STUDENT GRADE ANALYSIS PREDICTION

Tikam(2022542)   Kartikeya Malik(2022243)

Problem:
→ The objective of this project is to predict the final grades of IIIT Delhi students.
→ The problem statement involves using a dataset of students and their attributes to develop classification algorithms that can identify whether a student performs well in the final grade exam.
→ By predicting student grades, this project aims to provide valuable information to facilitate student retention in courses and help identify at-risk students early on.

Methods:
→ Since universities are prestigious places of higher education, student retention in these universities is a matter of high concern.
→ It has been found that most of the students drop-out from the universities during their first year is due to lack of proper support in undergraduate courses.
→ Due to this, the undergraduate student's first year is referred to as a "make or break" year.
→ Not getting any support on the course domain and its complexity may demotivate a student and cause them to withdraw from the course.
 → There is a great need to develop an appropriate solution to assist student retention at higher education institutions.
→ Early grade prediction is one of the solutions that have a tendency to monitor student's progress in the degree courses at the University and will lead to improving the student's learning process based on predicted grades.
→ Using machine learning with Educational Data Mining can improve the learning process of students.
→ Different models can be developed to predict student's grades in the enrolled courses, which provide valuable information to facilitate student retention in those courses.
→ Various models, including Linear Regression, ElasticNet Regression, Random Forest, Extra Trees, SVM, Gradient Boosted, and Baseline, are used to predict student's grades in enrolled courses.
→ This information can be used to early identify students at-risk based on which a system can 1 suggest the instructors to provide special attention to those students.
→ The predicted grades provide insights into age, urban/rural background, previous failures, family education, higher education aspirations, social activities, and reasons for choosing a college.
→ This information can also help in predicting the students grades in different courses to monitor their performance in a better way that can enhance the students retention rate of the universities.
→ These insights can help educators and institutions enhance student's learning process and improve retention rates.

Dataset:
→ The dataset used in this project focuses on student achievement in college.
→ It includes attributes such as student grades, demographic information, social factors, and college-related features.
→ The data was collected through college reports and questionnaires.
→ Two datasets are provided for the performance in two distinct subjects: Machine Learning and Mathematics
→ G4(4th period), has a strong correlation with G3, G2 and G1, which represent the grades in the 3rd, 2nd and 1st periods, respectively.
→ It is more difficult to predict G4 without G3, G2 and G1, but such prediction is much more useful.

Results:
→ To analyze the dataset and gain insights into the relationship between different attributes and the final grade, this project employs various visualization techniques, including KDE plots, box plots, histograms, and count plots.
→ The results provide valuable information on factors influencing student's performance, such as age, urban/rural background, previous failures, family education, higher education aspirations, social activities, and reasons for choosing college.
→ These insights can help educators and institutions improve learning and retention rates.