Abstract

The Federal Institute of Education, Science and Technology of Rio Grande do Norte (IFRN), has operations in the areas of teaching, research and extension; contributes more broadly to human and civic education; and stimulate socio-economic development, as it enhances scientific solutions, technical and technological, with commitment to extend benefits to the community.

This broad coverage throughout the territory north of Rio Grande contributes to position both IFRN as an educational institution, science and technology as their Camper as knowledge production links and social development. Thus ensures the maintenance of respectability in the communities in which they operate Camper and credibility built over the history of the institution [Political-Pedagogical Project IFRN, 2012].

In recent years, it has been observed, from educators and education managers, a high rate of repetition and dropout rates in some courses and even, there is a certain disincentives by students in certain courses, taught in several campus of the institution. It is a fact, for example, that some courses, classes of 3rd and 4th periods, arrive with average very low students, around 20%, in other words, these classes began with 40 students and the 4th period are with maximum 8 to 10 students. This has as a consequence, several teachers, teaching classes for a few students, decreasing in much the relationship student teacher, one of the index, which is used by the federal government to evaluate the Institute. Given this situation, the need to do further analysis of such problems. In other words, want to analyze the repetition rate, the dropout rate, the pass rate, completion rate and also the index of canceled registrations. With the purpose, trying to map what are the causes that involve, directly or indirectly in these indices.

Therefore, we are proposing this project the use of machine learning resources, to make an analysis on the mass of the academic system of the institution data, to try to map the profiles of repetition and dropout, two causes of concern and all education managers in Brazil, because according to data published by the PUND (because according to data published by the UNDP (United Nations Development Programme), Brazil, with a dropout rate of 24.3% (2013 data)), Brazil, with a dropout rate of 24.3% (data of 2013), It has the highest rate of school dropouts among 100 countries with the highest HDI (Human Development Index). In Latin America, only Guatemala (35.2%) and Nicaragua (51.6%) have higher dropout rates. It is also being proposed a Business system intellingence (BI), which will include a Data Warehouse (DW), which will be consolidated all data from various data sources, and with the greater purpose, to generate accurate reports for managers take quicker and more accurate way of making. In addition, it is also being proposed an educational counseling system, using this, an area of Artificial Intelligence (AI), known as Case-Based Reasonry - **CBR**) to assist the teaching staff in the guidance to students and their parents. Finally, as our goal is to improve the rates of repetition and dropout, it is also being proposed a game in the form of gamification, for the purpose of engaging increasingly our students in their courses and thus lower these rates.

**Key words**: Business Intellingence, Data Mining, Knowledge Discovery in Databases, Case Based Reasoning, Machine Learning, Gamification