

Input determination for models used in predicting student performance

Karlis Kruminis, Sarma Cakula
Vidzeme University of Applied Sciences

INTRODUCTION

Student performance prediction has become a viable means to improving academic performance and course content in online learning. Predictive models such as neural networks, decision trees and linear regression are used to transform inputs (e.g. past performance, social background, learning system usage patterns, test results) into outputs (course completion, expected grade, difficulties encountered, personalized suggestions). Often, the existing quantitative data drive model design, especially when applying such models to the conventional classroom and the person delivering the course, is a passive participant in designing models and delivering data.

In seeking to capture and code as much student behavior and environment as possible to apply learning analytics to a mostly conventional classroom, the most successful inputs (predictors) among existing models can be identified, categorized and their common characteristics determined. Together with a study of formative and summative assessment methods (e.g. types of feedback and how it can be captured) and factors affecting student performance in the classroom (e.g. environmental factors), this allows to identify the existing data in classrooms that are not captured by current learning management systems, thus allowing the expanded use of learning analytics and student performance prediction in traditional classrooms, with a focus on personalized suggestions.

The goal of the paper is to identify patterns among inputs used in existing models of student learning (based on online learning and learning management system data mining) that can then also be applied to the traditional classroom.

Research question: how can characteristics common to effective predictors of student performance be used to identify predictors among data produced in the traditional classroom?

MATERIAL AND METHODS

A literature review is performed where inputs captured and features discovered in existing learning analytics systems are characterised, along with methods used to identify those and the modelling approaches employed.

An attempt is made to identify measures in online learning that may have analogues in the traditional classroom (e.g., seating patterns and communication in chatrooms) or for which proxies may be found (e.g. screen size and lighting quality, where the proxy is the classroom number).

The corresponding outputs are recorded where possible, with a focus on those that allow providing feedback for individual students or for course/curriculum deliverers/designers (i.e. allow to improve the success of future students in this course).

RESULTS

Successful predictors and characteristics common to those are identified, so that they can be used in features engineering for student performance prediction models.

Predictors used in online learning are categorised, so that analogous inputs can be developed for use in traditional classrooms.

Types of feedback provided by existing models of learning are identified, where possible, along with the corresponding input (weights of inputs).

Studies are identified where learning personnel, not the researcher, were able to drive the model development process.

DISCUSSION

Recently, there has been increasing focus on increasing the visibility into models of learning and of involving learning personnel in designing, modifying and running those models. Providing inputs and recognizing the features they represent determines the success of such models. Therefore, recognizing existing successes and applying them to formative assessment methods may be a means of recognizing additional inputs to and features used in models, while involving educators. Applying learning models to the traditional classroom as an integrated part of the learning management (school record keeping/grading) systems may allow to expand their use, while simultaneously increasing the predictive power and effectiveness of (personalized) suggestions, both by using existing data, and by providing tools for educators to transform the existing feedback they provide into data that can be used as inputs for models.

CONCLUSION

Predictors used in learning models in online learning can be applied to the traditional classroom. Analogues may be found for predictors that are not available in the conventional classroom. Common characteristics and categorisation of predictors may be used to identify predictors among existing data, including data provided by students (e.g. formative feedback) that is not captured by the existing learning management systems used.

KEYWORDS

Online learning, input selection, prediction model.