Classification Exercise

Datasets:

1. Iris Dataset

2. Breast Cancer Wisconsin (Diagnostic) Dataset

3. Titanic Dataset

4. Pima Indians Diabetes Dataset

Data Loading and Preprocessing:

- Load the dataset.
- Handle missing values.
- Normalize/scale the features if necessary.
- Split the data into training and test sets.

Implement Classification Algorithms:

- Logistic Regression
- K-Nearest Neighbors (KNN)
- Support Vector Machines (SVM)
- Decision Tree
- Random Forest
- XGBoost

Model Evaluation:

Use evaluation metrics such as accuracy, precision, recall, F1 score, and confusion matrix to compare the performance of different models.

Hyperparameter Tuning:

Use GridSearchCV or RandomizedSearchCV to tune hyperparameters for the best models.

Visualization:

Visualize the results using plots such as confusion matrices, ROC curves, and feature importance.