■ Student Performance Prediction (Pass/Fail)

■ Project Objective

The objective of this project is to build a Supervised Machine Learning model that predicts whether a student is likely to pass or fail based on academic and lifestyle factors such as study hours, attendance, prior grades, and extracurricular activities. This system can help educators and parents identify students who may need additional support.

■ Dataset

We used a synthetic dataset (student_performance_extended.csv) containing 500+ records with the following features:

Feature	Description
student_id	Unique ID for each student
gender	Gender (M/F)
age	Age of student
study_hours_per_week	Average study hours per week
attendance_percent	Class attendance percentage
prior_grade	Previous academic grade (0-100)
assignments_submitted	Number of assignments submitted
family_support	Family support (Yes/No)
internet_access	Internet availability at home
extracurricular	Participation in extracurricular activities
health_score	Self-reported health condition (0-100)
sleep_hours	Average sleep per day
pass	Target variable → 1 = Pass, 0 = Fail

■ Methodology

1. Data Preprocessing: Encoded categorical variables (Yes/No, M/F) into numeric form and normalized inputs.
2. Model Selection: Used Random Forest Classifier due to its accuracy and ability to handle categorical + numerical data.
3. Training & Testing: Split dataset into train/test sets, trained with 100 estimators, and evaluated performance.
4. Deployment: Built a Streamlit dashboard for real-time predictions with visual feedback.

■ Results

• Achieved high prediction accuracy on the test dataset.
• Streamlit app provides:
- Real-time pass/fail prediction

- Probability confidence score
- Visual feedback via gauge meter and bar chart

■ Conclusion

This project demonstrates how Supervised Machine Learning can be applied in the education domain to predict student performance. The app can serve as an early warning system for struggling students, helping educators take corrective action in time.