ABSTRACT

Topic*:* *STUDENT ALCOHOL CONSUMPTION*

Educational data mining is the process of applying data mining tools and techniques to analyze data for educational purpose. This paper carries out educational data mining to study the student alcohol consumption through a public dataset which includes student attributes and their grades. The decision tree algorithm and the random forest algorithm are applied to perform classification and to analyze the variable importance. The regression model is then employed to illustrate the relationship between alcohol consumption level and the students' final grades. Our analysis provides knowledge on the relationship between student characteristics and alcohol consumption. The study also compares performance of the decision tree algorithm and the random forest algorithm.

In our society, there are number of problems arises due to the students consuming alcohol during its teen age. Retrieving exact and accurate students which consumes alcohol is the main task. It is a real-world problem in our society. The major challenge for finding alcohol addicted students with respect to given data is to find accurate and efficient method which takes less time to generate results. There is large amount of data available, but getting the right information accessible when needed is very important. The availability of educational data has been growing rapidly, and there is a need to analyze hedge amount of data generated from this educational ecosystem. Educational data mining (EDM) has been emerged as a process of applying data mining tools and techniques to analyze the data at educational institutions. This area of research is gaining popularity due to potential benefits to the educational field. Educational institutions use educational data mining (EDM) to gain deep and through knowledge to enhance its assessment, evaluation, planning, and decision making in its educational programs. EDM helps academic programs to identify and discover hidden patterns in the data. These extracted patterns can be used for finding students who are consuming alcohol and its affect on their academic performance. In our proposed system we will use some educational institutes student's data and generate prediction weather student is alcohol addicted or not, we will do this by using clustering, classification, and filtering methods of data mining.