ABSTRACT:

Educational institutions often struggle to identify students who may underperform at an early stage. This project proposes a machine learning–based system to predict student academic performance using historical data. A Flask-based web application is developed where educators can upload CSV datasets for analysis. The system preprocesses the data and trains multiple machine learning models, including Logistic Regression, Random Forest, and XGBoost. Model performance is evaluated using standard metrics and displayed through an interactive dashboard with visualizations. The proposed solution helps educators make data-driven decisions and identify at-risk students early.

INTRODUCTION:

Educational institutions generate large amounts of student data, but traditional evaluation methods lack predictive capabilities. As a result, identifying students who may underperform at an early stage becomes challenging. Machine learning provides efficient techniques to analyze historical data and predict academic performance. This project develops a Flask-based web application that uses multiple machine learning models to predict student performance and display results through an interactive dashboard, helping educators make data-driven decisions.

OBJECTIVE:

- The objective of this project is to develop a machine learning–based web application that predicts student performance using historical academic data.

- The system enables educators to upload student datasets in CSV format and automatically trains multiple machine learning models to analyze and compare prediction accuracy.

- The results are displayed through an interactive dashboard.

METHODOLOGY:

Data Input

- CSV file containing numerical features such as attendance, assignment scores, and exam marks

- Optional text-based data with labeled outcomes

Data Preprocessing

- Cleaning and handling missing numerical data

- Encoding class labels

- Vectorization of text data where applicable

Model Training

- Logistic Regression

- Random Forest Classifier

- XGBoost Classifier

Model Evaluation

- Accuracy

- Precision

- Recall

- F1-Score

Visualization

- Feature importance bar charts

- Class distribution pie charts

Deployment

- Flask-based web application

- User-friendly HTML templates (index.html, results.html)

Expected Outcomes :

- Identification of the best-performing machine learning model
- Generation of classification reports for each model
- Visual insights into factors affecting student performance
- Early identification of at-risk students

Applications:

- Academic performance monitoring
- Early intervention for weak students
- Data-driven decision making for educators
- Educational analytics

Tools & Technologies Used:

- **Programming Language:** Python
- **Framework:** Flask
- **Machine Learning Libraries:** Scikit-learn, XGBoost
- **Frontend:** HTML, CSS
- **Visualization:** Matplotlib, Seaborn

Conclusion:

This project provides a practical implementation of machine learning in the education domain. By combining predictive analytics with a web-based interface, the system offers an efficient and user-friendly solution for student performance prediction. It demonstrates the effective use of artificial intelligence to support educational institutions in improving academic outcomes.