# Predicting student satisfaction with courses based on log data from a virtual learning environment – a neural network and classification tree model

## Ivana Đurđević Babić[1,*]

[1] *Faculty of Education, Josip Juraj Strossmayer University of Osijek*
*Cara Hadrijana 10, 31000 Osijek, Croatia*
*E-mail:* ⟨ *idjurdjevic@foozos.hr* ⟩

**Abstract.** Student satisfaction with courses in academic institutions is an important issue and is recognized as a form of support in ensuring effective and quality education, as well as enhancing student course experience. This paper investigates whether there is a connection between student satisfaction with courses and log data on student courses in a virtual learning environment. Furthermore, it explores whether a successful classification model for predicting student satisfaction with course can be developed based on course log data and compares the results obtained from implemented methods. The research was conducted at the Faculty of Education in Osijek and included analysis of log data and course satisfaction on a sample of third and fourth year students. Multilayer Perceptron (MLP) with different activation functions and Radial Basis Function (RBF) neural networks as well as classification tree models were developed, trained and tested in order to classify students into one of two categories of course satisfaction. Type I and type II errors, and input variable importance were used for model comparison and classification accuracy.

The results indicate that a successful classification model using tested methods can be created. The MLP model provides the highest average classification accuracy and the lowest preference in misclassification of students with a low level of course satisfaction, although a t-test for the difference in proportions showed that the difference in performance between the compared models is not statistically significant. Student involvement in forum discussions is recognized as a valuable predictor of student satisfaction with courses in all observed models.

**Key words:** classification, neural networks, classification tree, course satisfaction, log data

## 1. Introduction

Academic educational institutions place great emphasis on the quality of education, which is essentially of a national public interest. In order to ensure

---

[*] Corresponding author

this, most university teaching staff gather information on student satisfaction with courses through various student questionnaires. This is usually done at the end of the course to obtain information that will serve as preparation for the next generation of students. Recognizing the unsatisfied as opposed to satisfied students on time could help teachers react more quickly and create a stimulating environment in their course that would enhance student satisfaction with the and eventually change a student's overall experience of the course.

There is a significant absence of literature dealing with the issue of student satisfaction. The reason being that this problem is primarily observed from a customer-orientated perspective, where students are identified as customers of academic institutions and research mainly focuses on satisfaction with student support services, tuition and accommodation or environmental factors. Very little research has observed this problem from a different aspect. Research generally concentrates on discovering correlations among student satisfaction and different personal psychological elements or the final grades and academic success of students.

By gaining new insights from student LMS course log data covering the issue of student satisfaction with courses, this research seeks to contribute to filling in the gap in literature on student satisfaction with courses.

The aim of this paper is to explore whether a successful classification model can be developed in order to predict student satisfaction with courses based on course log data. Classification decision trees, Multilayer Perceptron (MLP) and Radial Basis Function (RBF) neural network models are used for this purpose. NN models are used in this research due to their ability to define complex relationships between variables [5], high noise tolerance and ability to generalize results [33] as well as their function approximation capability [43]. Some researchers emphasize that the classification tree and neural network models do not require presuming statistical distributions of data while in traditional parametric statistical methods these assumptions are required [3]. In addition, when comparing the NN method with traditional statistics methods, most researchers have concluded that NN methods outperformed the traditional methods or produce results at least on par with traditional statistical methods (see for instance [12, 32, 30, 26]). Furthermore, in drawing a comparison between classification trees and traditional statistical methods, Fernandez [10] has noticed that classification trees are the preferred exploratory technique. He has specified their advantages over traditional statistical methods such as their ability to provide a simple rule for classification or to deal with large numbers of input variables, their straightforward approach as well as the already mentioned absence of requiring statistical data distributions.

The following section gives a brief overview of previous research. Section 3 describes the sample used in this research and methodology of neural networks

(NN) and classification tree models. The results of the NN and classification tree models with the highest average accuracy rate are presented and compared in Section 4, which is then followed by a discussion.

## 2. Review of previous research

Much of the research deals with a wide range of various issues concerning student satisfaction. For instance, some researchers endeavor to find out which factors affect student satisfaction (see, for instance [29, 41, 31]). Others explore the relationship between student satisfaction and other variables such as service quality (see, for instance [46, 44, 17]), instructional variables (see, for instance [20]) or instructional design and management style (see, for instance [23]). Still, others address topics on student satisfaction with respect to the learning environment (see, for instance [40]) or student satisfaction and lecture notes (see, for instance [27]).

Although a large number of research deals with the topic of student satisfaction, only a small number apply neural networks or other data mining techniques. Taking into consideration 9 variables (unsatisfactory rate, neutral rate, satisfactory rate, number of enrolled students, high distinction, credit, pass, fail), Guo [16] has used linear regression and MLP neural network in predicting student satisfaction with courses. He has concluded that MLP models outperform linear regression models predicting student satisfaction with courses and recommends the use of regression analysis as a means of identifying and excluding redundant variables from modelling. Dejaeger et al [9] have conducted research with the goal of identifying student satisfaction factors at two educational institutions in order to construct comprehensible data mining model for management at these institutions. They collected data using questionnaires that recorded reaction by students and used support vector machines, neural networks, decision trees and logistic regression for modelling. Although they obtained better results with logistic regression, the management of institutions found the decision tree model better and easier to understand.

Student behaviour in virtual learning environments (VLE) or learning management system (LMS) is automatically collected in the form of log files. These log files provide a source of data for which data mining techniques can be used to gain a considerable amount of useful information [36]. In order to confirm that pedagogically important information can be obtained from log files, Macfayden and Dawson [28] have analyzed log files of one particular LMS. In order to predict students' final grades in the course based on their global activity in that course, they used regression and identified 15 variables that show correlation with a student's final grade.

Thomas and Galambos [42] argue that the utilization of data mining techniques in higher education research has been insufficient. Sharma and Singh [38] emphasize important issues currently facing educational system (high dropout rates, personalization of teaching, etc.) and recommend data mining techniques as the means to overcome them by extracting information that will help understand students and their requirements. In recent years, several studies have focused on examining online student interaction and log files. He [19] applied educational data mining techniques and text mining techniques to explore utilization of online questions and chat messages in a live video streaming of a course. The research revealed a positive correlation between students' final grade and the number of student questions sent to an educator. Gobert et al [15] have used educational data mining on student log data in order to assess middle school student inquiry skills for designing and conducting experiments in a physical science simulation. García-Saiz and Zorrilla Pantaleón [13] have conducted research in which they have presented an application called E-learning Web Miner to ease discovering student behaviour in LMS courses.

However, no record of analyzing log data on student LMS courses for predicting student satisfaction with courses could be found. Therefore, this study opens up a new perspective to conduct research into this field by investigating whether it is possible to predict student satisfaction with courses based solely on general activity by students in LMS courses.

## 3. Sample and research methodology

In the 2013/2014 academic year, 154 students who had enrolled into third and the fourth year at the Faculty of Education in Osijek participated in this research. Input variables for models were collected from LMS Moodle courses log files. Only those general actions performed by students from both third and fourth year were taken into account as input variables for the models. Therefore, twelve input variables were used (V1 - assignment upload, V2 - assignment view, V3 - course view, V4 - forum discussion, V5 - forum view, V6 - page view, V7 - questionnaire submit, V8 - questionnaire view, V9 - resource view, V10 - url view, V11 - user view and V12 - user view all). A student questionnaire with nine statements was used to determine the level of student satisfaction with courses. The statements used in the questionnaire to establish student satisfaction with courses were obtained from an internal student survey at the Faculty of Education in Osijek [48] with permission from the Faculty Board and the author of the internal student survey. The students were asked to assess their level of agreement with these statements on a 5-point Likert scale (1 – strongly disagree, 5 – strongly agree). An overall measure of student satisfaction with courses was achieved by summing the scores for all 9

statements so that the higher score represents a higher level of student satisfaction with a course. The final scores of the student satisfaction with courses were divided into two categories: 1 representing the students with a higher level of course satisfaction who provided final scores ranging from 41 to 45, and 0 representing those students with a lower level of course satisfaction. All NN and classification trees models in this research had this binary value of student satisfaction with courses as the output variable. An average classification rate of both categories was used to assess model successfulness. The data were divided into a training sample and a test sample (see Table 1), which was used for evaluating model accuracy in a way so that the confusion matrix with a correct and incorrect hit rate for both categories 0 and 1 was computed. The training sample in NN models consisted of a train and selection subsample.

| Sample | Category 1 | | Category 0 | | Total | |
|---|---|---|---|---|---|---|
| | No. | % | No. | % | No. | % |
| Training (train and selection) | 62 | 40.25 | 62 | 40.25 | 124 | 80.52 |
| Test | 15 | 9.74 | 15 | 9.74 | 30 | 19.48 |

**Table 1**: *Sample distribution.*

## 3.1. MLP neural networks

Because of their ability to recognize hidden patterns between variables, neural networks are widely used in different areas [47]. Especially popular are feed-forward neural networks where the multilayer perceptron (MLP) is the most commonly used [18]. MLP comprises an input layer, one or more hidden layers and the output layer, where every unit in every layer is fully connected with weighted connections and only with the units in the next adjacent layer [7, 21]. As explained in [21], every unit in the input layer is multiplied by interconnection weights of the nodes. For the $x_i$ activation of the $i$th hidden node and $w_{ij}$ weight of the interconnection among the $j$th neurons in one layer and the $i$th neuron in the previous layer, output $h_j$ given by is:

$$h_j = f\left(\sum_i x_i w_{ij} + b\right) \tag{1}$$

where $f$ is an activation function and $b$ the bias for the neuron [21]. Upon commencing training, every weight is set to be an arbitrary small number [8]. Training data in the training process are presented to a network several times

and the obtained network output may differ from the target one. The difference between target and achieved output is a local error, which is used for adjusting the weights in the network and reducing the overall error of the MLP neural network [14]. The weight adjustment is repeated several times until a suitable output is gained. A backpropagation algorithm is generally used for training and error minimization of MLP neural networks, although several other algorithms such as Broyden-Fletcher-Goldfarb-Shanno (BFGS) and the conjugate gradient could also be used. BFGS is one of the most recommended techniques for training neural networks [39] and according to [18] is the best and most popular quasi-Newton algorithm. A summarized algorithm and example of BFGS can be found in [2].

Statistica 12 software with implemented training algorithms for MLP neural networks (backpropagation, BFGS (Broyden-Fletcher-Goldfarb-Shanno) and conjugate gradient algorithms) and various activation functions (sigmoid, tangent hyperbolic, exponential, linear) was used in this research. The number of hidden units altered from 4 to 25 and the number of training cycles was set to 200.

## 3.2. RBF neural networks

In [11] major differences between MLP and RBF neural networks are emphasised. Although they both are feed-forward neural networks, the RBF neural network has only one hidden layer where every unit has a radial basis function, usually Gaussian, with two essential parameters representing location of its center and width [11]. A unit in a hidden layer measures the distance between an input data vector and the center of its radial basis function [11]. An advantage of using RBF neural networks is the simplicity of architecture and the training method as mentioned in [1].

As explained in [6], learning in Radial Basis Function neural networks is performed firstly through adjusting hidden neuron function parameters and then connection weights and network thresholds. A greater distance between the input data and the center causes decrement of output in hidden nodes [4]. Approximating that distance can be performed by using diverse functions. However, most frequently used is Euclidean distance $d(x) = ||x - c_i||$ where $d(x)$ is the distance between vector $x$ and $c_i$. As described in [22] the output of the RBF neural network is obtained in such a manner that the output of the hidden layer, the radial basis function, is passed to the output layer where all the inputs from the hidden layer are summed. In [22] that is represented by the general equation:

$$y = bias + \sum_{i=1}^{n} w_i \varphi(d_i) \qquad (2)$$

where $w_i$ is the weight of the edge $i$, $bias$ is an independent weight value added to the output, $\varphi$ is the radial basis function and $d_i$ the distance between the input and the center of a radial function in a hidden layer.

Radial basis functions are usually presented as given in [6]:

$$\varphi_i(x) = \varphi \left( \frac{\|x - \rho_i\|}{d_i} \right) \text{ for } i = 1, \ldots, m \qquad (3)$$

where the center $i$ of the function $\varphi_i$ is $\rho_i = (\rho_{i1}, \ldots, \rho_{in}) \in R^n$, $d_i$ is the standard deviation and $\|x - \rho_i\|$ the distance between the center and input.

RBF neural networks in this research used the radial basis function training (RBFT) algorithm with Gaussian-based functions. Error functions used in training the networks were the sum of squares (SOS) or cross entropy (CE). The sum of squares error function is given by the sum of differences between the target and prediction output [39]

$$E_{SOS} = \sum_{i=1}^{N} (y_i - t_i)^2 \qquad (4)$$

where the cross entropy function is given by

$$E_{CE} = -\sum_{i=1}^{N} t_i ln \left( \frac{y_i}{t_i} \right) \qquad (5)$$

where $N$ is the number of training cases, $y_i$ is the outputs of the target value $t_i$ and target values of the $i^{th}$ data case [39].

## 3.3. Classification tree

A classification tree is used for classifying objects to certain established classes [35] and is suited for solving classification problems with a small number of categories [37]. It comprises root nodes, internal nodes and leaf nodes. Each root node and internal node represents a test on attribute of cases, while each leaf node represents an outcome of classification [45]. Learning is performed so that data are divided into a train set and a test set before the generation algorithm takes train data as input [45]. Many algorithms can be used for building a decision tree and their goal is to select a suitable splitting attribute [34]. Among

the most popular are discriminant-based univariate splits algorithms built on the QUEST (Quick, Unbiased, Efficient Statistical Trees) and classification and regression trees (CART) algorithm. Each split divides data into two categories and splitting continues repeatedly until a large tree is built, which is then pruned to a right-sized tree. As explained in [25], QUEST chooses the variable with the smallest significance as a split variable and then selects its split point. In order to do so, Statistica 12 uses the 2-means clustering algorithm of Hartigan and Wong, singular value decomposition methods and recursive quadratic discriminant analysis techniques [25, 40]. Additional information about the QUEST algorithm can be found in [24].

Discriminant-based univariate splits or CART style exhaustive search for univariate splits were used as a split section method in this research. Equal prior probabilities and equal misclassification costs were specified. As a stopping rule, pruning based on misclassification error (ME) with determined stopping parameters (minimum n=5 and standard rule error=1) or a FACT-style direct stopping rule (determined fraction of objects=0.1) were set and the 10-fold cross validation was used. The *p*-value for a split variable selection was set to 0.05. Evaluating decision tree accuracy relied on comparing predicted target values with tested data values.

## 4. Results

All models were constructed using all available variables as input variables. As shown in Table 2, the model with the highest average classification accuracy is the MLP neural network model. Its average hit rate, derived using the hyperbolic tangent function as an activation function, was 73.33 %. This model consists of 4 neurons in a hidden layer. The BFGS training algorithm with entropy in the form of an error function was used to train this MLP neural network model.

The hit rate of the most appropriate RBF neural network model was 70%. The model used an RBFT training algorithm with the sum of squares as an error function and the Gaussian radial basis function. The obtained results represented in Table 2 indicate that the RBF model has the highest efficiency in accurately classifying students expressing a high level of course satisfaction (80%), but the best MLP model has the highest ability to correctly recognize students expressing a low level of course satisfaction (93%).

The classification tree model with 12 splits (see Figure 1) had an average classification accuracy of 60%. It used the CART algorithm and FACT-style stopping rule with a fraction of objects set to 0.05. The Gini measure in this model was used as a goodness of fit.
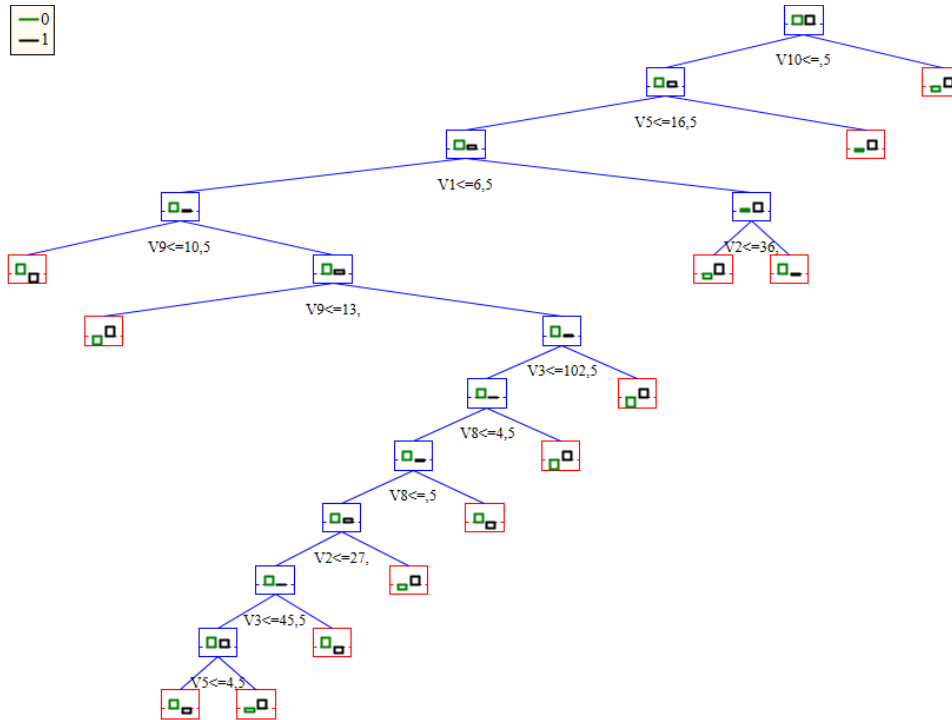
**Figure 1**: *Tree structure of CART model with 12 splits*

| Model architecture | Learning algorithm | Error function | Activation function | Accuracy for 1 (%) | Accuracy for 0 (%) | Average accuracy (%) |
|---|---|---|---|---|---|---|
| MLP 12-6-2 | BFGS | CE | linear | 33.33 | 93.33 | 63.33 |
| MLP 12-5-2 | BFGS | CE | logistic | 73.33 | 66.67 | 70.00 |
| MLP 12-5-2 | BFGS | CE | sigmoid | 46.67 | 80.00 | 63.33 |
| MLP 12-4-2 | BFGS | CE | tangent hyperbolic | 53.33 | 93.33 | 73.33 |
| MLP 12-12-2 | BFGS | CE | exponential | 33.33 | 100 | 66.66 |
| Model architecture | Learning algorithm | Error function | Radial basis function | Accuracy for 1 (%) | Accuracy for 0 (%) | Average accuracy (%) |
| RBF 12-21-2 | RBFT | SOS | Gaussian | 80.00 | 60.00 | 70.00 |

**Table 2**: *Results of the NN models obtained on a test sample.*

| Split method | Stopping rule | Accuracy for 1 (%) | Accuracy for 0 (%) | Average accuracy (%) |
|---|---|---|---|---|
| Discriminant-based univariate splits | ME | 6.67 | 86.67 | 46.67 |
| CART | FACT-style | 60.00 | 60.00 | 60.00 |

**Table 3**: *Results of the classification tree models obtained on a test sample.*

As is evident in Tables 2 and 3, the results obtained from applying the most successful classification tree models are lower than the results obtained from the most successful MLP and RBF neural network model. The accuracy of the CART, MLP and RBF models was conducted by applying the t-test on differences in proportions. The results of the t-test showed that the difference in these models was not statistically significant for a level of significance of 0.05. The obtained results are shown in Table 4.

| Hypothesis | Model | Average classification accuracy (%) | T-test p-value results |
|---|---|---|---|
| $H_0{:}p_{MLP}{=}p_{RBF}$ | MLP | 73.33 | p=0.3874 |
| | RBF | 70.00 | |
| $H_0{:}p_{MLP}{=}p_{CART}$ | MLP | 73.33 | p=0.1367 |
| | CART | 60.00 | |
| $H_0{:}p_{RBF}{=}p_{CART}$ | RBF | 70.00 | p=0.2084 |
| | CART | 60.00 | |

**Table 4**: *p-values from the conducted t-test.*

Type I and type II errors were also observed for the purpose of revealing effectiveness of the models. In comparison to the other two models, the MLP model tends to misclassify more students with a high level of course satisfaction into the category of students with a lower level of course satisfaction. This model has the highest type I error (0.47) among the observed models. The same model shows the lowest tendency in wrong classifications of students with a lower level of course satisfaction (type II error=0.07). RBF and CART models show the same preference for misclassification of students with a low level of course satisfaction. These two models have an equal type II error (0.40), but the RBF model has the smallest type I error (0.20), meaning that it has the lowest tendency of misclassification of students with a high level of course satisfaction in the observed models. In the CART model, there is no difference between type I and type II errors.

A sensitivity analysis was conducted for estimating models input variables importance. Variable ranking results in the models are presented in Table 5.

| Model | Input variables | | | | | | | | | | | |
|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|------|------|
|       | V1  | V2  | V3  | V4  | V5  | V6  | V7  | V8  | V9  | V10  | V11  | V12  |
| MLP   | 8   | 11  | 6   | 2   | 10  | 4   | 1   | 3   | 7   | 12   | 5    | 9    |
| RBF   | 2   | 12  | 10  | 4   | 1   | 5   | 11  | 9   | 3   | 7    | 6    | 8    |
| CART  | 6   | 1   | 2   | 4   | 8   | 9   | 10  | 5   | 3   | 11   | 7    | 12   |

Table 5: *Importance variable rank (1 - highest importance, 12 - lowest importance).*

Although all models had a completely different variable ranking, student engagement that provided opinions whether through questionnaire responses or through interactive correspondence on forums was recognized as an important indicator in detecting students with a high level of course satisfaction in all observed models. In the MLP model, variables that describe this active type of course communication are the most influential. Variables that represent a slightly passive approach to LMS learning, such as solely viewing course content data or given assignments, are identified as dominant in the CART model. However, considering the fact that the NN models were more successful in detecting students with a high level of course satisfaction, the conclusion is that student involvement in course forum communication is highly significant in detecting students with a high level of course satisfaction.

Variables that are not associated with this type of course communication and obligatory course learning materials, such as access to a participant's profile from a course in LMS, or access to additional learning materials available through external url, do not have such a powerful impact in detecting highly satisfied students in any of the models.

## 5. Conclusion

The aim of this research was to investigate whether a model to classify students according to their level of course satisfaction could be created based on student course log data. MLP neural networks, RBF neural networks and classification tree models were used. These techniques provided models with overall classification accuracy higher than 50%. Therefore, it has been shown a successful classification model for predicting student satisfaction with courses could be created based on their course log files. Furthermore, based on the obtained results, the conclusion is that there is a connection between student activities in an LMS course and their level of course satisfaction. Sensitivity analysis of input variables revealed student engagement in forum discussions as one of the basic activity predictors of student satisfaction with courses in all tested models although individually, the models had different variable rankings.

The results indicate that the MLP neural network model produced the highest average classification accuracy (73.33%). This model also showed the lowest preference for misclassification of students with a low level of course satisfaction (the lowest type II error) in the group with students expressing a high level of course satisfaction. The results obtained by the RBF model indicate that this model was the most successful in detecting students with a high level of course satisfaction, but its overall average classification accuracy was lower than in the MLP model. From an educational point of view, greater importance is given to not misclassifying students with a low level of course satisfaction than detecting those with a high level of course satisfaction. The results show that the best NN models perform better than the best classification tree model; although, the results of the t-test did indicate that there is no statistically significant difference on the level of significance of 0.05 in models performance.

The results obtained in this research have revealed that log files from student course activities can be used for predicting student satisfaction with courses. Teachers can benefit from using the research findings as a qualitative independent guiding principle in detecting students with a lower level of course satisfaction in order to ensure effective use of their teaching skills. This research can also contribute to the development of a more enhanced model, although the results obtained by tested methods showed acceptable overall classification accuracy.

The limitations of this research depend on the number of input variables including sample selection and distribution. Therefore, suggestions for future research mostly relate to increasing the sample and its selection, particularly by including students from other faculties in order to ensure generalization of the obtained results. Future research should also focus on improving model accuracy and reviewing additional parameters as input variables. Furthermore, given that the tested methods in this research do not significantly differ in their performance, it would be advisable to include other classifiers into the modelling method (such as support vector machines) and to consider a combined method approach.

## Acknowledgement

# References

[1] Alejo, R., Garcia, V., Sotoca, J. M., Mollineda, R. A. and Sánchez, J. S. (2006). Improving the classification accuracy of RBF and MLP neural networks trained with imbalanced samples. In: Intelligent Data Engineering and Automated Learning–IDEAL, 2006, (pp. 464–471). Springer Berlin Heidelberg. doi:10.1007/11875581_56.

[2] Arora, J. (2004). Introduction to Optimum Design. Academic Press.

[3] Ballı, S., and Tarımer, İ. (2013). An application of artificial neural networks for prediction and comparison with statistical methods. Electronics and Electrical Engineering, 19(2), 101–105. doi:10.5755/j01.eee.19.2.3478.

[4] Bouchoux, S., Brost, V., Yang, F., Grapin, J. C. and Paindavoine, M. (2002). Implementation of pattern recognition algorithm based on RBF neural network. In: International Symposium on Optical Science and Technology, (pp. 125–135). International Society for Optics and Photonics.

[5] Cardon, H. R. A., & van Hoogstraten, R. (1995). Key issues for successful industrial neural-network applications: An application in geology. In: Artificial Neural Networks (pp. 235–245). Springer Berlin Heidelberg. doi:10.1007/bfb0027032.

[6] Carrión, J. R. and Gamo, D. M., Artificial neural network. http://www.lia.upm.es/uploads/RNA2008/ArtificialNeuralNetworks.pdf [Accessed on 15 June 2014].

[7] Chen, S. and Billings, S. A. (1992). Neural networks for nonlinear dynamic system modelling and identification. International Journal of Control, 56(2), 319–346. doi:10.1080/00207179208934317.

[8] Dayhoff, J. E. and DeLeo, J. M. (2001). Artificial neural networks. Cancer, 91(S8), 1615–1635.

[9] Dejaeger, K., Goethals, F., Giangreco, A., Mola, L., and Baesens, B. (2012). Gaining insight into student satisfaction using comprehensible data mining techniques. European Journal of Operational Research, 218(2), 548–562. doi:10.1016/j.ejor.2011.11.022.

[10] Fernandez, G. (2010). Statistical Data Mining Using SAS Applications. CRC Press, Inc..

[11] Foody, G. M. (2001), Thematic mapping from remotely sensed data with neural networks: MLP, RBF and PNN based approaches. Journal of Geographical Systems, 3(3), 217–232. doi:10.1007/pl00011477.

[12] Gan, C., Clemes, M., and Weng, A. (2005). Consumer choice prediction: Artificial neural networks versus logistic models. Journal of Social Sciences, 1(4), 211–219. doi:10.3844/jssp.2005.211.219.

[13] García-Saiz, D., and Zorrilla, M. (2011). E-learning web miner: A data mining application to help instructors involved in virtual courses. Proceedings of the 4th International Conference on Educational Data Mining, July 6–8, Eindhoven, The Netherlands.

[14] Gardner, M. W., and Dorling, S. R. (1998). Artificial neural networks (the multilayer perceptron) – A review of applications in the atmospheric sciences. Atmospheric Environment, 32(14-15), 2627–2636. doi:10.1016/s1352-2310(97)00447-0.

[15] Gobert, J. D., Sao Pedro, M., Raziuddin, J., and Baker, R. S. (2013). From log files to assessment metrics: Measuring students' science inquiry skills using educational data mining. Journal of the Learning Sciences, 22(4), 521–563. doi:10.1080/10508406.2013.837391.

[16] Guo, W. W. (2010). Incorporating statistical and neural network approaches for student course satisfaction analysis and prediction. Expert Systems with Applications, 37(4), 3358–3365. doi:10.1016/j.eswa.2009.10.014.

[17] Hanaysha, J., Abdullah, H. H. and Warokka, A. (2011). Service quality and students' satisfaction at higher learning institutions: The competing dimensions of Malaysian Universities' competitiveness. Journal of Southeast Asian Research, 1, 1–9. doi:10.5171/2011.855931.

[18] Haykin, S. (1999). Neural Networks: A Comprehensive Foundation 2nd Edition. Upper Saddle River, NJ, the US: Prentice Hall.

[19] He, W. (2013). Examining students' online interaction in a live video streaming environment using data mining and text mining. Computers in Human Behavior, 90–102. doi:10.1016/j.chb.2012.07.020.

[20] Hong, K.-S. (2002). Relationships between students' and instructional variables with satisfaction and learning from a Web-based course. Internet and Higher Education, 5, 267–281. doi:10.1016/s1096-7516(02)00105-7.

[21] Kasgari, A. A., Divsalar, M., Javid, M. R., and Ebrahimian, S. J. (2013). Prediction of bankruptcy Iranian corporations through artificial neural network and Probit-based analyses. Neural Computing and Applications, 23(3-4), 927–936. doi:10.1007/s00521-012-1017-z.

[22] Kia, A. N., Fathian, M., and Gholamian, M. R. (2012). Using MLP and RBF neural networks to improve the prediction of exchange rate time series with ARIMA. International Journal of Information and Electronics Engineering, 2(4), 543–546. doi:10.7763/ijiee.2012.v2.157.

[23] Lee, H.-Y. and Rha, I. (2009). Influence of structure and interaction on student achievement and satisfaction in Web-based distance learning. Educational Technology & Society, 12(4), 372–382.

[24] Loh, W.-L. and Shih, Y.-S. (1997). Split selection methods for classication trees. Statistica Sinica, 7(4), 815–840.

[25] Loh, W.-Y. (2008), Classification and regression tree methods. Encyclopedia of Statistics in Quality and Reliability, Ruggeri, Kenett and Faltin (eds.), Wiley, 315–323. doi:10.1002/9780470061572.eqr492.

[26] López-Martín, C. (2014). Predictive accuracy comparison between neural networks and statistical regression for development effort of software projects. Applied Soft Computing, 27, 434–449. doi:10.1016/j.asoc.2014.10.033.

[27] Macedo-Rouet, M., Ney, M., Charles, S., and Lallich-Boidin, G. (2009). Students' performance and satisfaction with Web vs. paper-based practice quizzes and lecture notes. Computers & Education, 53(2), 375–384. doi:10.1016/j.compedu.2009.02.013.

[28] Macfadyen, L. P. and Dawson, S. (2010), Mining LMS data to develop an "early warning system" for educators: A proof of concept. Computers & Education, 54(2), 588–599. doi:10.1016/j.compedu.2009.09.008.

[29] Manzoor, H. (2013). Measuring student satisfaction in public and private universities in Pakistan. Global Journal of Management and Business Research Interdisciplinary, 13, 3.

[30] Nakhaei, F., Mosavi, M. R., Sam, A., & Vaghei, Y. (2012). Recovery and grade accurate prediction of pilot plant flotation column concentrate: Neural network and statistical techniques. International Journal of Mineral Processing, 110, 140–154. doi:10.1016/j.minpro.2012.03.003.

[31] Negricea, C., I., Edu, T. and Avram, E., M. (2014). Establishing influence of specific academic quality on student satisfaction. Procedia - Social and Behavioral Sciences, 116, 4430–4435. doi:10.1016/j.sbspro.2014.01.961.

[32] Paliwal, M., and Kumar, U. A. (2009). Neural networks and statistical techniques: A review of applications. Expert Systems with Applications, 36(1), 2–17. doi:10.1016/j.eswa.2007.10.005.

[33] Perai, A. H., Moghaddam, H. N., Asadpour, S., Bahrampour, J., and Mansoori, G. (2010). A comparison of artificial neural networks with other statistical approaches for the prediction of true metabolizable energy of meat and bone meal. Poultry Science, 89(7), 1562–1568. doi:10.3382/ps.2010-00639.

[34] Petrović, J., Ibrić, S., Betz, G., and Đurić, Z. (2012). Optimization of matrix tablets controlled drug release using Elman dynamic neural networks and decision trees. International Journal of Pharmaceutics, 428(1), 57–67. doi:10.1016/j.ijpharm.2012.02.031.

[35] Rokach, L., and Maimon, O. (2005). Top-down induction of decision trees classifiers-a survey. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, 35(4), 476–487. doi:10.1109/tsmcc.2004.843247.

[36] Romero, C., and Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. Expert Systems with Applications, 33(1), 135–146. doi:10.1016/j.eswa.2006.04.005.

[37] Rovnyak, S., Kretsinger, S., Thorp, J., and Brown, D. (1994). Decision trees for real-time transient stability prediction. Power Systems, IEEE Transactions on, 9(3), 1417–1426. doi:10.1109/59.336122.

[38] Sharma, S. and Singh, H. (2013). Data mining in education sector. International Journal of Electronics & Data Communication, 2(1), 4–8.

[39] STATISTICA help documentations. `http://documentation.statsoft.com` [Accessed on 10 September 2014].

[40] Stokes, S. P. (2001). Satisfaction of college students with the digital learning environment: Do learners' temperaments make a difference?, Internet and Higher Education, 4, 31–44. doi: s1096-7516(01)00046-x.

[41] Tessema, M. T., Ready, K. and Yu, W.-Y. (2012). Factors affecting college students' satisfaction with major curriculum: Evidence from nine years of data. International Journal of Humanities and Social Science, 2, 34–44.

[42] Thomas, E. H. and Galambos, N. (2004). What satisfies students? Mining student-opinion data with regression and decision tree analysis. Research in Higher Education, 45(3), 251–269. doi:10.1023/b:rihe.0000019589.79439.6e.

[43] Wang, G., Wu, J., Wu, J., & Wang, X. (2011). A comparison between the linear neural network method and the multiple linear regression method in the modeling of

continuous data. Journal of Computers, 6(10), 2143–2148. doi:10.4304/jcp.6.10.2143-2148.

[44] Wei, C. C. and Ramalu, S. S. (2011). Students satisfaction towards the university: Does service quality matters?, International Journal of Education, 3(2). doi:10.5296/ije.v3i2.1065.

[45] Yu, Z., Haghighat, F., Fung, B. and Yoshino, H. (2010). A decision tree method for building energy demand modeling. Energy and Buildings, 42(10), 1637–1646.

[46] Yunus, N. K. Y., Ishak, S. and Razak, A. Z. A. A. (2010). Motivation, empowerment, service quality and polytechnic students' level of satisfaction in Malaysia. International Journal of Business and Social Science, 1, 120–128.

[47] Yýldýrým, H., Altýnsoy, H. B., Barýpçý, N., Ergün, U., Oǧur, E., Hardalaç, F. and Güler, İ. (2004). Classification of the frequency of carotid artery stenosis with MLP and RBF neural networks in patients with coroner artery disease. Journal of Medical Systems, 28(6), 591–601. doi:10.1023/b:joms.0000044961.38008.97.

[48] University of Josip Juraj Strossmayer, Faculty of Education in Osijek, Internal student survey of the Faculty of Education in Osijek, `http://web.foozos.hr/images/dokumenti/osiguravanje-kvalitete/Obrazac_za_provodjenje_interne_studentske_ankete.doc` [Accessed on 26 March 2015].