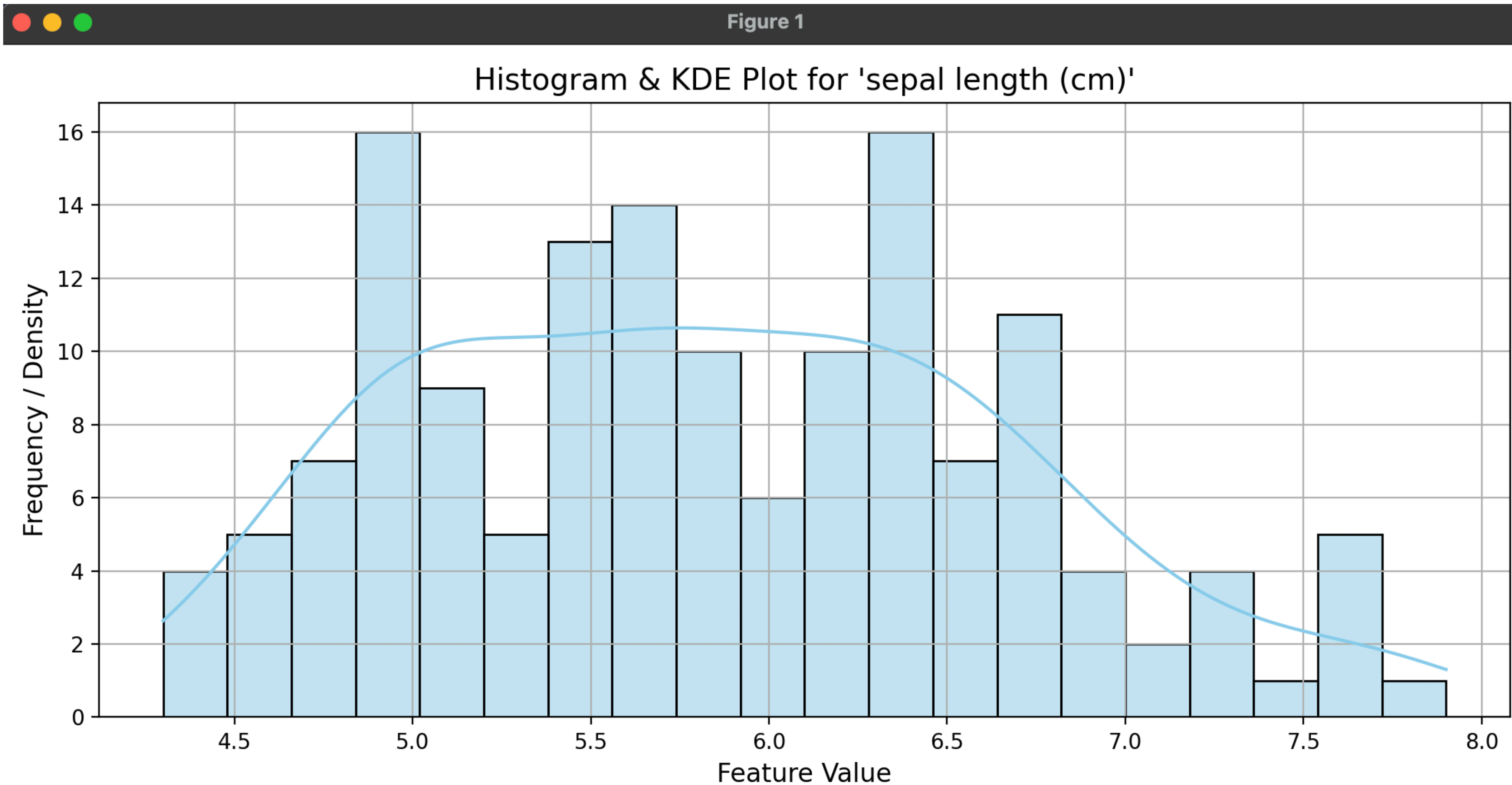# Data Visualisation

Er. Suman Shrestha
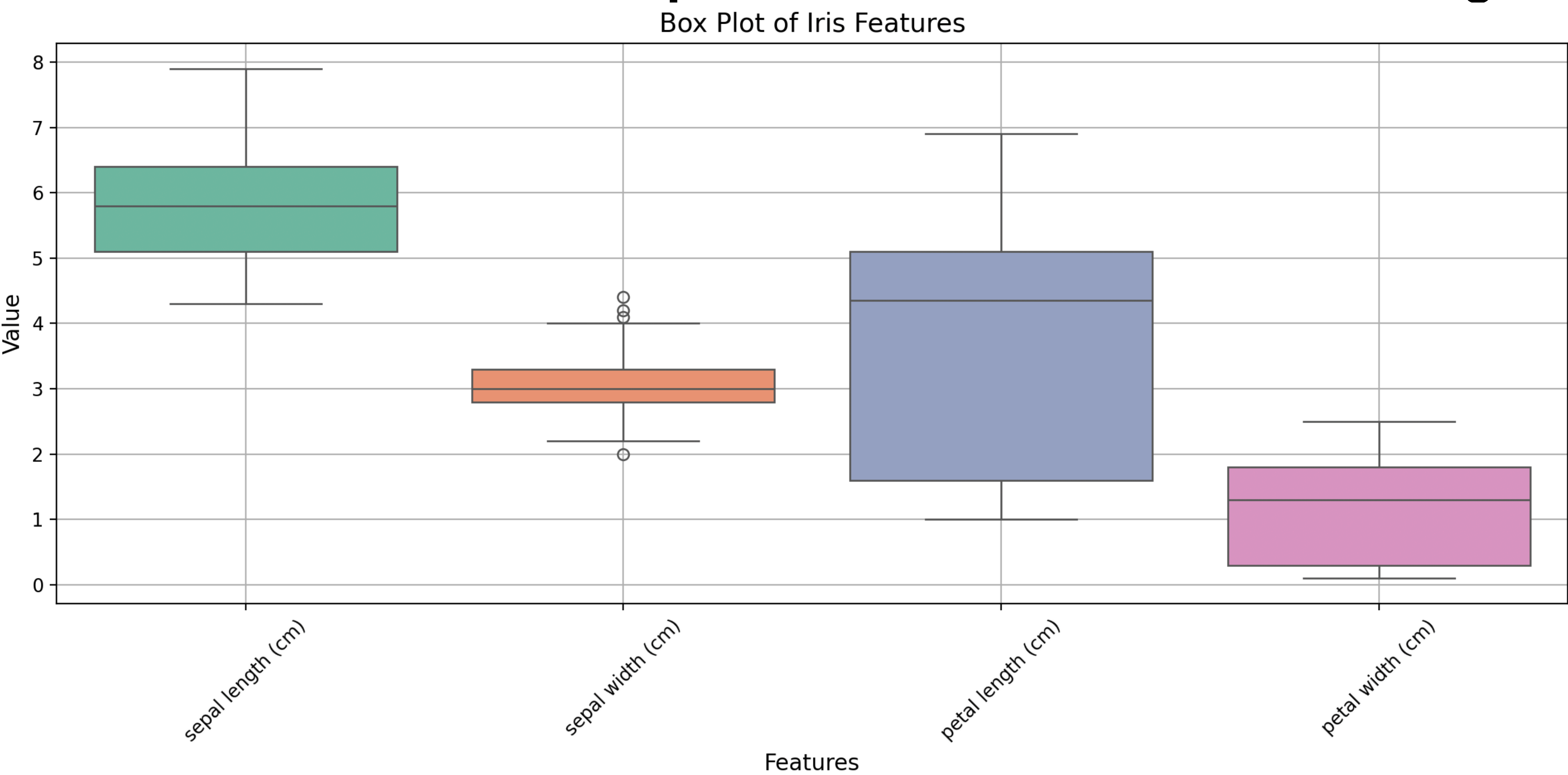
Histogram
To analyze feature distribution, skewness, outliers
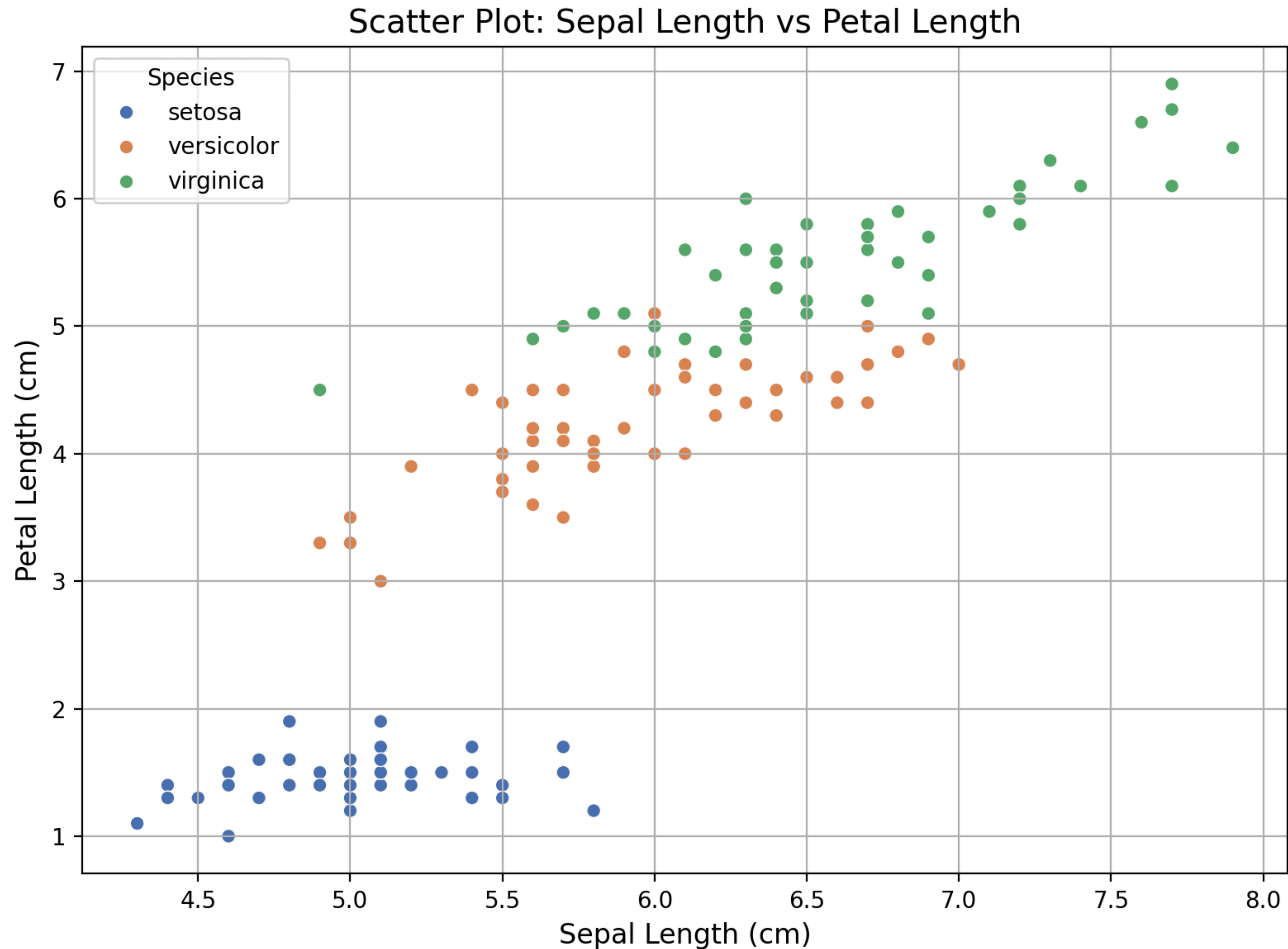
Box Plot
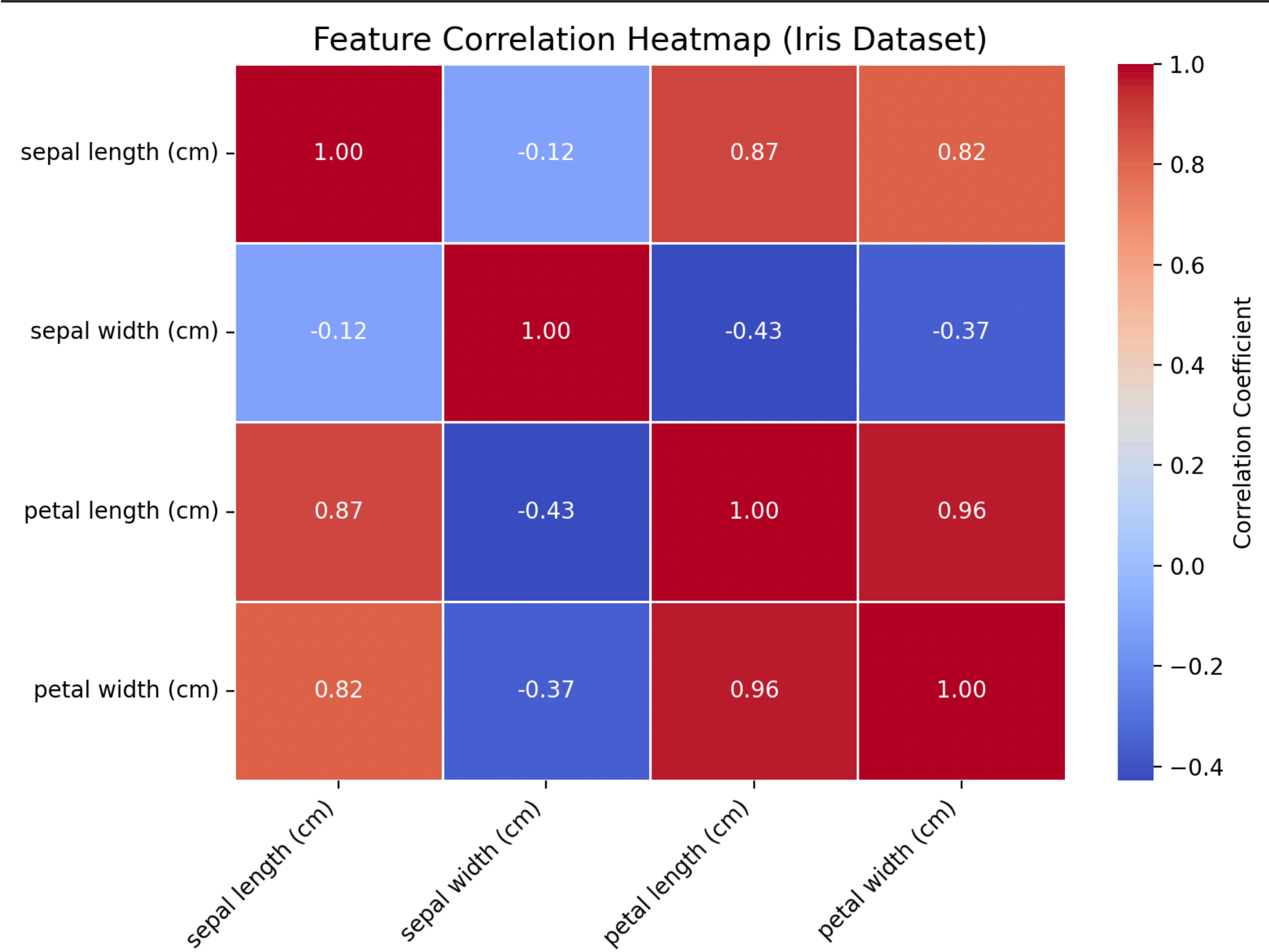## To detect outliers and compare distributions between categories.

Box Plot of Iris Features

Scatter Plot
To observe relationships between features
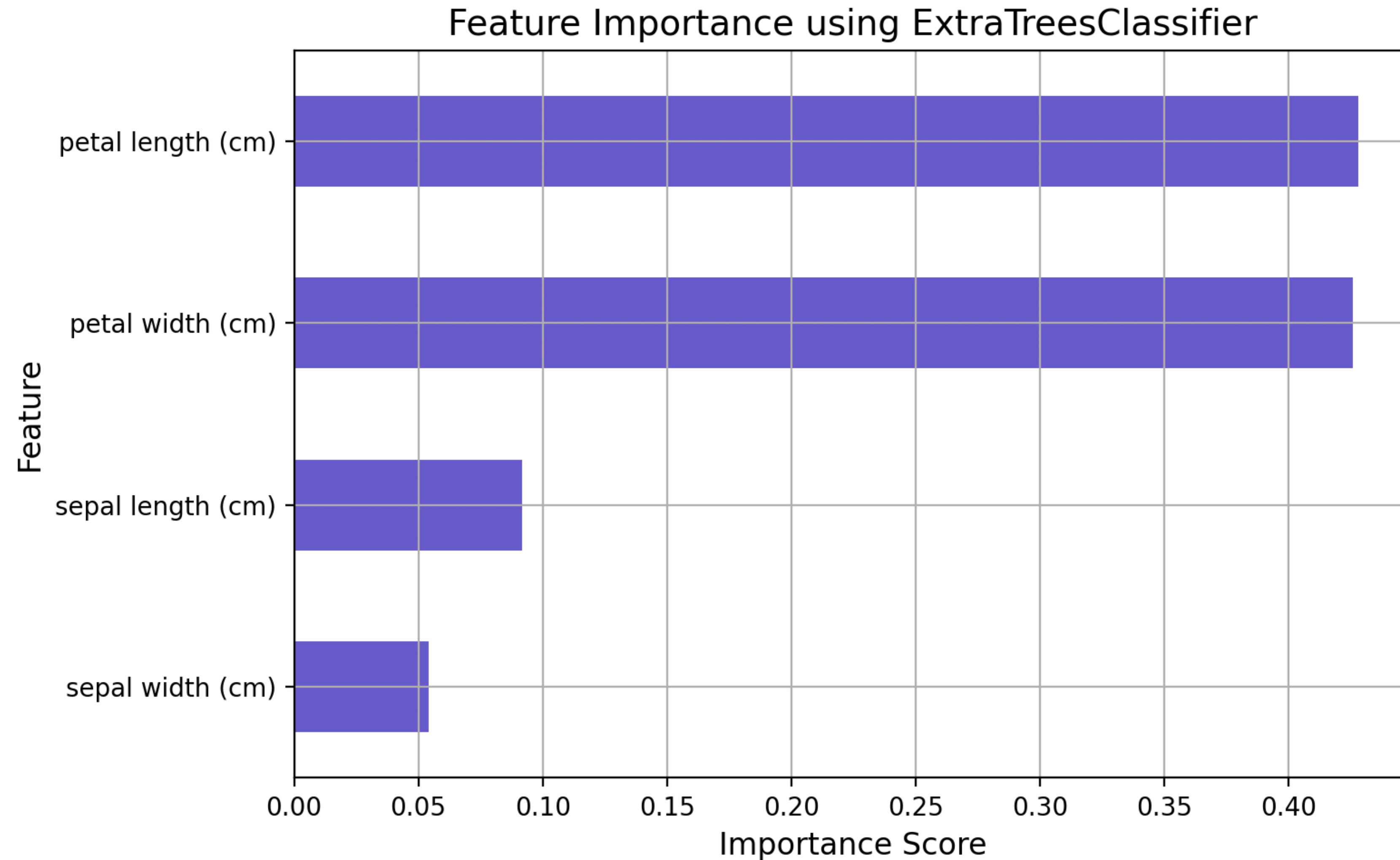
Heatmap
To detect multicollinearity among features.

Feature Correlation Heatmap (Iris Dataset)

Class Distribution Plot
To check label balance (for classification tasks).
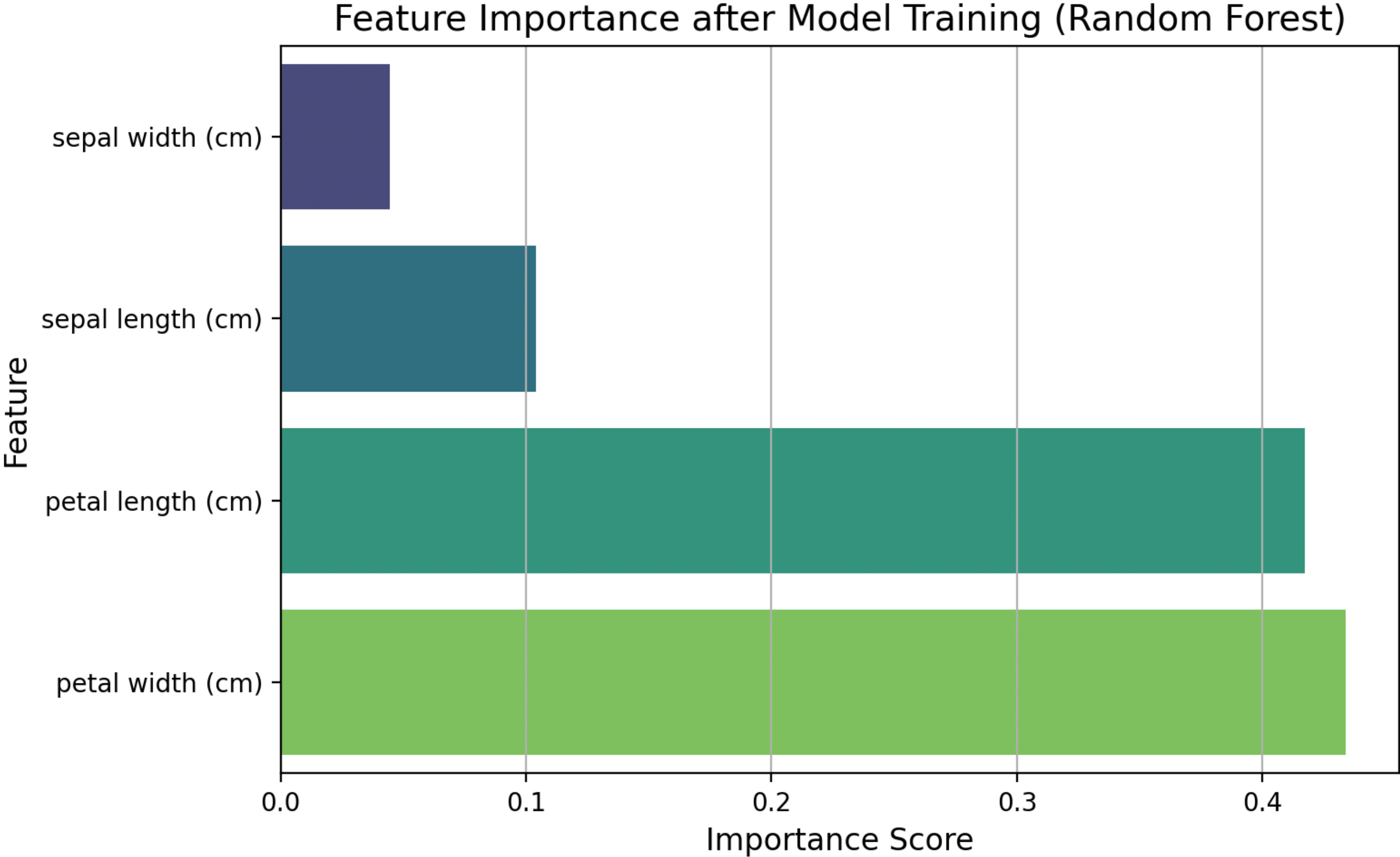

Class Distribution in Iris Dataset
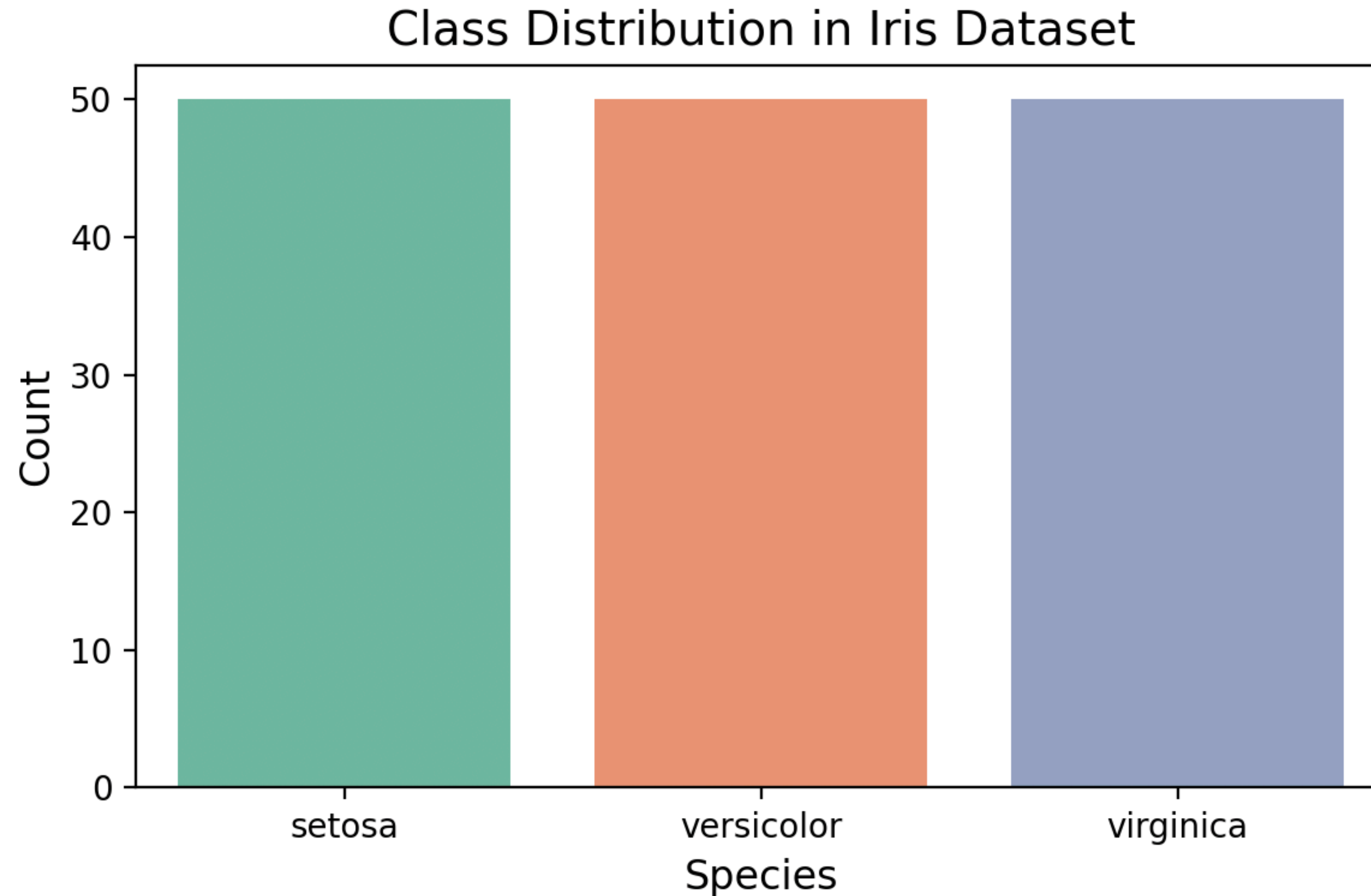
Feature Importance (before training)
To identify potentially useful features before training.
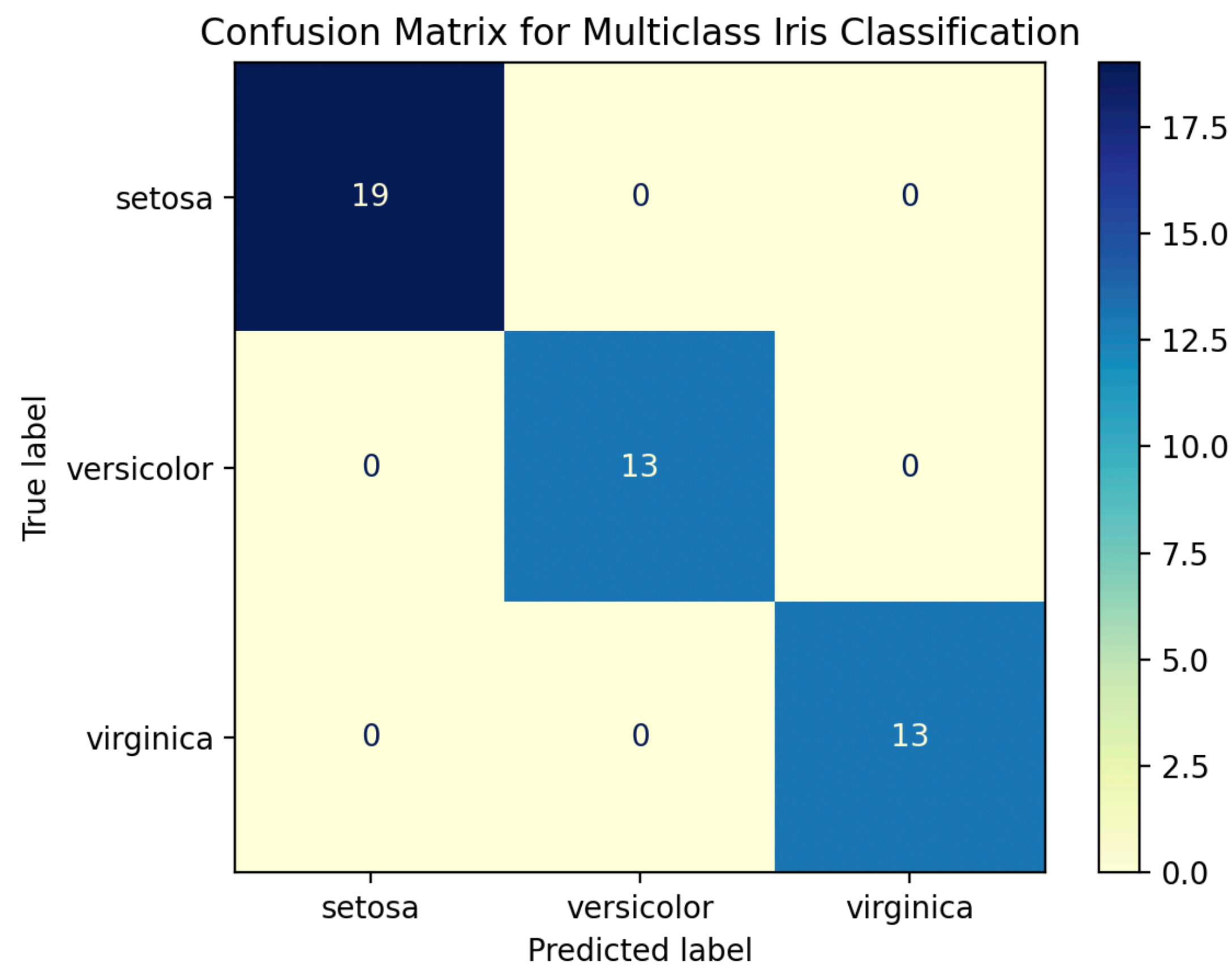

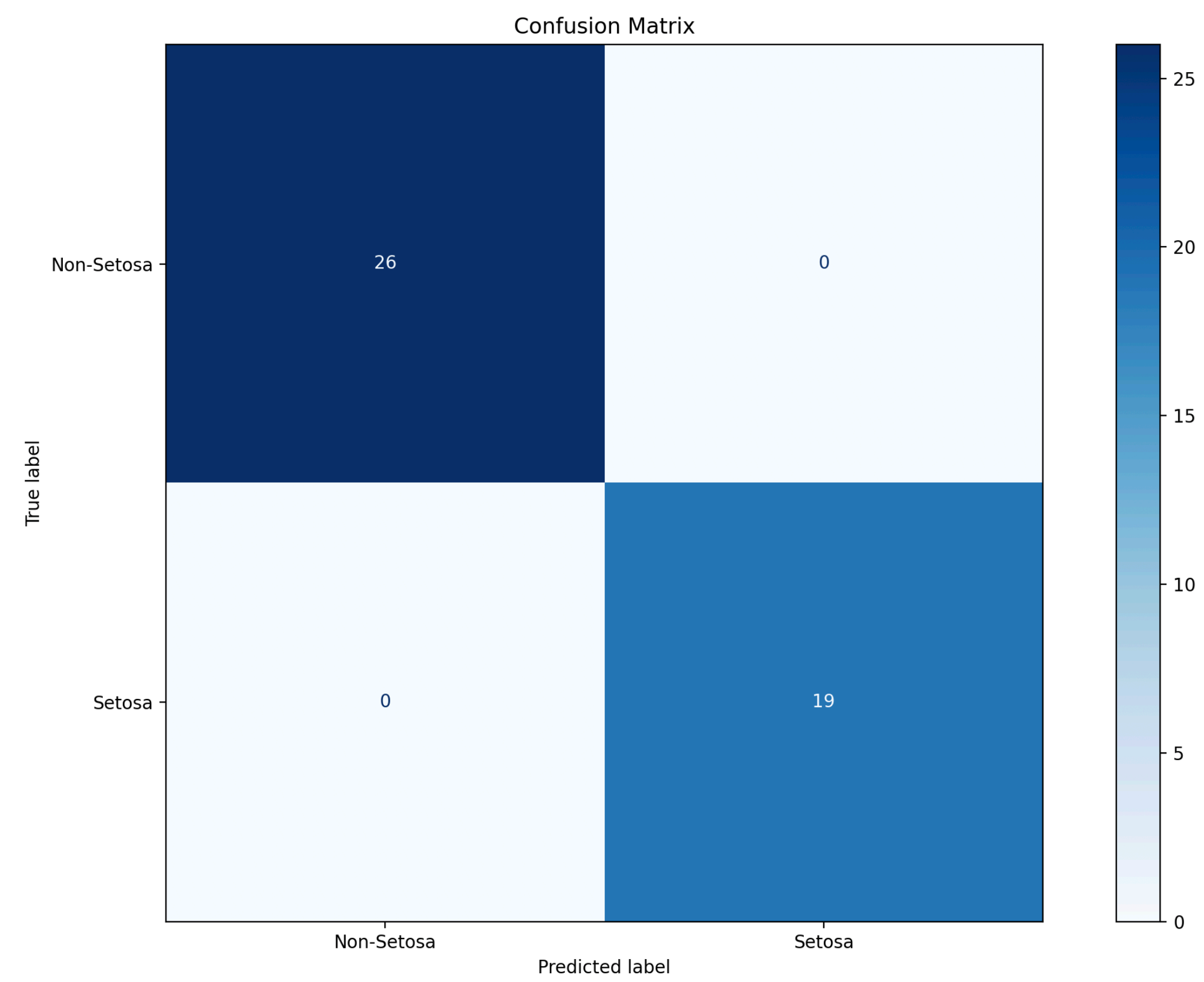Feature Importance using ExtraTreesClassifier

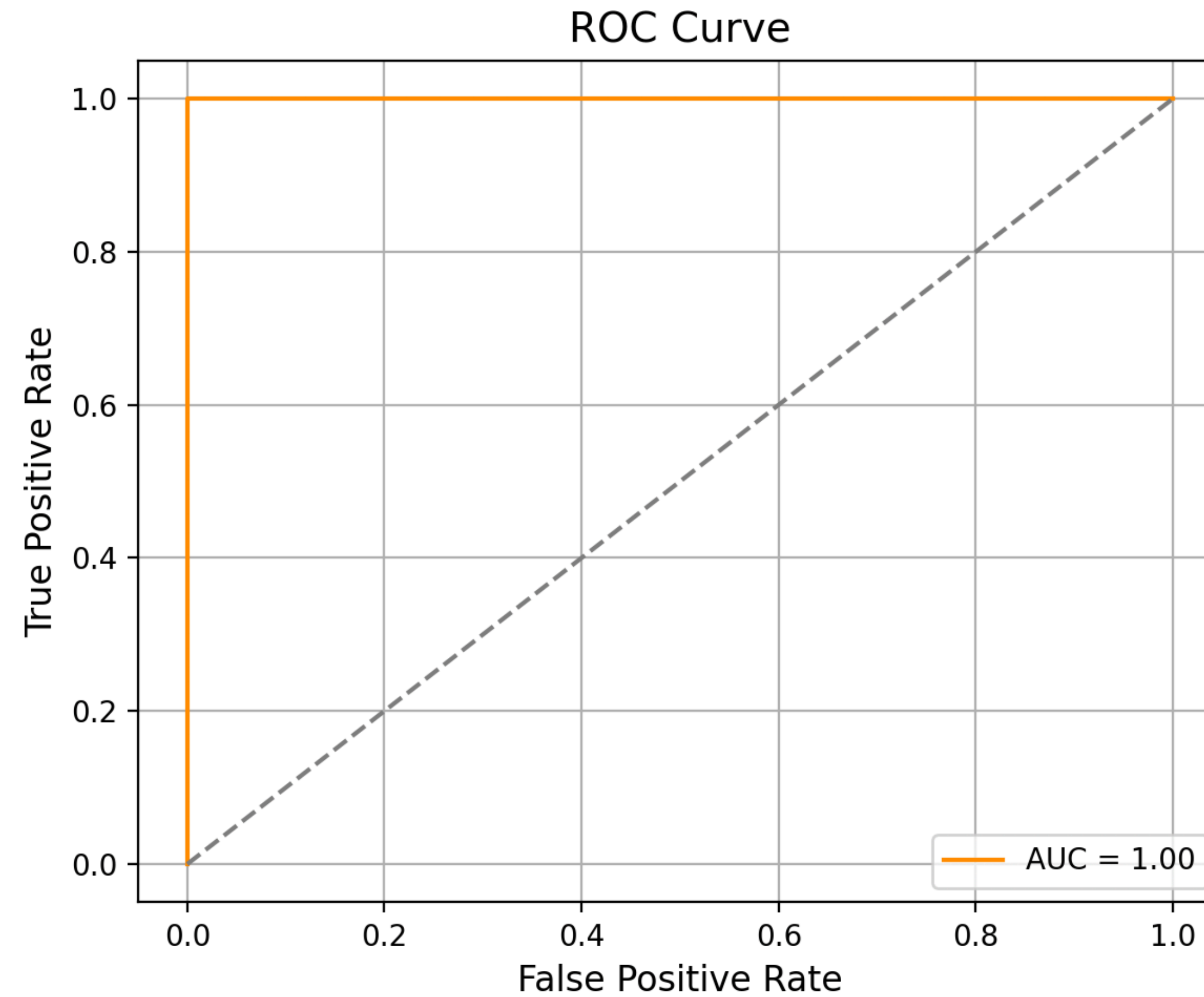# Feature importance after training

Confusion Matrix
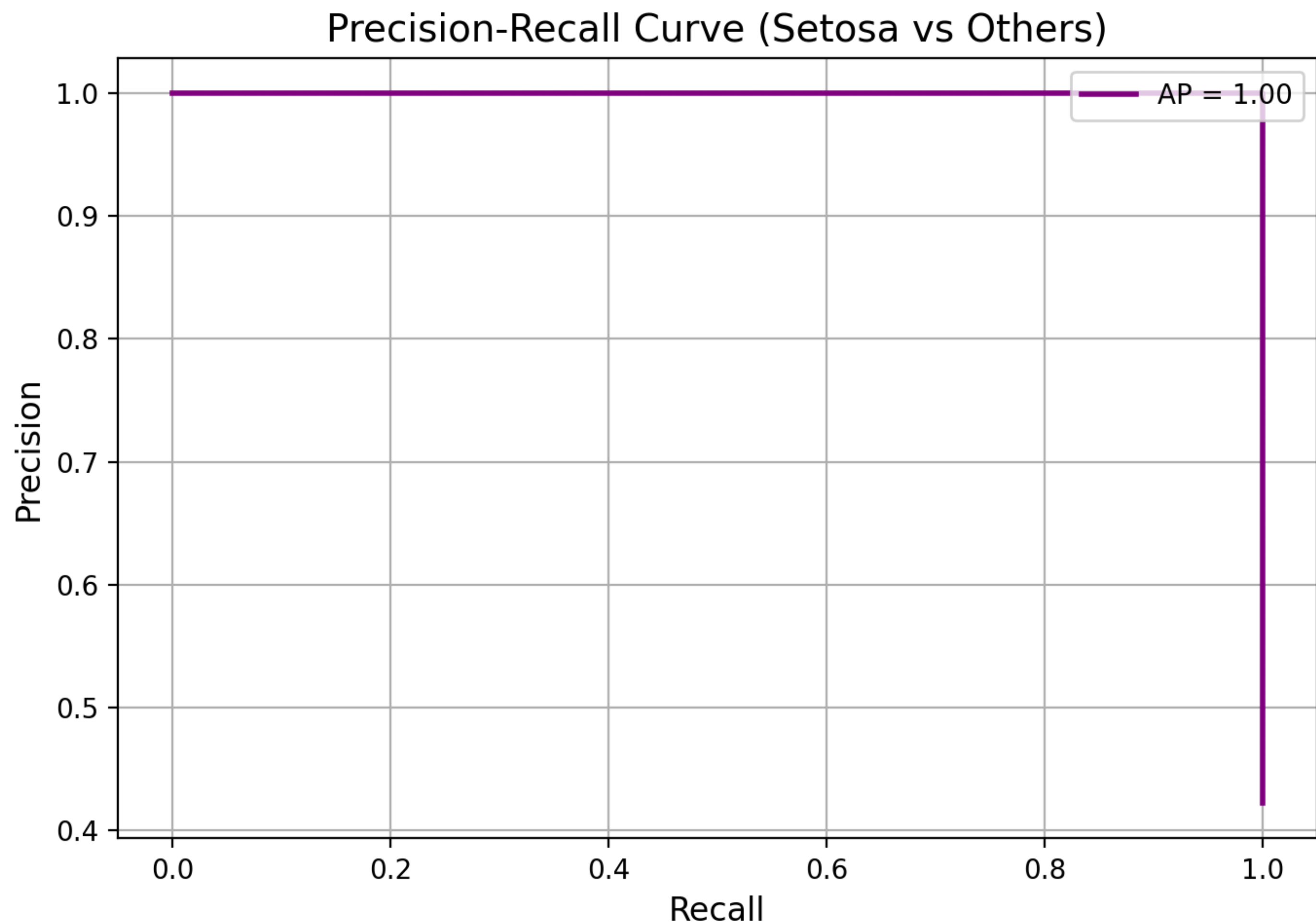## To evaluate classification performance in detail.

ROC Curve with AUC
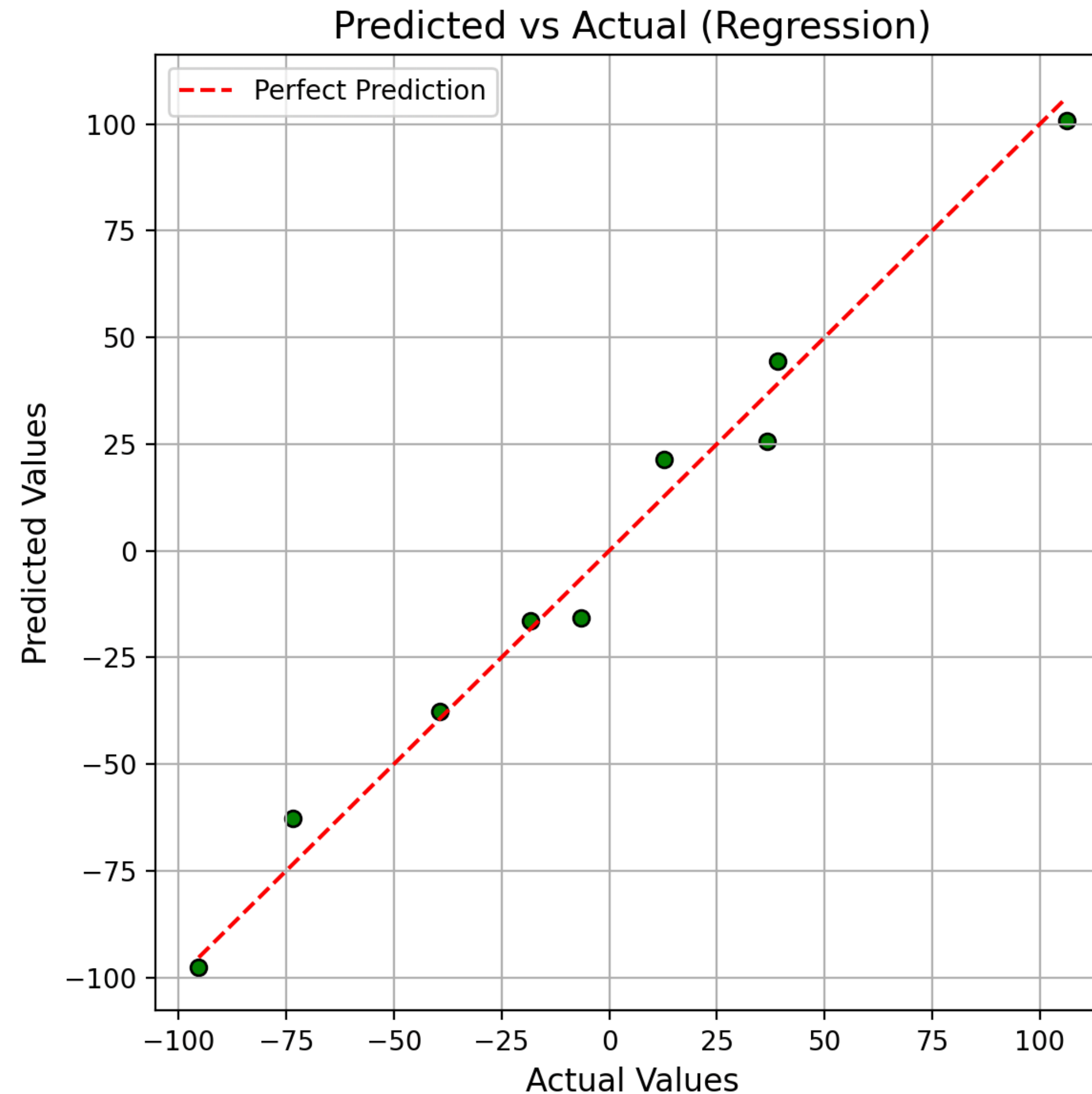## To monitor binary classifier performance.

Precision-Recall Curve
Useful especially when data is imbalanced.

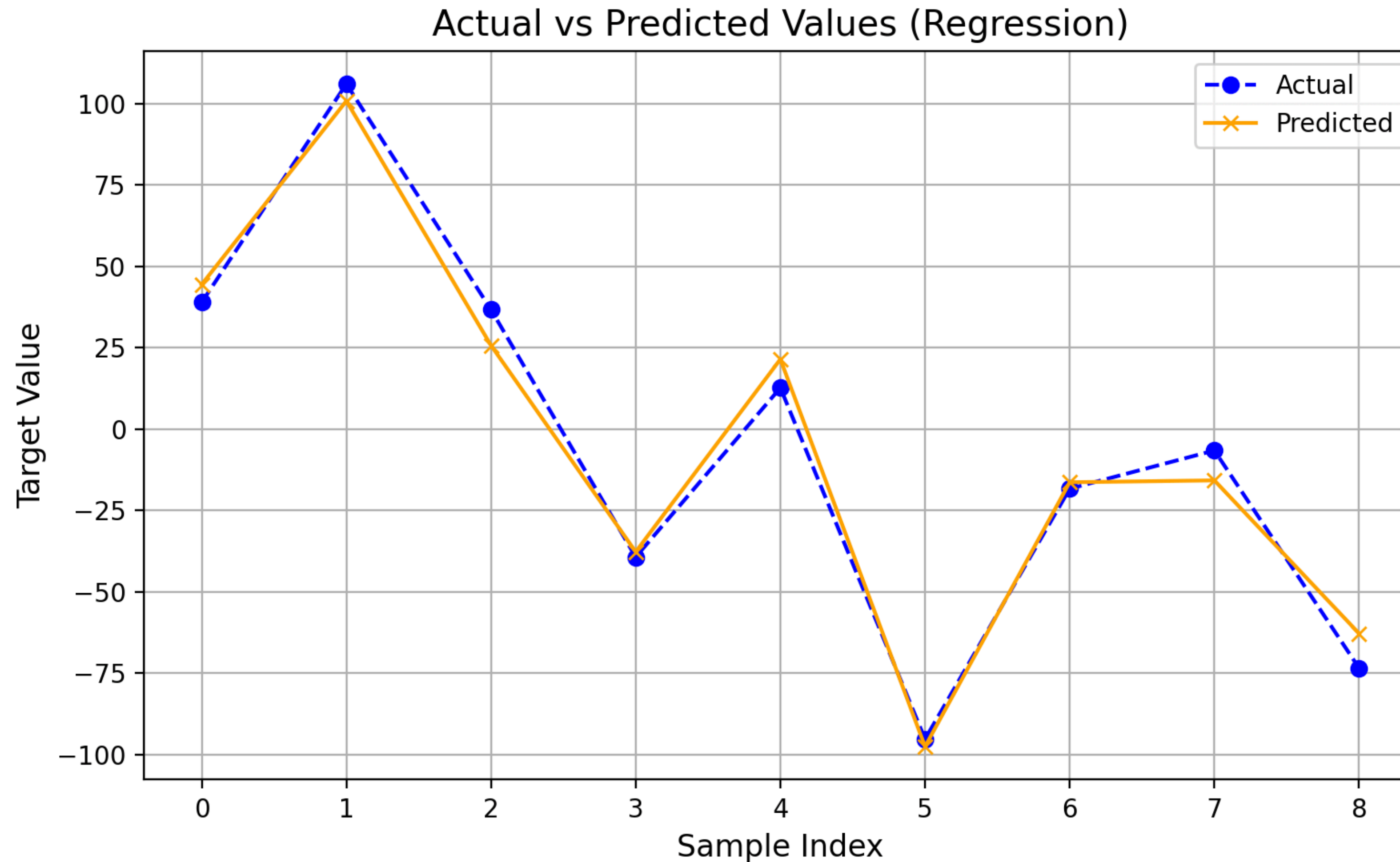Precision-Recall Curve (Setosa vs Others)

Predicted vs. Actual plot
For regression: shows how closely predictions match actual values.

Predicted vs. Actual plot
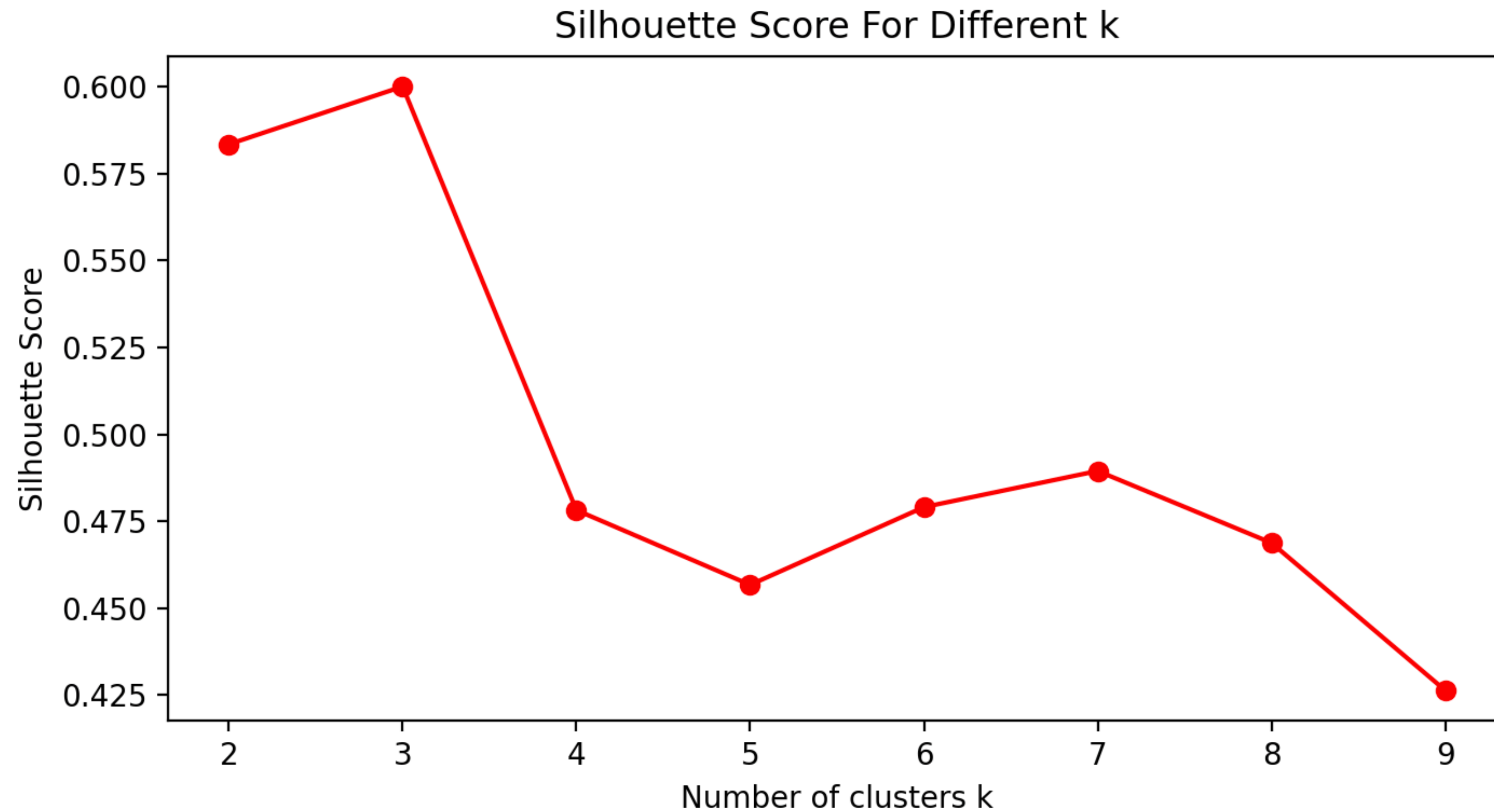## For regression: shows how closely predictions match actual values.

Actual vs Predicted Values (Regression)

Elbow Curve (for k-means)
To determine optimal number of clusters.
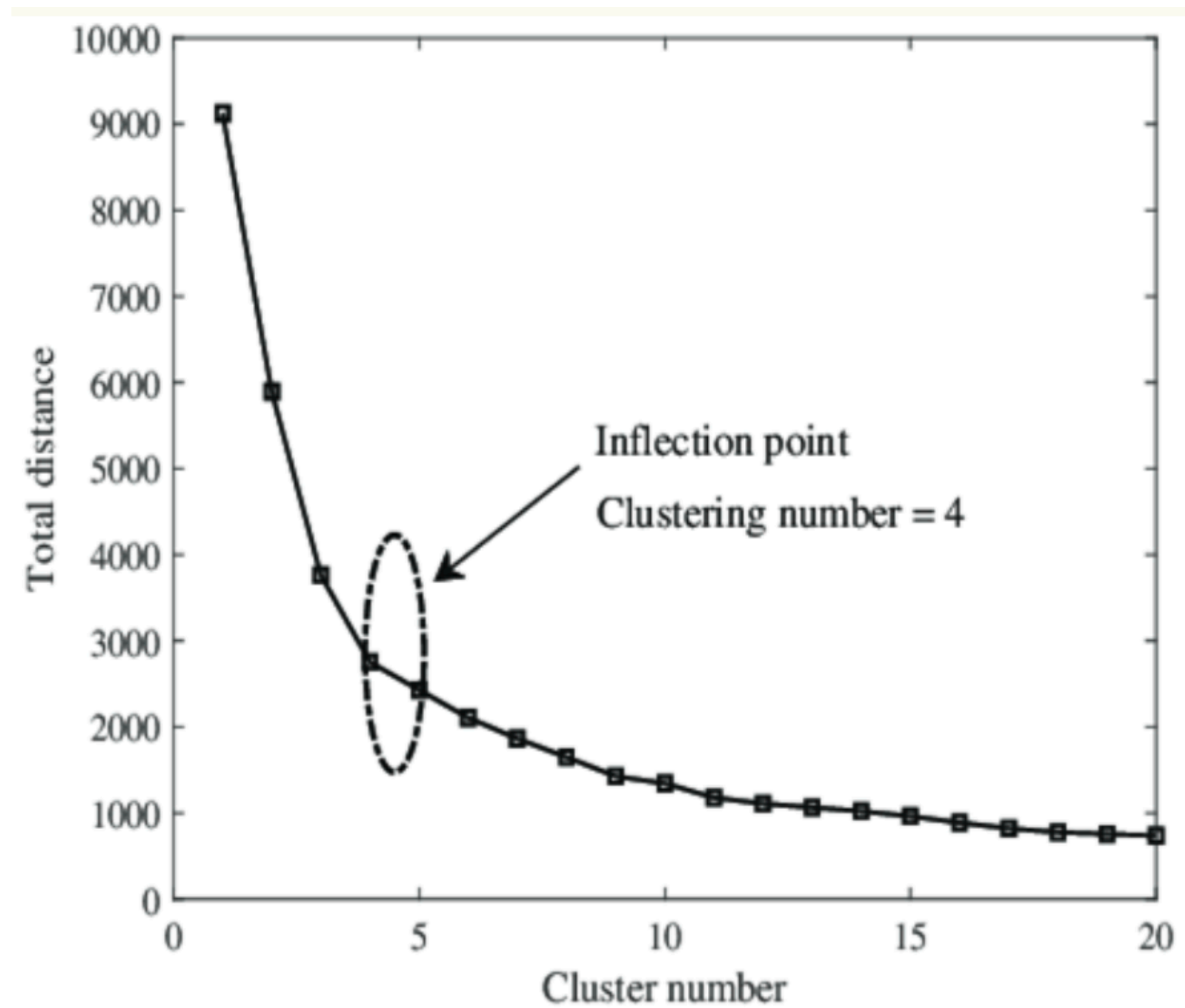
Silhouette Plot / Score
To Evaluate Clustering Quality

Silhouette Score For Different k

Dendrogram (for hierarchical clustering)
To visualize nested clustering structure.