2010), "A CHAID based performance prediction model in educational data mining", *International Journal of Computer Science Issues IJCSI*, Vol. 7, No. 1.
14. Roshan, K. H., Peiris, A. (2011), *Data Mining Applications in the education sector*, MSIT, Carnegie Mellon University.
15. Romero, C., Ventura, S. (2007), "Educational Data Mining: A Survey from 1995 to 2005", *Expert Systems with Applications*, Vol. 33, No. 1, pp. 135-146.
16. Schölkopf, B., Smola, A. J., Bach, F. (2002), *Learning with kernels: support vector machines, regularization, optimization, and beyond*, MIT press.
17. Shahiri, A. M., Husain, W. (2015), "A review on predicting student's performance using data mining techniques", *Procedia Computer Science*, Vol. 72, pp. 414-422.
18. Yadav, S. K., Pal, S. (2012), "Data mining: A prediction for performance improvement of engineering students using classification", *World of Computer Science and Information Technology Journal WCSIT*, Vol. 2, No. 2, pp. 51-56.

About the authors

Alisa Bilal Zorić, Ph.D. is a Lecturer at the University of Applied Sciences Baltazar Zaprešić. She received PhD in Business Science at the Faculty of Commercial and Business Sciences Celje with the dissertation thesis "Determinants of Efficacy of Studying in the Republic Croatia: Comparison of Neural Networks and Decision Trees". Alisa spent ten years designing and developing business information systems for Siemens and Banksoft. Her main research interests are information and communication technology and educational data mining. She published several scientific papers in international and national journals and participated in many scientific international conferences. The author can be contacted at abilal@bak.hr.