Literature Review

Several research studies have identified many benefits and applications of data mining in education. Romero and Ventura (2010) have identified student’s modelling, predicting students’ performances, data visualization, social network analysis, feedback for support management, planning and scheduling, grouping students and detection of undesirable behaviors as applications of educational data mining. Many studies such as those done by Baradway and Pal (2012), Kabackijeva (2012), Mardikyan and Badur (2012), and Hsia et al. (2008) used decision trees as a support tool for classification of big data.

Support tools such as decision trees and rule learners allow big data to be classified and evaluated. This can then be used in producing predictive models and identifying groups of students. In order to guide students and instructors, the students’ behavior needs to be understood. Without data mining, there is no use of collecting data as useful information has not yet been extracted and cannot be used to make decisions and conclusions. Multiple universities have taken advantage of data mining in order to enhance their educational system for their students. Purdue University and Austin Peay State University are reputable examples.

Purdue University’s Course Signals make use of a student’s grades, demographic characteristics, past academic history and effort to predict their performance. Instructors can then provide real-time feedback to a student without having to spend time analyzing each student’s data on their own. Course Signals present the information through a stoplight system. A red light would indicate a high likelihood of failure; yellow indicates a possible problem while green indicates a high likelihood of success within the course. Research indicates that there was a noticeable increase in satisfactory grades in courses that implement Course Signals, and a decrease in dissatisfactory grades and withdrawals. Thus, Purdue University has demonstrated that learning analytics provided by data mining has favorably impacted their students’ success.

Austin Peay State University has successfully applied data mining in education through their course recommendation system, Degree Compass. Tristan Denley (2012) made use of big data from past students’ grades and transcripts to pair current students with courses that are deemed best fitting to them. Degree Compass uses predictive analytics techniques to rank courses according to measured factors. Courses that apply directly to the student’s program of study is ranked and modelled to predict the courses that the student will excel the most in. Additionally, Degree Compass provides strategic reports that are used by department advisors. These reports allow targeted interventions, for instance, identifying students that would benefit from mentoring. The project had shown a significant difference between the grade distributions in students who used the Degree Compass to select their courses and students who did not. Tristan Denley (2012) has proven that data mining successfully guided both staff and students in their education journey.