Computers in Human Behavior xxx (2016) 1–11

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Full length article

Improving the expressiveness of black-box models for predicting student performance

Carlos J. Villagrá-Arnedo, Francisco J. Gallego-Durán, Faraón Llorens-Largo, Patricia Compañ-Rosique, Rosana Satorre-Cuerda, Rafael Molina-Carmona*

Department of Computer Science and Artificial Intelligence, University of Alicante, Carretera San Vicente del Raspeig s/n, 03690 Alicante, Spain

ARTICLE INFO

Article history:
Received 16 March 2016
Received in revised form
27 July 2016
Accepted 1 September 2016
Available online xxx

Keywords:
Black-box models
Prediction
Student performance
Graphical representation

ABSTRACT

Early prediction systems of student performance can be very useful to guide student learning. For a prediction model to be really useful as an effective aid for learning, it must provide tools to adequately interpret progress, to detect trends and behaviour patterns and to identify the causes of learning problems. White-box and black-box techniques have been described in literature to implement prediction models. White-box techniques require a priori models to explore, which make them easy to interpret but difficult to be generalized and unable to detect unexpected relationships between data. Black-box techniques are easier to generalize and suitable to discover unsuspected relationships but they are cryptic and difficult to be interpreted for most teachers. In this paper a black-box technique is proposed to take advantage of the power and versatility of these methods, while making some decisions about the input data and design of the classifier that provide a rich output data set. A set of graphical tools is also proposed to exploit the output information and provide a meaningful guide to teachers and students. From our experience, a set of tips about how to design a prediction system and the representation of the output information is also provided.

1. Introduction

Improving student performance, knowing their actual progress and trying to predict their results at the earliest stages of the learning process can be extremely important to act early and cut off the problems at the root. These needs have triggered several research lines in different fields aimed to find ways to predict the learning outcomes.

An interesting research line is modelling automatic prediction tools based on statistical analysis or machine learning techniques. An automatic prediction system is a technological tool that, from a number of variables that refer to behaviour or activity of students in a learning system, obtains in advance an estimation of the final student performance. This is a very general definition, so that each specific prediction system may differ in certain aspects: the techniques used to build the system, the input variables to measure the behaviour or activity of students and the output variables that offer an estimation of students' performance. The quality of the prediction system can also be measured in different ways: by focusing on purely metrical aspects (for example, the accuracy we get with the proposed method), or in terms of the expressiveness of the output to provide real help to the student.

Beyond the technical aspects of the model, the focus in this paper is on the expressiveness of the system. Modern educational theories advocate a student-centred training, with a truly formative assessment, and not just a classification. Therefore, for a prediction model to be really useful it should offer something more than a mere classification of students at the end of the learning process. It must also be able to provide tools to adequately interpret progress, to detect trends and behaviour patterns and to identify the causes of learning problems. This way, the prediction model will truly guide the students and detect the actual problems of the learning process.

There is a large amount of techniques to implement prediction models. In Learning Analytics, there are some techniques that require a priori models to explore (white-box techniques), and other more general techniques not requiring a priori models (black-box techniques). White-box techniques are easy to interpret since

* Corresponding author.
E-mail addresses: villagra@dccia.ua.es (C.J. Villagrá-Arnedo), fgallego@dccia.ua.es (F.J. Gallego-Durán), faraon@dccia.ua.es (F. Llorens-Largo), patricia@dccia.ua.es (P. Compañ-Rosique), rosana@dccia.ua.es (R. Satorre-Cuerda), rmolina@dccia.ua.es (R. Molina-Carmona).

http://dx.doi.org/10.1016/j.chb.2016.09.001
0747-5632/

the a priori model provides a straightforward explanation of the relationships between data. However this prior knowledge integrated in the model reduces the search space and so it may introduce an important bias. Moreover, they are difficult to generalize, since they follow an ad-hoc design. On the contrary, black-box techniques are easier to generalize, are suitable to discover unsuspected relations and oriented to automatic unsupervised computation. Unfortunately, the technical aspects are quite cryptic for most teachers and they are very difficult to interpret.

In this paper the use of black-box techniques is proposed to take advantage of the power and versatility of these methods, while making some decisions about the design of the system that allows a richer interpretation. Specifically, the design and construction of a prediction model based on a Support Vector Machine (SVM) is used (Villagra-Arnedo, Gallego-Durán, Molina-Carmona, & Llorens-Largo, 2015). The decisions about the input data set design, the number and distribution of the classes, the use of probabilistic classification and the exploitation of time factor, provide a richer output. Moreover, a set of graphical tools has been proposed to exploit the expressiveness of the output information and provide a meaningful guide to teachers and students.

In Section 2, a background of the research is presented, including previous related works, a discussion about the current research extent and what problems remain unsolved, and the research questions that guide the further decisions to design the proposed system. Section 3 is devoted to present the proposed prediction model, the design aspects, the results of the prediction and the representation of the output information. In Section 4, some tips about how to generalize the results of the experience are described. Finally, some conclusions and further research are presented in Section 5.

2. Background

Prediction can be defined as the inference of some information (the predicted or dependent variables) from a combination of other data (the predictor or independent variables). A prediction model is an analysis tool that obtains the predicted variables from a small sample of data, considering the statistical validity of the model so that it can be applied to the whole population (Berland, Baker, & Blikstein, 2014).

There is a large amount of techniques to implement prediction models. Kotsiantis (2012) makes an interesting review of the use of these techniques for educational purposes (classification and regression algorithms, association rules, sequential patterns analysis, clustering and web mining). This study identifies Machine Learning techniques as an emerging field that aims to develop methods of exploration of educational data and to find meaningful patterns. It specifically states that data collected from Learning Management Systems (LMS) or Intelligent Tutoring Systems (ITS) can be useful in developing prediction algorithms based on Machine Learning. It also points out that most research is about building ad-hoc models.

There are some techniques that require a priori models to explore (they are easier to interpret but designed ad-hoc), and other more general techniques non-requiring a priori models (suitable to discover unsuspected relations and oriented to automatic unsupervised computation but not so easy to understand). The former are known as white-box techniques and the latter as black-box techniques. The following paragraphs provide a brief review of some works about prediction based on white-box and black-box techniques in the field of education.

Macfadyen and Dawson (2010) develop a white-box system to predict students' performance using frequency of use and time spent on learning activities in a LMS. They perform a correlation analysis of several variables and the final grade, obtaining thirteen significant correlations. They admit that the predictive power of simple correlations used as a priori model is very limited, since the results show the obvious: most active students in the LMS end up being the most successful. Therefore, additional analyses are performed using multiple and logistic regression to obtain an accuracy of 70% in predicting students at risk of failure. Anyway, the need of an a priori model is an important limitation for the analysis.

Another example is the work of Pedro, Baker, Bowers, and Heffernan (2013). They analyse and predict enrolment in a course using data from an ITS. Again an a priori model is used, in this case logistic regression, to predict whether a student will enrol in a course using a combination of representative characteristics about emotional state, motivation, knowledge and other usage data. After conditioning these variables by statistical analysis, they find a combination thereof to detect students to be enrolled with an accuracy of nearly 70%. The solution is designed completely ad-hoc for the problem.

Decision trees technique is a habitual white-box method used in learning prediction. For instance, Hu, Lo, and Shih (2014) develop a system for early prediction of final student performance. Learning portfolios are used as input to the algorithm. The authors determine the time-dependent variables of the learning activities of students in a LMS and collect activity data for three different moments of the course (weeks 4, 8 and 13). Fourteen features are used, grouped into four types: access behaviour, use of online course materials, task status and participation in a discussion forum. They use three classification techniques based on decision trees, obtaining an overall accuracy of 95% in week 4. Combining these techniques with Adaboost (Freund, Schapire, & others, 1996, pp. 148–156) the accuracy is increased to almost 98%. The prediction system prototype is made up of a set of decision rules that automatically trigger alerts based on the values of the most significant dependent time-based variables. Despite the high accurate results, the proposal is highly dependent on the specific case, so it is difficult to be extrapolated to other experiences. Nevertheless, the fact of performing a prediction in several moments opens a new exploring path for building progressive prediction systems that can be used to detect trends.

Another example of the use of a priori models is the one proposed by Ley and Kump (2013) to collect data on the interaction of six people in a period of two months. Its aim is to find out if interactions with a professional learning system integrated into a workplace can be used to predict three levels of expertise (beginner, advanced or expert). They compare a task-based approach and a knowledge event-based approach. They use linear and logistic regression and conclude that combining both types of data as input improves the prediction significantly. The use of three classes instead of just a binary classification provides a more expressive model that can be considered in other research experiences.

Xing, Guo, Petakovic, and Goggins (2015) explore the development of more usable prediction models and representations. They use data from a collaborative geometry problem-solving environment to construct the prediction model. First, they link online learning with online participation, and operationalize activity theory to holistically quantify students' participation in the computer-supported course. As a result, 6 variables, Subject, Rules, Tools, Division of Labour, Community, and Object, are constructed. They make an interesting analysis of variables prior to the application of a model and, as a result, the data dimensionality is diminished and the data are systematically contextualized in a semantic background. They apply Genetic Programming as prediction model. As a result, the system has a high prediction rate and an interesting interpretability. Unfortunately the generalization of

Please cite this article in press as: Villagrá-Arnedo, C. J., et al., Improving the expressiveness of black-box models for predicting student performance, Computers in Human Behavior (2016), http://dx.doi.org/10.1016/j.chb.2016.09.001