

Predicting students' performance using artificial neural networks

Ioannis E. Livieris*, Konstantina Drakopoulou, Panagiotis Pintelas

Department of Mathematics, University of Patras, GR 265-00, Greece.

Educational Software Development Laboratory, Department of Mathematics, University of Patras, GR 265-00, Greece.

Abstract

Artificial intelligence has enabled the development of more sophisticated and more efficient student models which represent and detect a broader range of student behavior than was previously possible. In this work, we describe the implementation of a user-friendly software tool for predicting the students' performance in the course of "Mathematics" which is based on a neural network classifier. This tool has a simple interface and can be used by an educator for classifying students and distinguishing students with low achievements or weak students who are likely to have low achievements.

Keywords: Artificial neural networks, training algorithms, educational data mining, student's performance.

1. Introduction

During the last few years, the application of artificial intelligence in education has grown exponentially, spurred by the fact that it allows us to discover new, interesting and useful knowledge about students. Educational data mining (EDM) is an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational context. While traditional database queries can only answer questions such as "*find the students who failed the examinations*", data mining can provide answers to more abstract questions like "*find the students who will possibly succeed the examinations*". One of the key areas of the application of EDM is the development of student models that would predict student characteristics or performances in their educational institutions. Hence, researchers have begun to investigate various data mining methods to help educators to evaluate and improve the structure of their course context (see [17, 18] and the references therein).

The academic achievement of higher secondary school education (Lyceum) in Greece is a deciding factor in the life of any student. In fact, Lyceum acts like a bridge between school education and higher learning specializations that are offered by universities and higher technological educational institutes. Limiting the students that fail in the final examinations is considered essential and therefore the ability to predict weak students could be useful in a great number of different ways. More specifically, the ability of predicting the students' performance with high

*Corresponding author

Email address: livieris@upatras.gr (Ioannis E. Livieris)

accuracy in the middle of the academical period is very significant for an educator for identifying slow learners and distinguishing students with low achievements or weak students who are likely to have low achievements. By recognizing the students' weaknesses the educators are able to inform the students during their study and offer them additional support such as additional learning activities, resources and learning tasks and therefore increase the quality of education received by their students. Thus, a tool which could automatically recognize in time students' performance and especially students with learning problems is really important for educators.

However, the idea of developing an accurate prediction model based on a classifier for automatically identifying weak students is a very attractive and challenging task. Generally, datasets from this domain skewed class distribution in which most cases are usually located to the one class. Hence, a classifier induced from an imbalanced dataset has typically a low error rate at the majority class and an unacceptable error rate for the minority classes.

In this work, we propose the application of an artificial neural network for predicting student's performance at the final examinations in the course of "Mathematics". Our aim is to identify the best training algorithm for constructing an accurate prediction model. We have also evaluated the classification accuracy of our neural network approach by comparing it with other well-known classifiers such as decision trees, Bayesian networks, classification rules and support vector machines. Moreover, we have incorporated our neural network classifier in a user-friendly software tool for the prediction of student's performance in order to making this task easier for educators to identify weak students with learning problems in time.

References

- [1] C.M. Bishop. *Neural Networks for Pattern Recognition*. Oxford, 1995.
- [2] G. Brown, A. Pocock, M.J. Zhao, and M. Lujan. Conditional likelihood maximisation: A unifying framework for information theoretic feature selection. *Journal of Machine Learning Research*, 13:27–66, 2012.
- [3] W. Cohen. Fast effective rule induction. In *International Conference on Machine Learning*, pages 115–123, 1995.
- [4] P. Domingos and M. Pazzani. On the optimality of the simple Bayesian classifier under zero-one loss. *Machine Learning*, 29:103–130, 1997.
- [5] M.T. Hagan and M.B. Menhaj. Training feedforward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks*, 5(6):989–993, 1994.
- [6] S. Haykin. *Neural Networks: A Comprehensive Foundation*. Macmillan College Publishing Company, New York, 1994.
- [7] R. Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *IEEE International Joint Conference on Artificial Intelligence*, pages 223–228. AAAI Press and MIT Press, 1995.
- [8] B. Lerner, H. Guterman, M. Aladjem, and I. Dinstein. A comparative study of neural network based feature extraction paradigms. *Pattern Recognition Letters*, 20(1):7–14, 1999.
- [9] I.E. Livieris, M.S. Apostolopoulou, D.G. Sotiropoulos, S. Sioutas, and P. Pintelas. Classification of large biomedical data using ANNs based on BFGS method. In *13th Panellenic Conference of Informatics*, pages 87–91, 2009.
- [10] I.E. Livieris and P. Pintelas. An improved spectral conjugate gradient neural network training algorithm. *International Journal on Artificial Intelligence Tools*, 21(1), 2012.
- [11] G.D. Magoulas, V.P. Plagianakos, M.N. Vrahatis, and G.S. Androulakis. Neural network-based colonoscopic diagnosis using on-line learning and differential evolution. *Applied Soft Computing*, 4:369–379, 2004.
- [12] D. Nguyen and B. Widrow. Improving the learning speed of 2-layer neural network by choosing initial values of adaptive weights. *Biological Cybernetics*, 59:71–113, 1990.
- [13] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer-Verlag, New York, 1999.
- [14] J.C. Platt. *Advances in Kernel Methods - Support Vector Learning*. MIT Press, Cambridge, Massachusetts, 1998.
- [15] J.R. Quinlan. *C4.5: Programs for machine learning*. Morgan Kaufmann, San Francisco, 1993.
- [16] M. Riedmiller and H. Braun. A direct adaptive method for faster backpropagation learning: The Rprop algorithm. In *IEEE International Conference on Neural Networks*, pages 586–591, San Francisco, CA., 1993.
- [17] C. Romero and S. Ventura. Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications*, 33:135–146, 2007.

- [18] C. Romero, S. Ventura, and E. Garcia. Data mining in course management systems: Moodle case study and tutorial. *Computers & Education*, 51(1):368–384, 2008.
- [19] D.E. Rumelhart, G.E. Hinton, and R.J. Williams. Learning internal representations by error propagation. In D. Rumelhart and J. McClelland, editors, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, pages 318–362, Cambridge, Massachusetts, 1986.