**NAME : MANGESH A. GHADWAJE**

**ROLL NO:24**

**BATCH : B2**

**COURSE: ML PRACTICAL**

**Assginment No. 7**

**Problem Statement :**

**Implement K-Means Clustering on the following data set We have given a collection of 8 points. P1=[0.1,0.6] P2=[0.15,0.71]P3=[0.08,0.9] P4=[0.16,0.85] P5=[0.2,0.3] P6=[0.25,0.5] P7=[0.24, 0.1] P8=[0.3,0.2]. Perform the k-mean clustering with initial centroids as m1=P1=Cluster#1=C1 and m2=P8=cluster#2=C2. Answer the following a) Which cluster does P6 belongs to? b) What is the population of cluster around m2? c) What is updated value of m1 and m2?**

**Code :**

***import pandas as pd***

***import numpy as np***

***import matplotlib.pyplot as plt***

***data = pd.read\_csv('Mall\_customers.csv')***

***features = data[['Annual Income (k$)', 'Spending Score (1-100)']]***

***data\_points = (features - features.min()) / (features.max() - features.min()).values***

***data\_points = data\_points.values***

***centroid1 = data\_points[0]***

***centroid2 = data\_points[-1]***

***while True:***

***cluster1 = []***

***cluster2 = []***

***for point in data\_points:***

***dist1 = np.linalg.norm(point - centroid1)***

***dist2 = np.linalg.norm(point - centroid2)***

***if dist1 < dist2:***

***cluster1.append(point)***

***else:***

***cluster2.append(point)***

***new\_centroid1 = np.mean(cluster1, axis=0) if cluster1 else centroid1***

***new\_centroid2 = np.mean(cluster2, axis=0) if cluster2 else centroid2***

***if np.array\_equal(new\_centroid1, centroid1) and np.array\_equal(new\_centroid2, centroid2):***

***break***

***centroid1, centroid2 = new\_centroid1, new\_centroid2***

***cluster1 = np.array(cluster1)***

***cluster2 = np.array(cluster2)***

***plt.scatter(cluster1[:, 0], cluster1[:, 1], color='blue', label='Cluster 1')***

***plt.scatter(cluster2[:, 0], cluster2[:, 1], color='green', label='Cluster 2')***

***plt.scatter(centroid1[0], centroid1[1], color='red', marker='X', s=200, label='Centroid 1')***

***plt.scatter(centroid2[0], centroid2[1], color='red', marker='X', s=200, label='Centroid 2')***

***plt.title('K-Means Clustering')***

***plt.xlabel('Annual Income (scaled)')***

***plt.ylabel('Spending Score (scaled)')***

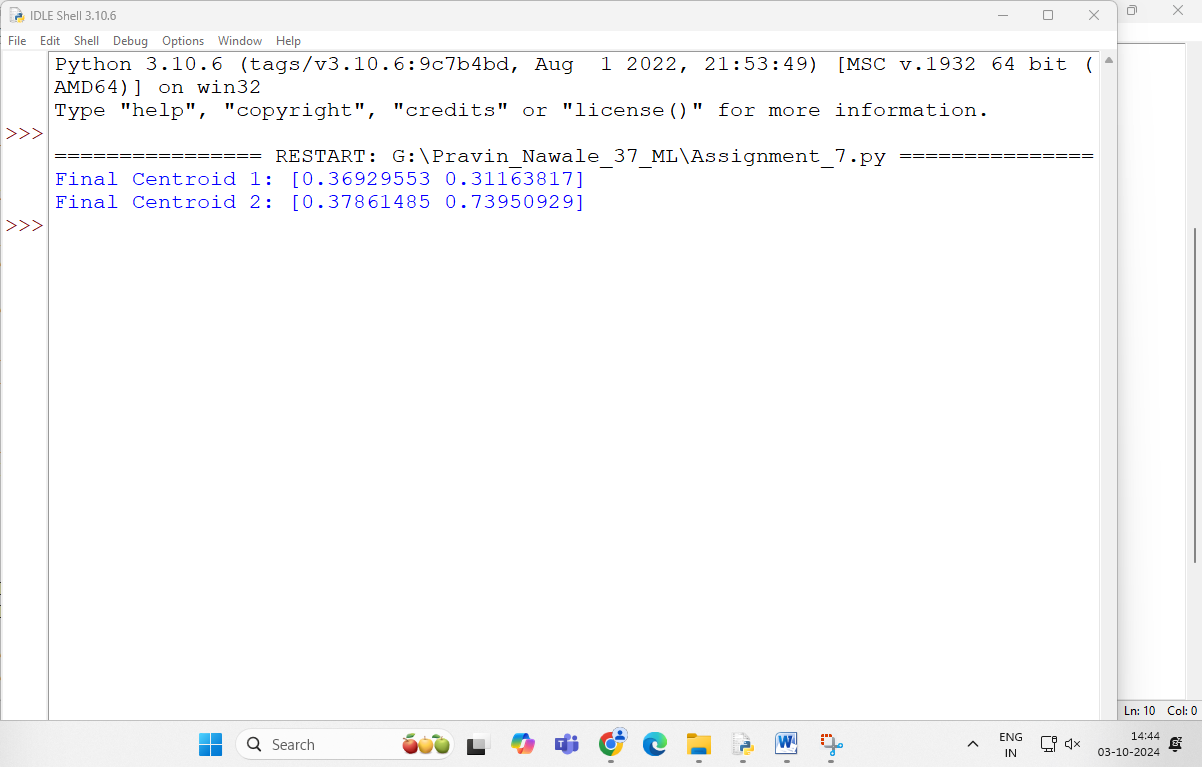
***plt.legend()***

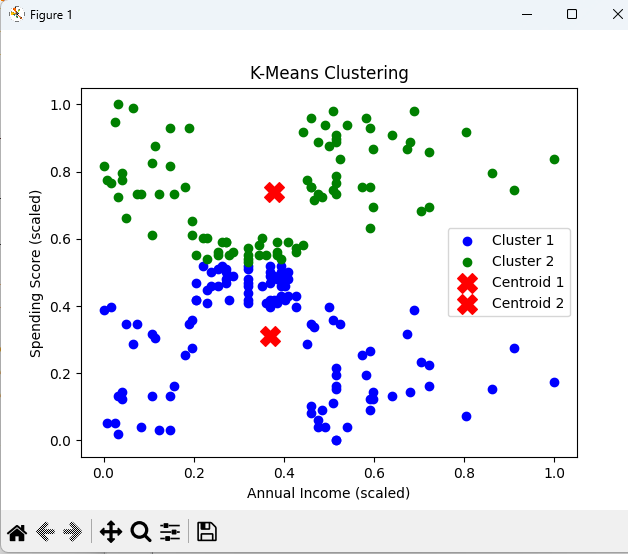
***plt.show()***

***print("Final Centroid 1:", centroid1)***

***print("Final Centroid 2:", centroid2)***

**Output :**

****

****