



## Extracting useful information from student usage data for predicting scholar failure and other interesting items

### Short description (what need was solved?)

This project aimed to carry out the incorporation of data mining processes to the different learning management systems that the Santillana group has in production all around the world (about 500.000 students in more than 15 countries). These data were stored in a database (*Datamart*), while all the information related to the students' evaluations are stored in another one (*Pleno*). The idea was to use this available information of interaction to extract new useful knowledge for the improvement of learning, and also to help the development of marketing strategies of said applications.

### What service(s) provided?

The project was developed in three phases:

- 1. Analysis of the available information.** During this phase we had access to the design of the information systems that support the different applications and, after un deep understanding of the available information (obtained by the careful study of the data sources and interviews with people from the company staff) we develop a report about the different knowledge discovery processes that could be developed. This document was analysed by the company in order to decide the following steps to give.
- 2. Preparation of the information.** Original database had more than 500000 registers taken from users all around the world. These data had to be prepared to be used for building data prediction models. The preprocessing included data
- 3. Development of predictive models.** Different prediction models of school failure were developed using the information corresponding to the activities carried out by the student in the system. The extracted knowledge was represented as decision trees, given the company's interest in the straightforward interpretation of the models obtained.

### Relationship with digitization

The developed models allow teachers to offer a series of recommendations and alerts, such as: problems of integration, problems of reading comprehension or attention deficit, increase the number of activities of a certain subject / topic, increase of theoretical resources to improve the preparation of exams, etc..

### Customer, details

Santillana Global S.L.  
Avda. de los artesanos, 6.  
Tres Cantos, 28760 – MADRID  
<http://www.santillana.es>  
<https://santillana.com>