Experiment No.: 9

Aim

Program to implement k-means clustering technique using any standard dataset available in the public domain

CO3

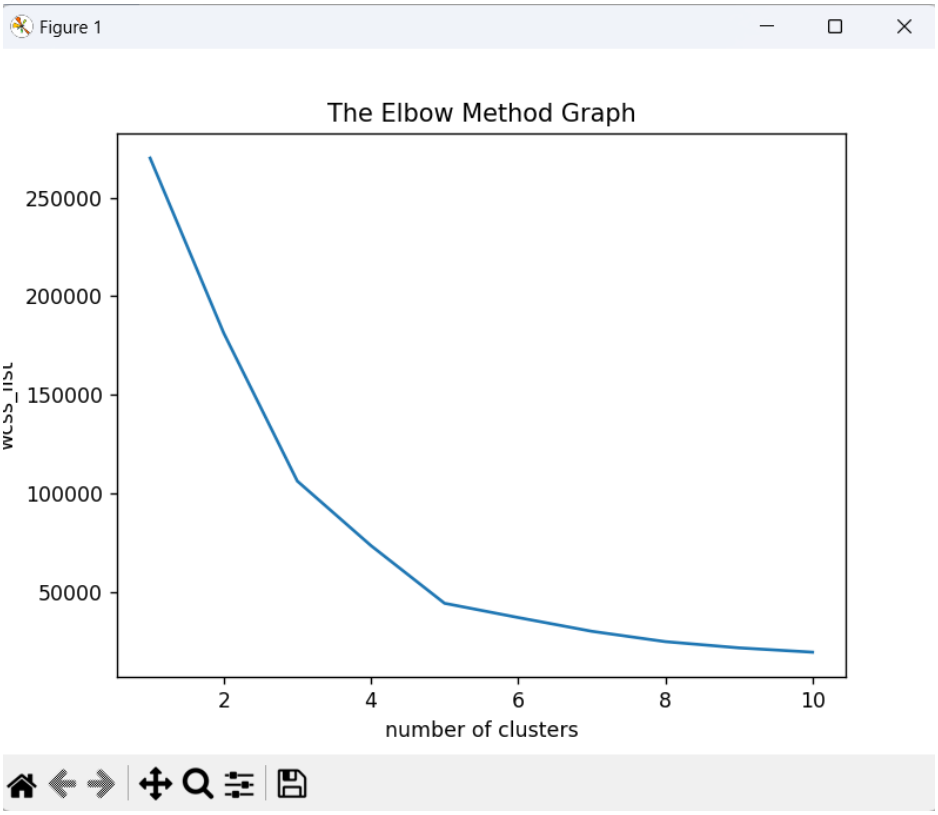
Use different packages and frameworks to implement text classification using SVM and clustering using k-means

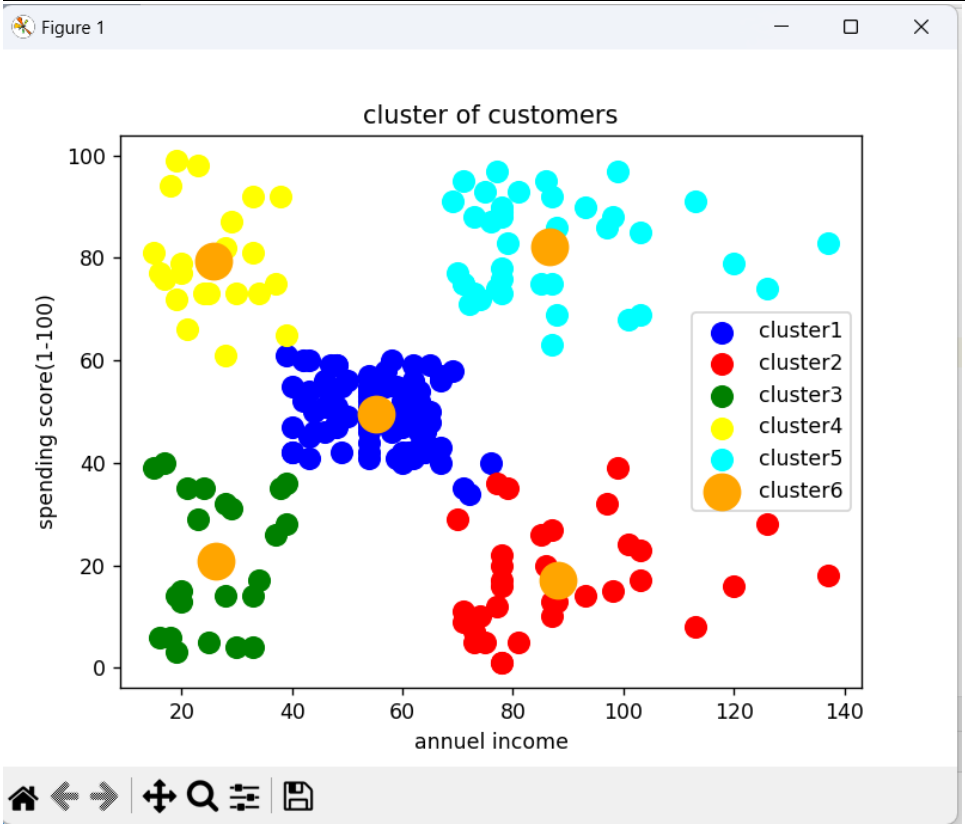
Procedure

```
import numpy as nm
import matplotlib.pyplot as mtp
import pandas as pd

dataset = pd.read_csv('Mall_Customers.csv')
x = dataset.iloc[:,[3,4]].values

from sklearn.cluster import KMeans
wcss_list = []
for i in range(1,11):
   kmeans = KMeans(n_clusters=i, init='k-means++',random_state=42)
   kmeans.fit(x)
   wcss_list.append(kmeans.inertia_)
mtp.plot(range(1,11),wcss_list)
mtp.title('The Elbow Method Graph')
mtp.xlabel('number of clusters')
mtp.ylabel('wcss_list')
mtp.show()

kmeans = KMeans(n_clusters=5,init='k-means++',random_state=42)
y_predict = kmeans.fit_predict(x)
print(y_predict)
```

```
mtp.scatter(x[y_predict == 0, 0], x[y_predict == 0, 1],s = 100,c = 'blue', label ='cluster1')
mtp.scatter(x[y_predict == 1, 0], x[y_predict == 1, 1],s = 100,c = 'red', label ='cluster2')
mtp.scatter(x[y_predict == 2, 0], x[y_predict == 2, 1],s = 100,c = 'green', label ='cluster3')
mtp.scatter(x[y_predict == 3, 0], x[y_predict == 3, 1],s = 100,c = 'yellow', label ='cluster4')
mtp.scatter(x[y_predict == 4, 0], x[y_predict == 4, 1],s = 100,c = 'cyan', label ='cluster5')

mtp.scatter(kmeans.cluster_centers_[:,0],kmeans.cluster_centers_[:, 1], s = 300, c = 'orange',
label = 'cluster6')
mtp.title('cluster of customers')
mtp.xlabel('annuel income')
mtp.ylabel('spending score(1-100)')
mtp.legend()
mtp.show()
```

Output Screenshot





Result

The program was executed and the result was successfully obtained. Thus CO3 was obtained.