LAB # 7

OPEN ENDED LAB

OBJECTIVE

To apply the foundational machine learning concepts learned in Labs 1–6 to design and implement a small, custom ML project in Python.

The goal is to integrate multiple concepts learned so far into a single, working prototype that performs real-world data analysis or prediction.

PROJECT DESCRIPTION

This is an open-ended lab where you will be told individually to work on the following project ideas and build a working ML solution in Python.
Your project should demonstrate the integration of at least two machine learning models and appropriate data preprocessing, visualization, and evaluation.

PROJECT IDEAS

1. Student Performance Predictor
Create a Python application that predicts whether a student will pass or fail based on study hours, attendance, and assignment scores.
Dataset: Student Performance Data Set (UCI)

- **Integration:**
  - Apply **data preprocessing** to handle missing data and scaling.
  - Use **Linear Regression** to predict final marks and **Logistic Regression** or **Decision Tree** to classify pass/fail.
  - Compare models using accuracy, precision, and recall.

---

2. Employee Attrition Analyzer
Analyze a dataset of employee records to predict whether an employee is likely to leave the company.
Dataset: IBM HR Analytics Employee Attrition & Performance
Integration:

- Clean and preprocess categorical and numerical data.
- Train and compare **Logistic Regression**, **Decision Tree**, and **Random Forest** models.
- Evaluate results using confusion matrix, F1 score, and feature importance visualization.

DELIVERABLES

- All source code for your project, clearly commented with Project name and roll no on top of the notebook.
- Your dataset description and preprocessing steps.
- The ML models used and how they were integrated.
- Model evaluation (metrics, confusion matrix, etc.).
- Code must be uploaded on GitHub and link paste on QOBE portal.