

Assignment 1

Name: Omkar Chavan

Approach and Methodologies:

1. Exploratory Data Analysis (EDA):

Objective: Understand the structure and characteristics of the Iris dataset.

Steps Taken:

Loaded the dataset using Pandas.

Displayed basic statistics to get an overview of the data.

Utilized histograms to visualize the distribution of each numerical feature.

Employed box plots to highlight variations in features across different species.

Utilized a pair plot to visualize relationships between numerical features.

2. Data Science Task:

Objective: Predict Iris species based on their characteristics.

Steps Taken:

Encoded the categorical target variable ('Species') using Label Encoding.

Split the dataset into features (X) and target variable (y).

Split the dataset into training and testing sets.

Chose the Decision Tree classifier for its simplicity and interpretability.

Trained the model on the training set and evaluated on the testing set.

Utilized metrics such as accuracy, precision, recall, and confusion matrix for evaluation.