Exercise: K-means Clustering and K-NN Classification on Iris Dataset

Objective

- Perform K-means clustering on the Iris dataset, including determining the optimal number of clusters using the elbow method and calculating the silhouette score.
- Perform K-NN classification on the Iris dataset, including splitting the dataset into training and testing sets, and evaluating the classifier's performance.

Tasks

Part 1: K-means Clustering

1. Load the Iris Dataset:
   - Use the load_iris function from sklearn.datasets to load the dataset.
2. Determine the Optimal Number of Clusters:
   - Use the elbow method to find the optimal number of clusters.
   - Plot the within-cluster sum of squares (WCSS) for different values of k.
3. Perform K-means Clustering:
   - Use the KMeans class from sklearn.cluster to perform clustering with the optimal number of clusters.
4. Calculate the Silhouette Score:
   - Use the silhouette_score function from sklearn.metrics to calculate the silhouette score for the clustering.
5. Visualize the Clusters:
   - Create a scatter plot to visualize the clusters.

Part 2: K-NN Classification

1. Split the Dataset:
   - Use the train_test_split function from sklearn.model_selection to split the dataset into training and testing sets.
2. Perform K-NN Classification:
   - Use the KNeighborsClassifier class from sklearn.neighbors to perform K-NN classification with 5 neighbors.
3. Evaluate the Classifier:
   - Calculate the accuracy of the classifier on the test set.
   - Optionally, display a confusion matrix to show the performance in more detail.