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. House Price & Category Prediction
Build a model that predicts both the **price (continuous)** and **price category (discrete)** of houses based on features like area, number of rooms, and location.
Dataset: House Prices: Advanced Regression Techniques
- **Integration:**
  - Use **Linear Regression** for price prediction.
  - Use **Random Forest** or **SVM** to classify the house as *Low*, *Medium*, or *High* value.
  - Include preprocessing (feature scaling, outlier removal) and performance comparison (R^2, accuracy).

2. Health Risk Classification
Design a model that predicts the likelihood of a patient having a medical condition (e.g., diabetes or heart disease) based on health indicators.
Dataset: Pima Indians Diabetes Database
- **Integration:**
  - Perform **data preprocessing** (handle missing values, normalize numeric data).
  - Apply **SVM** and **Decision Tree** for classification.
  - Compare both models' results and discuss overfitting and accuracy trade-offs.

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.