

Task 16

Machine Learning

Upload the .py or .ipynb extension file to GitHub public repo "100DaysofBytewise" and share the link in the submission form by July 31, 2024.

Dataset : Iris

1. Implementing K-Means Clustering

Task: Apply K-Means clustering to the Iris dataset and visualize the clusters using a scatter plot of two features. Evaluate the clustering by comparing it to the actual species labels.

2. Choosing the Optimal Number of Clusters

Task: Use the Elbow Method and Silhouette Score to determine the optimal number of clusters for the Iris dataset. Visualize the Silhouette Score for different cluster counts.

3. Cluster Visualization with PCA

Task: Use Principal Component Analysis (PCA) to reduce the Iris dataset to two dimensions. Visualize the clusters obtained from K-Means clustering in the PCA-reduced space.

4. Hierarchical Clustering: Dendrogram

Task: Implement hierarchical clustering using the Iris dataset. Plot a dendrogram to visualize the clustering process and choose an appropriate number of clusters.

5. Comparing Clustering Algorithms

Task: Compare the performance of K-Means and Agglomerative Hierarchical Clustering on the Iris dataset. Discuss the strengths and weaknesses of each approach based on the results.