Machine learning methods in predicting the student academic motivation

Ivana Đurđević Babić[1,†]

[1]*Faculty of Education, University of Osijek, Cara Hadrijana 10, 31 000 Osijek, Croatia*
E-mail: ⟨ *idjurdjevic@foozos.hr* ⟩

Abstract. Academic motivation is closely related to academic performance. For educators, it is equally important to detect early students with a lack of academic motivation as it is to detect those with a high level of academic motivation. In endeavouring to develop a classification model for predicting student academic motivation based on their behaviour in learning management system (LMS) courses, this paper intends to establish links between the predicted student academic motivation and their behaviour in the LMS course. Students from all years at the Faculty of Education in Osijek participated in this research. Three machine learning classifiers (neural networks, decision trees, and support vector machines) were used. To establish whether a significant difference in the performance of models exists, a t-test of the difference in proportions was used. Although, all classifiers were successful, the neural network model was shown to be the most successful in detecting the student academic motivation based on their behaviour in LMS course.

Keywords: academic motivation, machine learning, neural networks, decision tree, support vector machine

1. Introduction

The manner in which a person behaves in terms of academic performance is a subject of academic motivation [36]. In this research, academic motivation is perceived from the theory of self-determination [7], which observes academic motivation from the point of intrinsic, extrinsic and amotivational behaviour. It is widely recognized that academic motivation influences student academic performance and is mostly ignored in higher education circles, although the ability to detect student academic motivation might very well assist educators in improving student performance and ensuring high-quality learning.
Even though a considerable research deals with academic motivation, only a small number of researchers addressed the issue of predicting student academic motivation. In doing so, they generally use traditional statistical methods or structural

† Corresponding author

equation modelling. Student motivation in an online environment is also insuffi-ciently addressed in current literature.

This research has two goals. The main goal is to explore the possibility of creating an efficient classification model for predicting student academic motivation using three machine learning methods where only student log data from the LMS courses are used as input variables. The second goal is to compare the efficiency of obtained machine learning models in predicting student academic motivation to find the most appropriate model. The overall intent of this research is to give educators insight into information extracted from LMS log files concerning student academic motivation, which in turn would enable them to utilize this information and for improving the quality of teaching and learning. This research also attempts to shed light on the relationship between student behaviour in the LMS course and predicting academic motivation. The paper contributes to fulfilling the gap in the existing literature by focusing solely on student behaviour in learning management system (LMS) courses and exploiting machine learning methods in predicting student academic motivation to develop a stable and effective model. The next section provides an overview of previous research in this area. The sample and methodology of machine learning techniques used in the paper are explained in Section 3. Section 4 describes the results and is followed by a conclusion.

2. Literature review

Researchers have recognized the potential of information stored in LMS log files and have used it for achieving various goals in their research. Frequently, their focus has been on predicting student performance (e.g. [29], [30]), but some research goals were set on measuring a sense of community in courses [3], extrac-ting quality characteristics of the LMS courses [38] or even increasing student retention [20]. However, since this research focuses on predicting student academic motivation using machine learning methods and only raw log data, this review of literature presents only the results from several relevant studies that have used diverse predictors available from log files and various methods for predicting academic motivation within an online learning environment.

Although numerous researchers have explored the relationship between motivate-on and other variables, especially academic performance (for instance [1], [2], [9], [23], [27], [33]), there are only rare examples of research that used LMS log files for predicting student motivation. Cocea and Weibelzahl [4] calculated five variables (user identifier, percentage of correctly answered tests, time spent on pages, number of accessed pages, time spent performing tests and motivation) from log files and used decision trees tried to determine whether log files can be used for predicting student motivation. They assessed student motivation using their own rules and emphasized this particular manner of assigning motivational

levels as one of the main limitations of their research. These authors in their later work proposed other approaches in assessing the motivation levels ([5], [6]). The actions and timestamps from log files for only 48 students were used to rate student engagement levels as either engaged, neutral or disengaged based on criteria from student log activities [5]. Their research explored several data mining methods (Bayesian networks with K2 algorithm, logistic regression, simple logistic classification, instance based classification with IBk algorithm, Attribute Selected Classification using J48 classifier and Best First search, Bagging using reduced-error pruning tree classifier, Classification via Regression and Decision Trees with J48 classifier). They then extracted attributes related to reading pages and taking tests (number of false answers in a test, number of true answers in a test, number of tests, average time spent on a test, number of pages, average time spent on pages) as significant predictors of disengaged students [5]. The limitation of the studies were the manner in which student level of motivation was determined and a relatively small sample size.

Other researchers were also mainly focused on time-related variables. Hershkovitz and Nachmias [14] suggested an unvalidated framework for measuring motivation (through engagement, energization and source of motivation) and highlighted validation of classification related to the motivational levels as one of the limitations of their study. Seven variables calculated from log files (time on task percentage, average session duration, average pace of activity within sessions, average time between sessions, pace of word marking, exam activities percent, game activities percent) were classified to propose three motivational dimensions using a clustering process. Kularbphettong and Tongsiri [18] also drew conclusions about student motivation bead on student activities during classes. They explored log files, gathered 19 variables mostly related to time spent doing an activity and the grade received for doing that activity, applied a decision tree method to form association rules from which they extracted important variables that showed the impact on student behaviour. They concluded that data mining methods can be used for predicting student outcomes.

Keeping in mind limitations of past research in predicting student academic motivation based on log files, the Academic Motivation Scale (AMS) designed by Vallerand et al. [37] was used in this paper as a reliable scale for determining student motivation. The AMS is a widely used scale in studies that deal with assessing student academic motivation (e.g. [11], [16], [17], [21], [26]) but, so far, its use in past research on predicting student academic motivation from log files has not been found. In addition, this paper differs from past research in the choice of input variables, and while interdependency between variables in the dataset was not considered in past research, it is the main step in this research to ensure stability and efficiency of the proposed model.

3. Machine learning methods used in this research

3.1. Artificial neural networks (ANN)

The multilayer perceptron (MLP) and the radial basic function (RBF) neural networks were trained and tested in this research.

The MLP neural network consists of the input layer, one or more hidden layers and the output layer. The number of hidden layers can vary depending of the problem being solved, but the most common are MLP neural networks with only one hidden layer. The neurons that form the same layer are not interconnected, but neurons that are in adjacent layers are fully interconnected (a neuron in one layer is connected with all the neurons in an adjacent layer) with weighed connections. Haykin [13] emphasizes the basic characteristics of MLP stating that MLP models include a differentiable activation function, at least one hidden layer and a high degree of connectivity. The back-propagation algorithm is a commonly used method for training MLP networks. Marshland [22] reminds that training MLP networks involves producing outputs for specified inputs and current weights, and then changing the weights in compliance with the difference between the output of the network and the target output.

The signal flow in a MLP network is described by Hsieh [15]. If x_i are inputs and h_j neurons in the hidden layer, f and g activation functions, w_{ji} and $\widetilde{w}_{kj}$ weight parameter matrices, b_j and $\tilde{b}_k$ offset parameters, then the signal flow progress through the network can be described by the equations [15]:

$$h_j = f\left(\sum_j w_{ji} x_i + b_j\right)$$

(1)

$$y_k = g\left(\sum_j \widetilde{w}_{kj} h_j + \tilde{b}_k\right)$$

(2)

with y_k as the output.

The backpropagation algorithm (BFGS - Broyden-Fletcher-Goldfarb-Shanno) and conjugate gradient algorithms embedded in StatSoft Statistica 12 Data Miner were used in this research for training MLP neural network models. Exponential, hyperbolic tangent, logistic, identity or linear function were used as activation functions, whereas sum of squares or entropy were used as error functions in training. The number of hidden units varied from 3 to 20, while the number of training cycles was fixed to 200.

A popular alternative to the MLP neural network is the RBF neural network which is also known as a universal approximator [8]. It consists of three layers (input, hidden and output layer). As explained in [10] the input layer of the RBF neural network receives input data and a non-linear transformation is performed in the hidden layer. In the hidden layer, each unit is described by radial basis function [13]:

$$\varphi_j\,(x) = \varphi(\|x - x_j\|),\ \text{j=1, 2,..., N} \tag{3}$$

where x_j is the centre of the radial basis function and x is the pattern applied to the input layer. There are no weighted connections between the nodes of the input layer and the nodes in the hidden layer [13]. Several different types of functions can be used as radial-basis functions, but the most commonly used is the Gaussian function and is also used in this research. It follows that each unit in hidden layer is described as given in [13]:

$$\varphi_j\,(x) = exp\left(-\frac{1}{2\sigma_j^2}\|x - x_j\|^2\right),\ \text{j=1, .., N} \tag{4}$$

where σ_j is width of the jth Gaussian function with centre x_j.

As pointed out in [8], learning in a RBF neural network (RBFN) is based on determining radial-basis function centres and weights. Unsupervised and supervised learning are both used when training the RBFN [15]. Unsupervised learning is used for finding the centres and widths of the radial-basis functions, while supervised learning is used for minimization of the mean squared error (MSE) between the output data and target data [15].

In this research, the minimum number of hidden units in RBF neural networks was set to 3, while 23 was defined as the maximum number of hidden units. The sum of squares (SOS) or cross entropy (CE) were assigned as error functions.

3.2. Classification tree

Trees, which are used for classifying an object into predefined classes are called classification trees [28]. The main advantages of this method are highlighted in [15]. When using a classification tree it becomes easy to understand the dependence of the predictand on the predictors. This method is an appropriate solution for reducing the number of predictors [15]. The classification tree consists of the root node, internal or test nodes and the leaf or terminal nodes. The incoming and outgoing edges are fundamental in distinguishing the difference between these nodes. The root node has no incoming edges, while others have precisely one incoming edge [28]. Root and internal nodes have outgoing edges, while leaf nodes do not, but one class representing appropriate target is assigned to each leaf [28]. Each internal node contains split where the value of the predictor variable is tested [25]. The classification in decision trees starts at the root and moves

towards one of the leaf nodes were the classification decision occurs [22]. Adaptations of QUEST (Quick, Unbiased, Efficient Statistical Trees) and CART (Classification And Regression Tree) algorithms implemented in Statistica 12 were used as split selection methods in this research. As explained in [32], discriminant-based univariate splits works by establishing the best leaf node to split and predictor variable as a split variable calculating for leaf nodes p-values in order to determine the significance of relation between class membership and level of predictor variable. It chooses the predictor variable with the smallest p-value as a split variable [32].

When using CART, the popular goodness measures (splitting criteria) is a Gini measure of node impurity. As stated in [24], assuming that n is any node, $l \in [1, k]$, $p_l(n)$ a fraction of points at n belonging to class l, Gini is then defined as:

$$Gini\,(n) = \sum_{l=1}^{k} p_l(n)\big(1 - p_l(n)\big)$$

(5)

Besides this measure, two other applied goodness measures were the Chi-square measure and G-square measure. Discriminant-based univariate splits and the C&RT-style exhaustive search for univariate splits were used as split-selection methods with estimated prior probabilities and equal misclassification costs. FACT-style (Fast Algorithm for Classification Trees) direct stopping was applied as a stopping rule with a fraction of the objects set to 0.05. The p-value for split variable selection was set to 0.05 and 10-fold crossvalidation was applied.

3.3. Support vector machines

Support vector machines (SVM) can be used in linear or nonlinear models [34]. Hsieh [15] describes the development of SVM classifier for two-class problem in three steps. In the first step, the maximum margin classifier is introduced for problems where a linear decision boundary can be used for distinguishing two classes, and then in second step, the classifier is altered to allow misclassification, while in third step, using the kernel classifier is nonlinearly generalized [15]. As stated in [19], for a given set of data $(x_1,y_1)..(x_k,y_k)$, $x \in R^n$ as input space, $y_i \in \{-1,1\}$ as class labels, the objective is to form a hyperplane

$$\omega^T x + b = 0$$

(6)

that will divide classes by maximizing the distance between

$$\omega^T x + b = 1$$

(7)

and

$$\omega^T x + b = -1,$$

(8)

that is, by minimizing $\frac{1}{2}\omega^T\omega$, hence the problem can be represented by the following

$$min\frac{1}{2}\omega^T\omega + C\sum_{i=1}^{N}(\xi_i + \xi_i^*) \tag{9}$$

where ξ_i, ξ_i^* represent upper and lower constraints on the outputs and C represents a constant that regularizes the equation.

$$K(x_i, x_j) = \phi(x_i)\phi(x_j) \tag{10}$$

is a kernel function that performs mapping $x_i \in R^n$.

Three different kernel functions were applied in this research:

The Radial Basis Function (RBF) kernel [19]

$$K(x_i,x_j) = \exp\left(-\gamma\|x_i-x_j\|^2\right), \gamma \text{ is a width parameter } (\gamma > 0) \tag{11}$$

sigmoid kernel [35]

$$K(x_i,x_j) = \tanh\left(\gamma x_i x_j + \text{coefficient}\right), \gamma > 0, \text{coefficient} \geq 0 \tag{12}$$

and a polynomial kernel [35]

$$K(x_i,x_j) = \left(\gamma x_i x_j + \text{coefficient}\right)^{\text{degree}}, \text{degree} \in N, \text{coefficient} \geq 0, \gamma > 0. \tag{13}$$

10-fold crossvalidation was applied on all models to avoid overfitting.

4. Sample and research methodology

The research was conducted in the 2015/2016 academic year at the Faculty of Education in Osijek. A total of 129 students from all years of study participated in the research. A majority of the participants were female 126 (97.67%), and the remaining 3 (2.33%) were male. This is not surprising given that mostly females are enrolled at the Faculty of Education in Osijek. The distribution of participants by year of study is given in Table 1.

Study year	Number of students from each study year	Percentage (%)
1st	34	26.36
2nd	36	27.91
3rd	33	25.58
4th	8	6.20
5th	18	13.95

Table 1: *Description of participants by year of study*

At first, 7 variables describing the main behaviour of the students in LMS was considered for the predictors (assign view, course view, forum view discussion, forum view forum, questionnaire submit, questionnaire view, resource view). Following a scatterplot examination suggesting that a relationship (mostly positive) existed between the variables and the first two steps of the three-step proposed modelling procedure in [39], a correlation analysis was conducted.

The results of the correlation analysis are shown in Table 2. The obtained correlation coefficients (see Table 2) suggest that for $p < 0.05$, the observed variables have either weak or moderate correlations, or their correlations are not significant (i.e. negligible). Due to these results, only four variables (V1 – assign view, V2 – forum view discussion, V3 – questionnaire view, V4 – resource view) were taken as predictors for the purpose of modelling, while academic motivation measured using the Academic Motivation Scale (AMS) was set as a dependent variable and presented using binary representation.

	assign view	forum view discussion	questionnaire view	resource view
assign view	1.00	0.29	0.39	0.20
forum view discussion	0.29	1.00	0.14	0.51
questionnaire view	0.39	0.14	1.00	-0.09
resource view	0.20	0.51	-0.09	1.00

Table 2: *Summary of correlation analysis (marked correlations are significant at p<0.05)*

Description of input variables used in modelling procedure is provided in Table 3.

Variable	Valid number (N)	Mean (M)	Standard deviation (SD)
assign view (V1)	129	19.33	23.79
forum view discussion (V2)	129	6.19	6.13
questionnaire view (V3)	129	3.29	5.51
resource view (V4)	129	25.17	28.33

Table 3: *Descriptive statistics of input variables*

The Academic Motivation Scale (AMS) (college version) designed by Vallerand et al. [37] was used for assessing student academic motivation in this research. The AMS consists of 28 items where participants express their agreement with the statement on a 7- point Likert scale ranging from *does not correspond at all* to *corresponds exactly.* The AMS addresses three levels of motivation (intrinsic, extrinsic, amotivation) and its items form seven motivational subscales (intrinsic motivation to know, intrinsic motivation toward accomplishments, intrinsic motivation to experience stimulation, extrinsic motivation – identified, extrinsic motivation – introjected, extrinsic motivation - external regulation, amotivation) [37]. A higher score obtained for each subscale is a sign that a particular motivation type has a higher level. An overall measure of each participant's strength in academic motivation was obtained by following the approach and calculations described in Hackett [12] and calculating the Self Determination Index (SDI) that is used as an overall measure of academic motivation. Since the mean SDI score for all participants was 5.06 (SD=3.05), for the purpose of this research participants whose SDI was lower than 5.06 were considered those with a below-average level of academic motivation and were categorised into the category labelled as 0. Those participants whose SDI was higher than 5.06 were in this research considered as those with an above-average level of academic motivation and they were categorised into the category labelled with 1.

This instrument was used in a number of researches and several of them tested the concurrent and its construct validity (see [31]). An acceptable level of internal consistency of the AMS subscales (the reported values of Cronbach alpha for all subscale ranged from 0.60 to 0.86) and temporal stability were reported in [37].

With the intention of comparing models, the sample was randomly split into training (80%) and testing (20%) subsamples. For neural networks, the training subsample included train and selection subsample.

4. Results

The obtained results (see Table 4) showed that among MLP models, the highest total classification accuracy was obtained when hyperbolic tangent was used as an activation function (61.54%) while the RBF model achieved the highest total

classification accuracy among all tested models (76.92%). The classification tree model with the highest total classification accuracy was obtained by using CART as a split-selection method. The obtained tree was quite large and complex (for the tree structure see Appendix 1). It consisted of 17 splits and 18 terminal nodes. Several SVM models obtained the same total accuracy (57.69%). Among them, the SVM type II model with the RBF kernel resulted with the highest classification accuracy for students with below-average level of academic motivation (63.64%) while the SVM type II model with the RBF kernel resulted in the highest classification accuracy for students with an above-average level of academic motivation (60.00). The SVM type II model with RBF kernel was considered the best SVM model in this research.

Model	Accuracy for students with a below-average level of academic motivation - category 0 (%)	Accuracy for students with an above-average level of academic motivation - category 1 (%)	Total accuracy (%)
MLP - exponential	9.09	93.33	57.69
MLP – tangent hyperbolic	45.45	73.33	61.54
MLP - logistic	9.09	93.33	57.69
MLP - sine	9.09	86.67	53.85
MLP – identity (linear)	9.09	86.67	53.85
RBF	100.00	60.00	76.92
decision tree -CART	53.33	81.81	65.38
decision tree - discriminant-based univariate splits	60.00	63.36	50.00
SVM type I - linear	0.00	100.00	57.69
SVM type I - RBF	0.00	100.00	57.69
SVM type I- sigmoid	0.00	100.00	57.69
SVM type I- polynomial	0.00	100.00	57.69
SVM type II - linear	54.55	46.67	50.00
SVM type II- RBF	63.64	53.33	57.69
SVM type II- sigmoid	54.55	60.00	57.69
SVM type II- polynomial	81.82	6.67	38.46

Table 4: *Results obtained by models on the test sample*

To compare models and determine whether the difference between their total accuracy is statistically significant, a t-test of differences in proportions was appli-

ed. The results of the t-test showed that there was no statistical significance in performance between models at the level of significance of 0.05. The obtained results are given in Table 5.

Hypothesis	p-value results
H_0 : RBF = CT	p =0.18
H_0 : SVM = CT	p =0.28
H_0 : RBF = SVM	p =0.07

Table 5: *Results of t-test of differences in proportions*

Sensitivity, specificity, type I and II errors as well as positive (PPV) and negative predictive values (NPV) were calculated for evaluating and comparing models. According to the obtained confusion matrix for the models (see Table 6), the RBF model has the highest ability in identifying students with a below-average level of academic motivation (100%), while the CT model has the highest ability in detecting students with an above-average level of academic motivation (87%). This model also has the highest type II error (64%) and the lowest type I error (0.13). The SVM model has the lowest ability in detecting students with an above-average level of academic motivation (53%) and accordingly, the highest type I error (0.47) among the observed models and the highest ability in misclassifying students with an above-average level of academic motivation.

Model		0	1
RBF	0	1.00	0.40
	1	0.00	0.60
Classification tree	0	0.36	0.13
	1	0.64	0.87
SVM	0	0.64	0.47
	1	0.36	0.53

Table 6: *Confusion matrix for observed models*

The RBF model, on the other hand, did not showed characteristics of misclassifying students with a below-average level of academic motivation (0.00) and therefore it had the highest ability in correctly recognizing students that truly had a below-average level of academic motivation (see Table 7) and the highest negative predictive rate.

Model	Sensitivity	Specificity	PPV	NPV
RBF	1.00	0.60	0.65	1.00
CT	0.36	0.87	0.67	0.65
SVM	0.64	0.53	0.50	0.67

Table 7: *Measures of observed and evaluated models*

The achieved results suggest that the RBF model is the most suitable model for predicting student academic motivation. To identify links between predicted category of student academic motivation and their behaviour in the LMS course, the Box & Whisker plots of the variables based on values of the RBF output are presented in Figure 1 and Figure 2. The results revealed that students with an above-average level of academic motivation fewer times viewed assignments, questionnaires and forum discussions in LMS courses then students identified as students with a below-average level of academic motivation by the RBF model in this research.
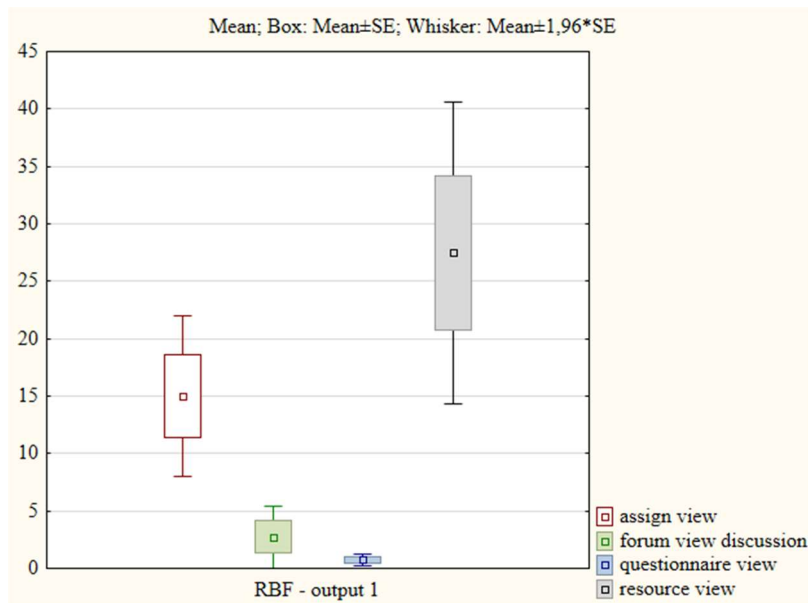

Figure 1: *Box plot of variables based on the value 1 for the RBF output*
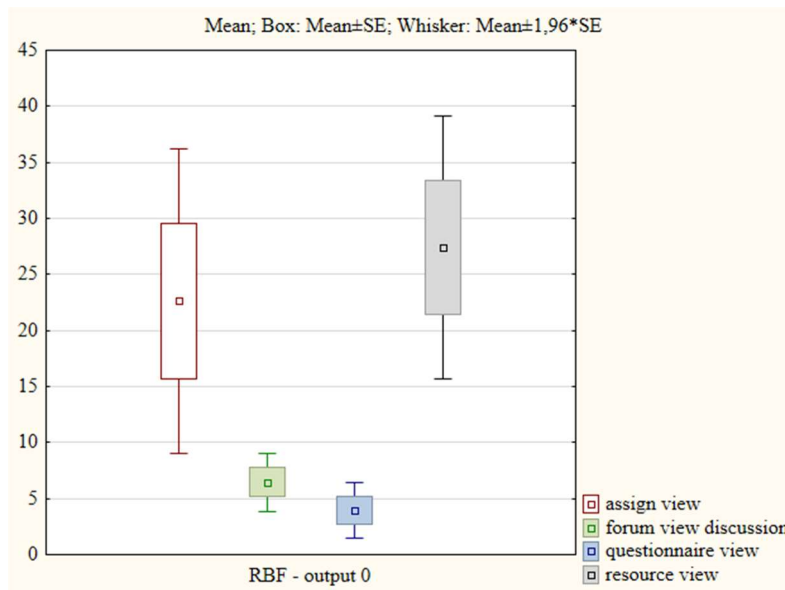
Figure 2: *Box plot of variables based on the value 0 for the RBF output*

5. Conclusion

In this research, the goals of the research were to explore the possibility of creating an efficient classification model for predicting student academic motivation with three machine learning methods when student LMS course data are used as input variables and to find the most suited model by comparing the efficiency of the obtained models. Before conducting the modelling process, the relationship among all potential predictors and then between potential predictors and the target variable was considered which resulted in a decrement of predictors. Three machine learning techniques (neural networks, classification tree, support vector machines) were applied for modelling and all three methods gained acceptable results. The results showed that the RBF neural network model performed better than the best obtained models using the other two techniques. This model produced an overall classification accuracy of 76.92% and used measures for evaluating models which showed that among the tested models, this model was capable of identifying all students with a below-average level of academic motivation (100%) and had the highest negative predictive value (100%) than the two other tested models. Its positive predictive value was satisfactory (65%), but it was not the highest among the observed models. A statistically significant difference in performance at a 5% significance level was not detected with the t-test of differences in proportions between the observed models.

Although this research used different predictors, different algorithms and a different approach in detecting student academic motivation, the obtained results are in accordance with previous research and confirm that machine learning methods can produce an effective model for predicting student academic motivation. However, the neural network model provided the most efficient model in this research.

Academic teachers and other educators could potentially benefit from the results of this research and early recognition of students with a below-average or above-average level of academic motivation. For instance, by recognizing students with below-average level of academic motivation early on, teachers are then able to advise and pair them up for educational purposes (e.g. pair or group projects) with students who possess an above-average level of academic motivation in order to influence motivation, actively engage them in learning and enhance their academic outcomes.

The results obtained in this research are limited to the observed sample which should be extended in future research to make the results more widely applicable. Future research could also be expanded by including some of the other parameters as input variables and other machine learning techniques or a combination of them in the modelling process.

References

[1] Areepattamannil, S. (2014). Relationship between academic motivation and mathematics achievement among Indian adolescents in Canada and India. The Journal of general psychology, 141(3), 247–262.

[2] Areepattamannil, S., Freeman, J. G., and Klinger, D. A. (2011). Intrinsic motivation, extrinsic motivation, and academic achievement among Indian adolescents in Canada and India. Social Psychology of Education, 14(3), 427–439.

[3] Black, E. W., Dawson, K., & Priem, J. (2008). Data for free: Using LMS activity logs to measure community in online courses. The Internet and Higher Education, 11(2), 65–70.

[4] Cocea, M., & Weibelzahl, S. (2006). Can log files analysis estimate learners' level of motivation? In LWA. University of Hildesheim, Institute of Computer Science.

[5] Cocea, M., & Weibelzahl, S. (2007, July). Eliciting motivation knowledge from log files towards motivation diagnosis for Adaptive Systems. In International Conference on User Modeling (pp. 197–206). Springer, Berlin, Heidelberg.

[6] Cocea, M., & Weibelzahl, S. (2011). Disengagement detection in online learning: validation studies and perspectives. IEEE Transactions on Learning Technologies, 4(2), 114–124.

[7] Deci, E. L., Vallerand, R. J., Pelletier, L. G., and Ryan, R. M. (1991). Motivation and education: the self-determination perspective. Educational psychologist, 26(3–4), 325–346.

[8] Du, K. L., and Swamy, M. N. S. (2013). Neural networks and statistical learning. Springer Science & Business Media.

[9] Ennen, N. L., Stark, E., and Lassiter, A. (2015). The importance of trust for satisfaction, motivation, and academic performance in student learning groups. Social Psychology of Education, 18(3), 615–633.

[10] Fyle, C. (2000). Artificial Neural networks and Information Theory. Department of Computing and Information Systems, The University of Paisley.

[11] Guay, F., Ratelle, C. F., Roy, A., & Litalien, D. (2010). Academic self-concept, autonomous academic motivation, and academic achievement: mediating and additive effects. Learning and Individual Differences, 20(6), 644–653.

[12] Hackett, M. (2017). The Impact of Racial-Ethnic Identity on Academic Motivation of African American High School Students (Doctoral dissertation). Retrieved from: http://digitalcommons.liberty.edu/cgi/viewcontent.cgi?article=2526&context=doctoral

[13] Haykin, S. (2009). Neural networks and learning machines (Vol. 3). Upper Saddle River, NJ, USA, Pearson.

[14] Hershkovitz, A., & Nachmias, R. (2008, June). Developing a log-based motivation measuring tool. In Educational Data Mining 2008.

[15] Hsieh, W. W. (2009). Machine learning methods in the environmental sciences: Neural networks and kernels. Cambridge university press.

[16] Komarraju, M., Musulkin, S., & Bhattacharya, G. (2010). Role of student–faculty interactions in developing college students' academic self-concept, motivation, and achievement. Journal of College Student Development, 51(3), 332–342.

[17] Komarraju, M., Karau, S. J., & Schmeck, R. R. (2009). Role of the Big Five personality traits in predicting college students' academic motivation and achievement. Learning and individual differences, 19(1), 47–52.

[18] Kularbphettong, K., & Tongsiri, C. (2012). Mining educational data to analyze the student motivation behavior. World Academy of Science, Engineering and Technology, 68(2012), 1256–1259.

[19] Labjar, H., Cherif, W., Nadir, S., Digua, K., Sallek, B., and Chaair, H. (2015). Support vector machines for modelling phosphocalcic hydroxyapatite by precipitation from a calcium carbonate solution and phosphoric acid solution. Journal of Taibah University for Science.

[20] Lauría, E. J., Baron, J. D., Devireddy, M., Sundararaju, V., & Jayaprakash, S. M. (2012, April). Mining academic data to improve college student reten-

tion: an open source perspective. In Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (pp. 139–142). ACM.

[21] Lee, E. (2005). The relationship of motivation and flow experience to academic procrastination in university students. The Journal of Genetic Psychology, 166(1), 5–15.

[22] Marsland, S. (2015). Machine learning: an algorithmic perspective. CRC press.

[23] Matthews, A. R., Hoessler, C., Jonker, L., and Stockley, D. (2013). Academic motivation in calculus. Canadian Journal of Science, Mathematics and Technology Education, 13(1), 1–17.

[24] Mohri, M., Rostamizadeh, A., and Talwalkar, A. (2012). Foundations of machine learning. MIT press.

[25] Murthy, S. K. (1998). Automatic construction of decision trees from data: A multi-disciplinary survey. Data mining and knowledge discovery, 2(4), 345–389.

[26] Orsini, C., Binnie, V., Evans, P., Ledezma, P., Fuentes, F., & Villegas, M. J. (2015). Psychometric validation of the academic motivation scale in a dental student sample. Journal of dental education, 79(8), 971–981.

[27] Próspero, M., and Vohra-Gupta, S. (2007). First generation college students: motivation, integration, and academic achievement. Community College Journal of Research and Practice, 31(12), 963–975.

[28] Rokach, L., and Maimon, O. (2015). Data mining with decision trees: theory and applications. World scientific.

[29] Romero, C., Espejo, P. G., Zafra, A., Romero, J. R., & Ventura, S. (2010). Web usage mining for predicting final marks of students that use Moodle courses. Computer Applications in Engineering Education, 21(1), 135–146.

[30] Romero, C., Ventura, S., & García, E. (2008). Data mining in course management systems: Moodle case study and tutorial. Computers & Education, 51(1), 368–384.

[31] Smith, K. J., Davy, J. A., and Rosenberg, D. L. (2012). An empirical analysis of an alternative configuration of the academic motivation scale. Assessment in education: Principles, policy & practice, 19(2), 231–250.

[32] StatSoft, Classification Trees. Available at: http://www.statsoft.com/Textbook/Classification-Trees [Accessed 25/8/2016]

[33] Struthers, C. W., Perry, R. P., and Menec, V. H. (2000). An examination of the relationship among academic stress, coping, motivation, and performance in college. Research in higher education, 41(5), 581–592.

[34] Suthaharan, S. (2016). Machine learning models and algorithms for big data classification. Boston: Springer.

[35] Tseng, T. L. B., Aleti, K. R., Hu, Z., and Kwon, Y. J. (2016). E-quality control: a support vector machines approach. Journal of Computational Design and Engineering, 3(2), 91–101.

[36] Usher, E. L., and Morris, D. B. (2012). Academic motivation. In Encyclopedia of the Sciences of Learning (pp. 36–39). Springer US.

[37] Vallerand, R. J., Pelletier, L. G., Blais, M. R., Briere, N. M., Senecal, C., and Vallieres, E. F. (1992). The Academic Motivation Scale: A measure of intrinsic, extrinsic, and amotivation in education. Educational and psychological measurement, 52(4), 1003–1017.

[38] Valsamidis, S., Kontogiannis, S., Kazanidis, I., Theodosiou, T., & Karakos, A. (2012). A Clustering Methodology of Web Log Data for Learning Management Systems. Educational Technology & Society, 15 (2), 154–167.

[39] Zekic-Sušac, M., and Đurđević Babić, I. (2015). Discovering patterns of student behaviour in e-learning environment. In Z. Kolar-Begović, R. Kolar-Šuper & I. Đurđević Babić (Eds.) Higher Goals In Mathematics Education (94), Zagreb: Element.

Appendix 1: Classification tree table presentation.

Node	Left branch	Right branch	Predict. class	Split constant	Split variable
1	2	3	1	21,50	assign view
2	4	5	1	26,50	resource view
3	6	7	1	21,00	questionnaire view
4	8	9	1	1,50	assign view
5	10	11	0	6,50	forum view discussion
6	12	13	1	0,50	forum view discussion
7	--	--	0	--	--
8	14	15	0	1,50	forum view discussion
9	16	17	1	9,50	resource view
10	--	--	1	--	--
11	--	--	0	--	--
12	--	--	0	--	--
13	18	19	1	22,50	forum view discussion
14	20	21	1	1,50	questionnaire view
15	--	--	0	--	--
16	--	--	1	--	--
17	22	23	1	7,50	forum view discussion
18	24	25	1	8,50	forum view discussion
19	--	--	0	--	--
20	26	27	1	1,50	resource view
21	--	--	0	--	--
22	28	29	0	12,50	resource view
23	--	--	1	--	--
24	30	31	1	27,00	resource view
25	--	--	1	--	--
26	32	33	1	0,50	resource view
27	--	--	1	--	--
28	--	--	0	--	--
29	--	--	1	--	--
30	--	--	1	--	--

31	--	--	0	--	--
32	34	35	1	0,50	assign view
33	--	--	0	--	--
34	--	--	1	--	--
35	--	--	0	--	--