

# Student Performance Prediction using Machine Learning

---

## Objective:

To develop a machine learning model that can predict whether a student will pass or fail based on their study hours, attendance, and previous scores.

## Introduction:

Artificial Intelligence and Machine Learning are transforming how data is analyzed and predictions are made. In the education sector, predicting student performance can help educators provide timely intervention. This project uses a decision tree classifier to predict student performance.

## Tools and Technologies Used:

- Programming Language: Python
- Libraries: Pandas, Scikit-learn, Matplotlib
- IDE: Jupyter Notebook / Google Colab

## Algorithm Used:

Decision Tree Classifier – A supervised machine learning algorithm that is simple and effective for classification problems.

## Methodology:

1. Data Collection: Synthetic dataset with attributes like study hours, attendance, and previous marks.
2. Data Preprocessing: Clean and format data.
3. Model Training: Use decision tree classifier.
4. Prediction: Test the model on new data.
5. Evaluation: Check accuracy and performance.

## Sample Code:

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier
```

```

from sklearn.metrics import accuracy_score

# Sample data
data = {
    'Study_Hours': [2, 3, 4, 5, 1, 6, 2, 3, 4, 5],
    'Attendance': [60, 70, 80, 90, 50, 95, 55, 65, 85, 92],
    'Previous_Score': [45, 55, 60, 70, 30, 90, 35, 50, 65, 80],
    'Result': ['Fail', 'Pass', 'Pass', 'Pass', 'Fail', 'Pass', 'Fail', 'Pass', 'Pass', 'Pass']
}

df = pd.DataFrame(data)

# Features and target
X = df[['Study_Hours', 'Attendance', 'Previous_Score']]
y = df['Result']

# Train-test split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=42)

# Model training
model = DecisionTreeClassifier()
model.fit(X_train, y_train)

# Prediction
y_pred = model.predict(X_test)

# Accuracy
print("Accuracy:", accuracy_score(y_test, y_pred))

```

### Result:

The model was able to predict student outcomes with reasonable accuracy. It shows that factors like attendance and previous performance significantly affect outcomes.

### Conclusion:

The project demonstrates a basic application of machine learning in the field of education. Predictive models like this can help educators support at-risk students early.

**Future Scope:**

- Include more parameters like parent involvement, health, etc.
- Use larger real-world datasets.
- Apply advanced algorithms (e.g., Random Forest, SVM).