SDG 4-“Student Final Grade Performance”

Why SDG 4 is given Quality Education title ?

SDG 4 is given the title "Quality Education" because it is essential for achieving all of the other SDGs. Education is the key to poverty reduction, economic growth, and sustainable development. It empowers people to make informed decisions, participate in their communities, and build a better future for themselves and their families.

Education liberates the intellect, unlocks the imagination and is fundamental for self-respect. It is the key to prosperity and opens a world of opportunities, making it possible for each of us to contribute to a progressive, healthy society. Learning benefits every human being and should be available to all.

● Approach to solve the problem

Problem Statement : The Government of India's initiative for UN Sustainable Development Goal 4, our university is leveraging advanced analytics and machine learning to analyze a comprehensive education dataset. Focused on socio-economic factors, our objective is to develop a predictive model.

How we have approach the problem statement: Upon visualizing regression plots, we observed a strong correlation among the variables G1, G2, and G3 in student performance. It was evident that G3 is influenced by G2, and G2, in turn, is dependent on G1. During analysis, we identified 53 instances where G1, G2, and G3 were all equal to zero, allowing us to discern patterns and potential areas for improvement. In light of these findings, we propose the implementation of a threshold value to proactively identify students with lower performance, aiming to employ targeted interventions and strategies for performance enhancement.

Solution:- Our aim is develop a predictive model to forecast the final performance or grade of students based on various input factors. Additionally, our objective includes implementing a recommendation system to identify and advise students whose projected final performance falls below a certain threshold. The ultimate goal is to enhance academic support and intervention strategies, thereby promoting overall student success.

● Findings and analysis made from available data.

The Datasets Consists of 1044 rows and 33 features. These are features in the dataset 'school', 'sex', 'age', 'address', 'famsize', 'Pstatus', 'Medu', 'Fedu', 'Mjob', 'Fjob', 'reason', 'guardian', 'traveltime', 'studytime', 'failures', 'schoolsup', 'famsup', 'paid', 'activities', 'nursery', 'higher', 'internet', 'romantic', 'famrel', 'freetime', 'goout', 'Dalc', 'Walc', 'health', 'absences', 'G1', 'G2', 'G3' .

The academic performance of students, as reflected in the grades G1, G2, and G3, exhibits a strong correlation, indicating a progressive influence on each other. This suggests that early

academic achievements (G1) significantly contribute to subsequent performance (G2), ultimately impacting the final grades (G3).

Simultaneously, the observation that students engage in alcohol consumption on both weekdays and weekends implies a potential influence on their overall well-being and, consequently, academic outcomes.

Moreover, the importance of maternal education emerges as a key factor for student growth, underlining the pivotal role that a mother's educational background plays in shaping a student's academic journey. Additionally, the significant role played by mothers as guardians underscores their impactful involvement in the holistic development and academic success of students. This interconnectedness highlights the need for comprehensive support systems and interventions that recognize and address the multifaceted aspects influencing students' educational experiences and achievements.

• Best predictive model selected based on experimentation, along with justifications and performance evaluation.

In the conducted analysis, we aimed to identify the best predictive model for student performance (G3) by considering 15 selected features.

Model into Consideration 'Linear Regression', 'Ridge Regression', 'Lasso Regression', 'Elastic Net Regression', 'Decision Tree Regressor', 'Random Forest Regressor', 'Gradient Boosting Regressor', 'XGBoost Regressor', 'Support Vector Regression', 'Multi-layer Perceptron Regressor', 'Bagging Regressor', 'Extra Trees Regressor', 'KNeighbors Regressor', 'Huber Regressor', 'SGD Regressor'.

The models, including Linear Regression, Ridge Regression, Huber Regressor, SGD Regressor, and Gradient Boosting Regressor, were trained and evaluated across multiple random states. The average accuracy was computed for each model, with 'Linear Regression' consistently demonstrating the highest average accuracy across random states.

This indicates that, on average, Linear Regression provided the most reliable predictions among the considered models. The choice of Linear Regression as the best model is justified by its consistently strong performance, making it a robust candidate for predicting student performance based on the selected features.

• Present the findings of the study through an interface and make use of the model for estimating student performance based on various factors, that is convenient to an end user.

We have developed the predictive model for the Final Grade of the Students, Also we are recommending features for better Education Growth. Utilizing a Graphical User Interface (GUI) for data analysis enhances accessibility and comprehension.

Addressed Key Questions:-

1) Do all the features available in the dataset really affect the Quality of Education?

- **NO**

2) What are the predicting variables actually affecting the Quality of Education?

- **The Best 15 Selected Features are 'address', 'Medu', 'Fedu', 'Mjob', 'traveltime', 'studytime', 'failures', 'schoolsup', 'higher', 'internet', 'Dalc', 'Walc', 'absences', 'G1', 'G2'**

3) How does a student's social life affect the quality of education?

'Schoolsup', 'famsup', 'activities', 'internet','goout','romantic'

- Surprisingly, the availability of internet access at home ('internet') doesn't show a significant association with final grades, challenging the assumption that online resources alone may directly correlate with academic success.

- The quality of education appears to be influenced by a complex interplay of social and support structures, emphasizing the need for tailored interventions to address individual needs and challenges in the academic journey.

Team work:

Ladli and Dashrath focusing on Data, Data preprocessing, Data visualization and Data Modelling.

George and Amrutha: GUI Part

Document and Presentation was done collectively by whole team.