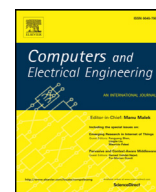




Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compelecengData mining for modeling students' performance: A tutoring action plan to prevent academic dropout[☆]Concepción Burgos^a, María L. Campanario^a, David de la Peña^b, Juan A. Lara^{b,*}, David Lizcano^b, María A. Martínez^b^a Madrid Open University, MOU – Facultad de Ciencias Económicas y Empresariales, Ctra. De la Coruña, km 38.500 – Vía de Servicio, 15 – 28400, Collado Villalba, Madrid, Spain^b Madrid Open University, MOU – Escuela de Ciencias Técnicas e Ingeniería, Ctra. De la Coruña, km 38.500 – Vía de Servicio, 15 – 28400, Collado Villalba, Madrid, Spain

ARTICLE INFO

Article history:

Received 7 March 2017

Accepted 8 March 2017

Available online xxx

Keywords:

E-learning

Student dropout prediction

Educational data mining

Logistic regression model

Temporal data

Student dropout prevention

Tutoring action plan

ABSTRACT

E-learning systems generate huge amounts of data, whose analysis may become a daunting task which makes it necessary to use computational analytical techniques and tools.

We propose the use of knowledge discovery techniques to analyse historical student course grade data in order to predict whether or not a student will drop out of a course. Logistic regression models are used for the purpose of classification. Experiments conducted with data on over 100 students for several distance learning courses confirm the predictive power of our proposal. Using the resulting predictive models we have designed a tutoring action plan. Applying this plan, we managed to reduce the dropout rate by 14% with respect to previous academic years in which no dropout prevention mechanism was applied. Our main contribution is a tool and a tutoring plan that can be used by our educational institution (and others) to reduce dropout rate in e-learning courses.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

E-learning systems have huge benefits but they also raise major challenges for the educational community, including student dropout detection and prevention without direct contact between instructors and students [1].

Current methodologies stress the importance of continuous assessment and student coursework. As students complete the different activities, instructors start to gather information on student grades for activities. This information is potentially very interesting, for example, for implementing tutoring actions to prevent student dropout. For example, Fig. 1¹ shows the grades (used as a benchmark for this research) attained by a group of students (rows) for three specific activities (columns), taken from the Moodle platform.

[☆] Reviews processed and recommended for publication to the Editor-in-Chief by Guest Editor Dr. S. A. Aljawarneh.

* Corresponding author.

E-mail address: juanalfonso.lara@udima.es (J.A. Lara).

¹ The table is in Spanish because it is a real excerpt from a gradebook of a course that we teach. The English equivalents of the main items in Fig. 1 follow:

- Actividad: Activity
- Caso: Case
- AEC: Continuous Assessment Activity