Predicting students' results in higher education using a neural network

Oancea, Bogdan and Dragoescu, Raluca and Ciucu, Stefan

April 2013

Predicting students' results in higher education using a neural network

Bogdan Oancea[1], Raluca Dragoescu[2], Stefan Ciucu[3]

[1]*Finance and Accounting Department, "Nicolae Titulescu" University, Calea Vacaresti 185, sector 4, Bucharest, Romania*
Bogdan.oancea@gmail.com
[2]*Statistics and Econometrics Department, The Bucharest University of Economic Studies, 6, Piata Romana, 1st district, Bucharest, Romania*
dragoescuraluca@gmail.com
[3]*IT Department, "Nicolae Titulescu" University, Calea Vacaresti 185, sector 4, Bucharest, Romania*
stefanciucu@univnt.ro

Abstract: A significant problem in higher education is the poor results of students after admission. Many students leave universities from a variety of reasons: poor background knowledge in the field of study, very low grades and the incapacity of passing an examination, lack of financial resources. Predicting students' results is an important problem for the management of the universities who want to avoid the phenomenon of early school leaving. We used a neural network to predict the students' results measured by the grade point average in the first year of study. For this purpose we used a sample of 1000 students from "Nicolae Titulescu" University of Bucharest from the last three graduates' generations, 800 being used for training the network and 200 for testing the network. The neural network was a multilayer perceptron (MLP) with one input layer, two hidden layers and one output layer and it was trained using a version of the resilient backpropagation algorithm. The input data were the students profile at the time of enrolling at the university including information about the student age, the GPA at high school graduation, the gap between high school graduation and higher education enrolling. After training the network we obtained MSE of about 1.7%. The ability to predict students' results is of great help for the university management in order to take early action to avoid the phenomenon of leaving education.

Keywords: higher education, neural networks, prediction.

Introduction

Romania is among European countries with the highest school dropout rate. The development of strategies to prevent school dropout should be among the priorities for educational institutions. A significant problem in higher education in Romania is the poor results of students after admission. Because of the poor results during the first year of study many students leave universities. The reasons for these poor results are various: poor background knowledge in the field of study, very low grades and the incapacity of passing an examination, lack of financial resources and the increasing costs of education. That's why one of the main aims of the management system of the universities is to avoid the phenomenon of early school leaving. The ratio between graduating students and enrolled students is one of the indicators used for accreditation of higher education institutions in Romania. The universities' management tries to keep this ratio at a good value by means of early intervention and supporting those students with problems and that are candidates to leave school. In this regard, predicting students' results is an important problem and has a great value to universities for early intervention to avoid school leaving.

Many researchers tried to predict the students' results based on various data. Predictions were made using different statistical methods like multivariate regression, path analysis or discriminant analysis. None of these methods have the power of discovering potential data patterns as neural networks. Feed forward neural networks are applied in many fields like financial forecasting, medical diagnosis, bankruptcy prediction, OCR for regression or classification purposes because they are one of the best functional mappers. The good results of applying neural networks in classification problems lead us to use them for predicting students' results in higher education.

In this paper we describe an implementation of a feed forward neural network used to predict the GPA after the first year of study. The paper is organized as follows. The next section reviews the related work in the field of using neural network for prediction and classification purposes in education. Then we describe our implementation of the neural network and the data set used for training and testing. Next, the results of using the network are presented, showing the good capabilities for classification of the NN. The final section presents some conclusions and ideas for further development of the study.

Related work

Neural networks were used by many researchers for predictions of students' results. In (Cooper, 2010) the author presents a neural network-based decision support system that identifies students who are "at-risk" of not retaining to their second year of study. The system correctly predicted retention for approximately 70% of the students. (Halachev, 2012) presents a neural network used for prediction of the outcome indicators of e-Learning, based on Balanced ScoreCard. The author obtained a 3-4% prognosis error which is acceptable from a practical point of view.

(Livieris et al., 2012) describes the implementation of a user-friendly software tool based on neural network classifiers for predicting the student's performance in the course of "Mathematics" of the first year of Lyceum. The authors of the paper compared many training algorithms namely the Broyden-Fletcher-Goldfarb-Shanno (BFGS), the Levenberg-Marquardt (LM), the Resilient Backpropagation (Rprop) and the modified spectral Perry (MSP). The performances of the neural network classifier were also compared with other classifiers such as Bayesian networks and support vector machines and the authors proved the better results of the neural network classifier.

Data registered by Moodle were used in (Calvo-Flores et al., 2006) for predicting students' marks. The authors used a RBF neural network and information logged by Moodle regarding the number and types of educational resources accesses to predict the students' marks at a discipline. They obtained an accuracy of prediction of about 75-80%. The model developed in this paper shows that it is possible to predict those students with problems to pass a course and to give professors an indication to pay more attention on those students that probably will fail passing a course.

In (El Moucary et al., 2011) the authors present a use Neural Networks and Data Clustering based method designed to predict students' GPA according to their foreign language performance for those students who study in a foreign language. In a second stage the students are grouped in well-defined cluster for further advising. They obtained a maximum error of prediction less than 10% for GPA.

Neural networks were also used in (Naik et al., 2004) to predict MBA student success. The authors classified applicants to MBA program into successful and marginal student pools based on undergraduate GPA, undergraduate major, age, GMAT score using a neural network with three layers. They obtained overall prediction accuracy for their model of about 89%. To assess the ability of the neural network for classification of student, the authors compared the results obtained using the neural network with a logit and probit regression model. The overall rate of accuracy in prediction of the logit model was about 73% while the accuracy of the probit was 73.37%.

(Oladokun et al., 2008) used a neural network in order to determine those factors that influence students' performance. They classified students in three classes according to their results. The accuracy of prediction obtained by the authors of the paper was about 74%. (Karamouzis et al., 2008) used a three-layer perceptron network trained by backpropagation to predict student graduation rate. The network model build by the authors had an accuracy of 70.27% for successful graduates and 66.29% for unsuccessful graduates.

Methodology

For our study, we used a sample of 1000 students from the last three graduates' generations from "Nicolae Titulescu" University from Bucharest. Most of the students that leave studies take their decision after the first year. In our data sample, about 70% of the students who left university took this decision after the first year of study. Most of them had a very low GPA, usually lower than 6 (in Romania, grades range from 1 to 10). In order to avoid the phenomenon of early school leaving, we build a model to predict the GPA of the students after the first year of study and classify them in three classes according to theirs GPA. A lower GPA is a serious indication that the student has difficulties in finding his/her educational path and is a premise for early school leaving.

Our model uses a neural network with one input layer, two hidden layer and one output layer. As input data for predicting the GPA after first year we used:

- Type of the study program: distance education (part time) or full time education;
- Gender of the student;

- High-school graduation GPA;
- Age of the student;
- Difference in years from the moment the student graduates high-school until he/she enrolls at university.

We classified students according to theirs GPA after the first year of study in three classes:

- POOR RESULTS – those students with GPA lower than 6;
- MEDIUM RESULTS– those students with GPA between 6 and 8;
- GOOD RESULTS – those students with GPA greater than 8.

Our task was to predict the class a student belongs based on the five input variables. The first layer of the neural network comprises 7 neurons: two for type of education (0/1 full time/part time), two for gender (0/1 – M/F), one for high school graduation GPA, one for the age of the student and one for the difference in years from the moment the student graduates high-school until he/she enrols at university. The last three variables were normalized to [0, 1] interval. We conducted a series of tests in order to establish the number of hidden layers and the number of neurons in each hidden layer. Our tests give us that the best results are obtained with two hidden layers, the first layer having 50 neurons and the second layer having 400 neurons. The output layer has three neurons, one for each of the three classes. The input layer and each of the hidden layers also receive a bias signal of 1. Every neuron in a layer is fully connected to every neuron in the next layer.

Each neuron accumulates the input from the neurons in the preceding layer and calculates an output signal according to:

$$y_i = f\left(\sum w_{ij} x_j + b\right) \tag{1}$$

where *b* is the bias input and *f* is the activation function of the neuron. We used the *tanh* as an activation function for the two hidden layers and *softmax* function for the output layer. We used MSE (mean square error) as a network error function.

We tested several training algorithms: backpropagation, quick propagation, classical resilient propagation, scaled conjugate gradient. We find that the best results are obtained using a version of the Resilient backpropagation called iRPROP+ which is one of the best performing first-order learning methods for feed forward neural networks (Igel, 2000).

The classical RPROP training algorithm controls the weight update for each connection thus maximizing the update step size and minimizing oscillations. The direction of weight update is based on the sign of the partial derivative $\frac{\partial E}{\partial w_{ij}}$, where *E* is the error function and w_{ij} is the weight from neuron *j* to neuron *i*. The update size is different for each weight and is independent of the absolute value of the partial derivative. If the partial derivative $\frac{\partial E}{\partial w_{ij}}$ has the same sign for consecutive steps, the step-size is increased, otherwise it is decreased.

The weight updates are computed according to the change in sign of the partial derivative: if the sign has not changed, the weight update is done normally but if the sign has changed the previous weight update is reverted (Riedmiller, 1993).

The iRPROP+ training algorithm modifies the rule of weight updating and reverts only weight updates that have caused sign changes of the partial derivate and an error increase. This rule combines information the sign of the error function which is error surface information with the magnitude of the network error when the decision of reverting an update step is taken.

Results

We have implemented the neural network with the Encog framework (Heaton, 2011) using Java programming language. We used a sample of 1000 records from the last three graduates' generations: 800 records where used for training the network and 200 records were used for testing the network.

Choosing number of the neuron and the layers is a difficult problem. A small number of neurons and layers will lower the mapping power of the network. On the other hand, a large number of hidden layers and neurons could give the network the power to fit very complex data but it will slow down the training process. The network structure was found on a trial and error basis. We started with a small network and gradually increase its size. Finally we found that the best results are obtained for a network with the following structure: 7I-50H-400H-3O, i.e. 7 input neurons, first hidden layer with 50 neurons, a second hidden layer with 400 neurons and an output layer with 3 neurons.

We trained the network for 100.000 epochs on a computer with INTEL I7 processor and 4 GB of RAM memory under the Windows 7 operating system. Before training the network, the order of the training data records was randomized. The MSE obtained after training the network was 1.7%. The mean square error for the test data set was 1.91%.

In our data set 30.1% of students belong to the class "POOR RESULTS", 50.9% of them to the class "MEDIUM RESULTS" and 19% to the class "GOOD RESULTS".

Table 1 shows the number of students in each class predicted by our neural network using the test data set compared with the real number of students in each class.

Table 1

The results of the neural network

CLASS	Number of students	Predicted Values	Predicted value (percent)
POOR RESULTS	60	52	86,6%
MEDIUM RESULTS	105	99	94.2%
GOOD RESULTS	35	30	85,7%

The results presented in table 1 are encouraging: 86.6% of the poor results students were predicted by the network. This prediction could be used by

Conclusion

Prediction is an important tool and represents a first step in intervention from the university management to avoid the phenomenon of early school leaving. We used the classification power of a neural network to predict the potential students with problems in continuing their education. Our network achieved an accuracy of over 86%.

We used the Encog framework for building the network that was a feed forward Multi Layer Perceptron with one input layer, two hidden layers and one output layer. The activation function of the hidden layers was *tanh* and for the output layer we used the *softmax* function. We tested several training algorithms and we found out that the best results were obtained using the iRPROP+ algorithm.

The average predictability rate was 86% for the "POOR RESULTS" class of students which represents the pool of potential students that are candidates for leaving the university. This predictability rate is comparable with other results presented in the beginning of the paper. This encourages us to continue the research, first of all by expanding the number of input variables included in each student' record. A larger number of input variables will increase the predictability power of the network. Another improvement would be to increase the number of training records which is also useful for cross-validation.

Another direction for a future research would be to identify those parameters from the input data that influence in a greater extent the students' decision to leave education. This will be helpful for educational administrators to intervene in the shortest possible time. Eliminating these parameters will contribute to an increasing predictability power of the network since they are only noise.

References

1. Baker, B.D. and Richards, C.E., 1999. A comparison of conventional linear regression methods and neural networks for forecasting educational spending, *Economics of Education Review,* 18, pp. 405-415.
2. Calvo-Flores, M.D, Galindo, E.G., Pegalajar Jiménez, M.C. and Piñeiro, O.P., 2006. Predicting students' marks from Moodle logs using neural network models, A. Méndez-Vilas, A. Solano Martín, J.A. Mesa González and J. Mesa González (*eds*.) Current Developments in Technology-Assisted Education, pp. 586-590.
3. Cooper, C.I., 2010. A Generalizable Neural Network for Predicting Student Retention, Professional Papers Proceedings Archive, Available at: http://www.asee.org/documents/zones/zone1/2010/professional/A-Generalizable-Neural-Network-for-Predicting-Student-Retention.pdf, 17.01.2013.
4. Halachev, P., 2012. Prediction of e-Learning Efficiency by Neural Networks, *Cybernetics and information technologies*, Volume 12, No 2, pp. 98-108.
5. Heaton, J., 2011. Programming Neural Networks with Encog3 in Java, 2nd Edition, Heaton Research.
6. Igel, C., Hüsken, M., 2000. Improving the Rprop Learning Algorithm, *Proceedings of the Second International Symposium on Neural Computation*, pp. 115–121, ICSC Academic Press.
7. Karamouzis S.T., and Vrettos, A., 2008. An Artificial Neural Network for Predicting Student Graduation Outcomes, *Proceedings of the World Congress on Engineering and Computer Science*, pp. 991-994.
8. Livieris, I.E., Drakopoulou, K., Pintelas, P., 2012. Predicting students' performance using artificial neural networks, *Proceedings of the 8th Pan-Hellenic Conference "Information and Communication Technology in Education*", Available at : http://hcicte2012.uth.gr/main/sites/default/files/proc/Proceedings/LivierisEtAl.pdf, 17.01.2013.
9. El Moucary, C., Khair, M., Zakhem, W., 2011. Improving Student's Performance Using Data Clustering and Neural Networks in Foreign-Language Based Higher Education, *The Research Bulletin of Jordan ACM*, vol II, pp. 27-34.
10. Naik, B., Ragothaman, S., 2004. Using Neural Network to Predict MBA Student Success, *College Student Journal*, Vol. 38, No. 1. pp. 1-4.
11. Oladokun, V.O., Adebanjo, A.T., and Charles-Owaba, O.E., 2008. Predicting Students' Academic Performance using Artificial Neural Network: A Case Study of an Engineering Course, *The Pacific Journal of Science and Technology*, Volume 9, Number 1, pp. 72-79.
12. M. Riedmiller and H. Braun, 1993. A direct adaptive method for faster backpropagation learning: The RPROP algorithm. In *Proceedings of the IEEE International Conference on Neural Networks*, pages 586–591. IEEE Press.