Program No:12

Date:05-01-2022

Aim: Program to implement KMeans clustering technique using any standard data set available in the public domain

Program Code

import numpy as np

import matplotlib.pyplot as mtp

import pandas as pd

dataset= pd.read\_csv('world\_country\_and\_usa\_states\_latitude\_and\_longitude\_values.csv')

x = dataset.iloc[:,[1,2]].values

print(x)

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(k)')

mtp.ylabel('wcss\_list')

mtp.show()

kmeans = KMeans(n\_clusters=3, 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 = 'Cluster 1')

mtp.scatter(x[y\_predict == 1, 0], x[y\_predict == 1, 1], s = 100, c = 'green', label = 'Cluster 2')

mtp.scatter(x[y\_predict == 2, 0], x[y\_predict == 2, 1], s = 100, c = 'red', label = 'Cluster 3')

mtp.scatter(kmeans.cluster\_centers\_[:, 0], kmeans.cluster\_centers\_[:, 1], s = 300, c = 'yellow', label = 'Centroid')

mtp.title('clustors of customers')

mtp.xlabel('latitude')

mtp.ylabel('longitude')

mtp.legend()

mtp.show()

OUTPUT





