Forecasting Students’ Performance Through Self-Regulated Learning Behavioral Analysis

Rodrigo Lins Rodrigues, Universidade Federal Rural de Pernambuco, Recife, Brazil
Jorge Luís Cavalcanti Ramos, Universidade Federal do Vale do São Francisco, Petrolina, Brazil
João Carlos Sedraz Silva, Universidade Federal do Vale do São Francisco, Petrolina, Brazil
Raphael A. Dourado, Universidade Federal de Pernambuco, Recife, Brazil
Alex Sandro Gomes, Universidade Federal de Pernambuco, Recife, Brazil

ABSTRACT

The increasing use of the Learning Management Systems (LMSs) is making available an ever-growing, volume of data from interactions between teachers and students. This study aimed to develop a model capable of predicting students’ academic performance based on indicators of their self-regulated behavior in LMSs. To accomplish this goal, the authors analyzed behavioral data from an LMS platform used in a public University for distance learning courses, collected during a period of seven years. With this data, they developed, evaluated, and compared predictive models using four algorithms: Decision Tree (CART), Logistic Regression, SVM, and Naïve Bayes. The Logistic Regression model yielded the best results in predicting students’ academic performance, being able to do so with an accuracy rate of 0.893 and an area under the ROC curve of 0.9574. Finally, they conceived and implemented a dashboard-like interface intended to present the predictions in a user-friendly way to tutors and teachers, so they could use it as a tool to help monitor their students’ learning process.

KEYWORDS

Educational Data Mining, Learning Analytics, Learning Management Systems, Learning Systems, Self-regulated Learning

1. INTRODUCTION

Given the evolution in the development of technological artifacts for education, computational environments for teaching and digital contents start to be seen as essential instruments in the teaching process. Also, the use of these online learning platforms is contributing to the expansion of distance learning courses (Wang, Doll, Deng, Park, & Yang, 2013), which in turn generate huge volumes of data as a result of the users’ interactions with the system. These data reflect the students’ behavior and, if correctly explored, can offer important clues about how students interact and learn (Koedinger, et al., 2010).

However, keeping track of all students’ interactions in different tools is one of the most exhausting and challenging tasks for teachers nowadays. Given the growing size and heterogeneity of classes and the limited time available, teachers tend to adopt a passive posture, limited to answering messages
and grading assignments. In this way, noticing the individuality of each student and guiding his/her learning process becomes challenging in the distance learning context.

In virtual learning environments, the students can count on flexibility of time, space, and features, which also brings the need to be more responsible and autonomous. As a result, students have a greater level of control over their own learning process (Pintrich, 1999). It is also necessary to take into account that each student has different levels of proficiency in establishing self-regulated learning strategies. On one hand, those who are more skilled in this ability can control the rhythm and direction of their learning process, and manage a set of characteristics in the learning environment — thus making choices that can aid in their cognitive structure development (Broadbent and Poon, 2015).

On the other hand, many students have difficulties self-regulating and self-monitoring their routine learning tasks. These students, in most cases, come from a past of dependency on teachers and administrative staff for regulation, which is common in traditional pedagogy. Furthermore, the conditioning of these students by traditional face-to-face education makes them more dependent on personalized guidance and direct supervision from teachers in order to manage their assignments.

In this context, recent findings by authors like Kizilcec et al. (2016) state that such socioemotional abilities —— like self-regulation and self-monitoring —— influence individual’s capacity to learn. Such research aim to understand which abilities affect the learning process, based on each student behavioral profile in the virtual learning platforms. This phenomenon have also been addressed by Cognitive Psychology scholars in efforts to understand the self-regulation learning process by studying the individuals’ strategies to manage resource, time, effort, environment, interaction, and help-seeking (Pintrich, 1999).

In this way, our work aims to answer the following research question: Is it possible to develop a predictive model capable of detecting students’ academic performance based on their self-regulatory behavior? From this main research question, we derived the following sub-questions:

- What are the behavioral variables recorded in LMS environments that best describe the students’ self-regulated learning strategies?
- What are the relationships between self-regulation skills and students’ academic performance?

By answering these research questions, this study offers two main contributions. The first one is the development and validation of a model for predicting student’s performance. The second one is the design of a dashboard that teachers and tutors can use to monitor their students’ self-regulatory behavior, improve their pedagogical decisions, and help mitigate student dropout and failure.

The remainder of this paper is organized as follows: Section 2 describes the background and related work regarding the relationship between self-regulated learning (SRL) and academic performance. In Section 3, we present the methodology used to conduct the research. In Section 4, we present the results obtained from the use of educational data mining techniques. In Section 5, we describe the application developed to present the data mining results in a user-friendly way. In Section 6, we discuss the results and its implications and, finally, in Section 7, we draw some conclusions and perspectives for future work.

2. BACKGROUND

2.1. Self-Regulated Learning (SRL)

Self-regulated learning is a cognitive process that allows the development of strategies to face tasks and challenges (Wang, Shannon & Ross, 2013). According to this understanding, the autonomous, active, and planned actions of an individual in the teaching-learning process is crucial.

Other researchers define SRL as a process by which individuals, after establishing goals that reflect their expectations, develop strategies to achieve these goals, thus creating the necessary conditions
Collaborative Learning in Dynamic Group Environments
[www.igi-global.com/chapter/collaborative-learning-dynamic-group-environments/53513?camid=4v1a](www.igi-global.com/chapter/collaborative-learning-dynamic-group-environments/53513?camid=4v1a)

Intelligent Language Tutoring System: Integrating Intelligent Computer-Assisted Language Learning Into Language Education
[www.igi-global.com/article/intelligent-language-tutoring-system/229018?camid=4v1a](www.igi-global.com/article/intelligent-language-tutoring-system/229018?camid=4v1a)