Project Title: Predicting Student Results Based on Study Hours Using Logistic Regression

Problem Statement:

The goal of this project is to develop a predictive model that can accurately forecast whether a student will pass or fail based on the number of study hours per week. This model aims to identify students at ris k of failing and provide insights for targeted interventions to improve educational outcomes.

Objective:

- To create a logistic regression model that predicts whether a student will pass or fail based on st udy hours per week.
- To identify the threshold of study hours that significantly impacts student performance.
- To support educators and policymakers in designing effective strategies to enhance educational quality and equity.

Data Source:

- School Records: Attendance and grades of student.
- Surveys: Information on study habits of student , study time of student ,total study hours per week of student

Methodology:

1. **Data Collection**: Gather data from the sources mentioned above.
2. **Data Preprocessing**:
    - Handle missing values and outliers.
    - Normalize and standardize the data.
    - Encode categorical variables.
3. **Feature Selection**: Identify and select the most relevant features that impact student performan ce.
4. **Model Development**:
    - Split the data into training and testing sets.
    - Train a logistic regression model on the training set.
5. **Model Evaluation**:
    - Evaluate the model using metrics such as encoding dataset, feature-scaling ,accuracy & confusion matrix.

Focus on SDG 4: Quality Education

This project aligns with the United Nations Sustainable Development Goal (SDG) 4: Quality Education. By predicting student results and identifying at-risk students, the project aims to:

Ensure inclusive and equitable quality education for all students

- Promote lifelong learning opportunities by providing targeted support and interventions.
- Assist educators and policymakers in making data-driven decisions to improve educational outcomes