



21st International Conference on Knowledge Based and Intelligent Information and Engineering Systems, KES2017, 6-8 September 2017, Marseilles, France

Decision tree learning used for the classification of student archetypes in online courses

Alexandru Topîrceanu^{a,*}, Gabriela Grosseck^b

^aDepartment of Computer and Information Technology, Politehnica University Timisoara, Bd. V. Parvan 2, 300223 Timisoara, Romania

^bFaculty of Psychology and Sociology, West University Timisoara, Bd. V. Parvan 4, 300223 Timisoara, Romania

Abstract

With ubiquitous Internet access nowadays, individuals have the ability to share more information than before, and it allows young people to collaborate and learn from a distance, so that educational systems are constantly being reshaped. Understanding eLearning is important, and so is the typology of students who participate in this trend with increasing dedication. Yet, we consider that this accelerated pace of propagation of online education has left behind an important aspect needed for the act of teaching, namely studying and understanding student archetypes. By this we mean the common patterns which define the interaction type, dedication amount, and finalization perspective of courses. This paper introduces an original set of student profiles specific to online courses, and it does so by means of data mining and supervised learning. We use the responses from an online questionnaire to gather detailed opinion from 632 students from Romania regarding the advantages and disadvantages of MOOCs, as well as the reasons for not joining online courses. Based on the extracted statistics, we present six decision trees for classifying the finalization and participation rates of online courses based on the students individual traits. Furthermore, we discuss these profiles and explain the implications of this study. We believe our findings to bring consistent novelty both in understanding the needs of modern students, as well as in optimizing the way eLearning is further developed.

© 2017 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of KES International

Keywords: classification, decision trees, student profiles, eLearning, social media analysis;

1. Introduction

The current learning environment is characterized by openness and dynamism, so that a significant proportion of students have a declared preference for flexible learning^{22,19}, through which they can fulfil their academic pursuits, as well as job responsibilities and family chores. As a consequence, educational models like eLearning evolved in order to offer such flexibility¹⁹. While eLearning systems have major strengths, they also pose many challenges to the educational community. One noteworthy difficulty is the lack of spatial and temporal synchronization between the teacher and student, which is an obstacle for students trying to follow up on teachers. The amount of data generated

* Corresponding author. Tel.: +40-256-403261;
E-mail address: alextp@cs.upt.ro

by virtual learning systems sometimes overwhelms educators, who are unable to process the information without the support of special business intelligence tools and techniques specific for large data and Big Data analysis and visualization⁶.

The performance of students enrolled in both offline and online education is important for many institutions, because their strategic programs can be planned to improve this performance. There are studies in this sense, taking into consideration the average grades upon graduation, or track completion^{2,24}, based specifically on data mining techniques in order to predict the drop out rate of students. In particular, decision tree techniques^{20,30} are applied to create surveys that predict the likelihood to drop out from college, then these are turned over to management for direct or indirect intervention²⁴.

To support this direction of research in educational science, we rely on decision tree learning techniques to go beyond the simple statistical analysis and profiling of students done in eLearning^{7,27}. The methodological novelty presented in this study consists on applying supervised learning on a survey with results from 632 participants, in order to define student profiles or archetypes specific to online course and massive online open courses (MOOCs). The dataset consists of mostly undergraduate students from Romania, and quantifies their overview on eLearning and MOOCs in particular. By analysing important metrics, like participation rate, finalization rate and certification, with the help of recursive partitioning⁵, we are able to easily express the educational outcomes with the help of student expressed feedback. This in turn, makes it possible to define representative profiles of students which can help in the further development of online education.

Studies also try to understand the motivation of learners, and one such project shows the need for better alignment between teacher belief and student expectations¹⁰. Namely, it is shown that the instructors *believe* that learner performance is tied to instructor actions focused on course content, but student satisfaction is more likely tied to their feeling that their interpersonal communication needs are met. Learners consider items focused on communication needs and being treated as individuals as most important, which is often difficult to quantify in most surveys, feedback or statistics.

To improve the efficiency of eLearning, the goal of this study is set to define a set of archetypes which can quickly assess any student, so that educators better understand their inner drives to participate and finalize a course. Using our defined profiles we consider that a more personalized educational experience may be automatically offered, to each student in turn, using an online learning framework. Our proposed classification extracts valuable knowledge regarding the advantages, disadvantages of participation in eLearning, as well as expectations and reasons for not participating, all seen from the perspective of students.

2. State of the art in decision tree learning and student profiling

In this paper we employ several decision trees learning algorithms to obtain insights into empirical data based on the supervised machine learning principle.

Decision tree learning is used in^{20,30} to link observations about entities (represented in the tree's branches) to conclude upon the entities' target value (represented in the tree's leaf nodes). It is consistently used in data mining^{7,27}, and represents a predictive modelling approach. Based on what values the target variable can take, there are classification trees, where the target variable can take a finite set of values, respectively regression trees, where the target variable can take continuous values; we only make use of classification trees.

Since all the data is available in the survey dataset, in form of questionnaire answers given by students, we employ supervised learning, in the sense that we pick a few labels of interest (e.g. whether the course was finalized or the participation amount) and classify the corresponding students based on some grouped input parameters. The resulting decision trees are useful for online course managers to evaluate their existing students, to assign them to our defined profiles, knowing their possible strengths and weaknesses.

Recursive partitioning is a strong statistical tool used for multivariable analysis⁵. Recursive partitioning tries to create a decision tree that can classify individuals of a population, by splitting the population into smaller populations based on certain independent variables. Well known methods of recursive partitioning include Ross Quinlan's ID3 algorithm²⁵, and C4.5 and C5.0¹⁵. The advantages of this method are the generation of more intuitive models that do not require the user to perform calculations¹⁶; it allows varying prioritizing of misclassifications in order to create a decision rule that has more sensitivity or specificity⁸; and it can be very accurate^{17,15}. The major disadvantages

Download English Version:

<https://daneshyari.com/en/article/4960582>

Download Persian Version:

<https://daneshyari.com/article/4960582>

[Daneshyari.com](https://daneshyari.com)