Experiment No 8 :Implementing K-Means Clustering Algorithm

AIM: To implement K-Means Clustering Algorithm using python

Code:

```python
import pandas as pd
import numpy as np
from sklearn.preprocessing import LabelEncoder
import csv
import seaborn as sns
import matplotlib.pyplot as plt

#reading and processing the data
data=pd.read_csv('abc.csv')
data=data.drop(['CustomerID'],axis=1)

user_list=data.iloc[:,0]
encoder=LabelEncoder()
data.iloc[:,0]= encoder.fit_transform(user_list)
data=data.iloc[:,:]

from mpl_toolkits.mplot3d import Axes3D

sns.set_style("white")
fig = plt.figure(figsize=(20,10))
ax = fig.add_subplot(111, projection='3d')
ax.scatter(data.Age, data["Annual_Income_(k$)"],
data["Spending_Score"], c='blue', s=60)
ax.view_init(30, 185)
plt.xlabel("Age")
plt.ylabel("Annual_Income_(k$)")
ax.set_zlabel('Spending_Score')
plt.show()


#Finding the apt value of k using elbow metod
from sklearn.cluster import KMeans

wcss = []
for k in range(1,11):
    kmeans = KMeans(n_clusters=k, init="k-means++")
```

```
        kmeans.fit(data.iloc[:,1:])
        wcss.append(kmeans.inertia_)
plt.figure(figsize=(15,10))
plt.grid()
plt.plot(range(1,11),wcss, linewidth=2, color="red", marker ="8")
plt.xlabel("K Value")
plt.xticks(np.arange(1,11,1))
plt.ylabel("WCSS")
plt.show()

#clustering
km = KMeans(n_clusters=5)
clusters = km.fit_predict(data.iloc[:,1:])
data["label"] = clusters
fig = plt.figure(figsize=(20,10))
ax = fig.add_subplot(111, projection='3d')
ax.scatter(data.Age[data.label == 0],
data["Annual_Income_(k$)"][data.label == 0],
data["Spending_Score"][data.label == 0], c='blue', s=60)
ax.scatter(data.Age[data.label == 1],
data["Annual_Income_(k$)"][data.label == 1],
data["Spending_Score"][data.label == 1], c='red', s=60)
ax.scatter(data.Age[data.label == 2],
data["Annual_Income_(k$)"][data.label == 2],
data["Spending_Score"][data.label == 2], c='green', s=60)
ax.scatter(data.Age[data.label == 3],
data["Annual_Income_(k$)"][data.label == 3],
data["Spending_Score"][data.label == 3], c='orange', s=60)
ax.scatter(data.Age[data.label == 4],
data["Annual_Income_(k$)"][data.label == 4],
data["Spending_Score"][data.label == 4], c='purple', s=60)
ax.view_init(30, 185)
plt.xlabel("Age")
plt.ylabel("Annual_Income_(k$)")
ax.set_zlabel('Spending_Score')
plt.show()
```

Result : You have successfully executed K-Means Clustering Algorithm using python .

Screenshot:

GRAPHICAL REPRESENTATION