Project Proposal: Student Performance Improvement Analysis

Overview: The objective of this project is to enhance student performance on standardized exams by identifying key factors that influence academic success. Using data-driven approaches, we aim to provide actionable insights and interventions that can help educators and policymakers improve student outcomes.

Objectives:

1. Identify key factors affecting students' exam scores.
2. Develop predictive models for student performance.
3. Segment students into clusters based on performance-related attributes.
4. Provide targeted recommendations for intervention strategies.

Data Sources: The dataset includes 6,607 records with variables categorized as follows:

- **Study & Academic Factors:** Hours Studied, Attendance, Previous Scores, Tutoring Sessions.
- **Parental & Socioeconomic Factors:** Parental Involvement, Family Income, Parental Education Level.
- **Personal & Lifestyle Factors:** Sleep Hours, Motivation Level, Physical Activity.
- **School Environment Factors:** Teacher Quality, Peer Influence, School Type, Distance from Home.
- **Outcome Variable:** Exam Score.

Methodology:

1. **Exploratory Data Analysis (EDA):**
    - Compute summary statistics.
    - Identify trends and patterns in student performance.
    - Perform data cleaning and feature engineering.
2. **Clustering Analysis:**
    - Apply K-Means to segment students.
    - Use silhouette scores to determine the optimal number of clusters.
3. **Predictive Modeling:**
    - Use Linear Regression, Decision Trees, Random Forest, and Gradient Boosting to predict student scores.
    - Evaluate models using RMSE, MAE, and R-squared metrics.
4. **Recommendations:**
    - Derive actionable insights based on model results.
    - Develop strategies for intervention, such as additional tutoring or resource allocation.

Expected Outcomes:

- A predictive model that estimates student performance based on selected features.
- Identification of high-risk students who may require additional support.
- Data-driven recommendations for educators and policymakers to improve academic success.
- Visualization dashboards to present key findings.