NEGI AMIT SINGH
23552006

Unsupervised Machine Learning: Flat Clustering using K-Means Clustering Algorithm

Problem Statement: We have a set of unlabeled 2D points.

- The job is to label those 2D points in several clusters by using K-Meansclustering algorithm

Code:

```
# Importing libraries
import numpy as np
from sklearn.cluster import KMeans
import matplotlib.pyplot as plt
import pandas as pd

# Generate random 2D points for demonstration
data = np.random.rand(100,2)

# shape of dataset
data.shape
```

- There are 100 rows and 2 column

```
data

kmeans = KMeans(n_clusters=2,random_state=42)
kmeans.fit(data)

centroid = kmeans.cluster_centers_
```

```
labels = kmeans.labels_



labels


colors = ["g.", "r."]
for i in range(len(data)):
    plt.plot(data[i][0], data[i][1], colors[labels[i]], markersize=10)


plt.scatter(centroid[:, 0], centroid[:, 1], marker="x", linewidths=5, s=150, zorder=10)


# Convert NumPy array to Pandas DataFrame
df = pd.DataFrame(data, columns=['X', 'Y'])
df['Label'] = labels


# Export DataFrame to CSV with coordinates and labels
df.to_csv('coordinates_with_labels.csv', index=False)


plt.savefig('cluster_plot.png')
plt.show()
```

TESTING OF MODEL

```
new_data = np.random.rand(4,2)
print('input_data for testing model',new_data)
new_labels = kmeans.predict(new_data)
print("New Data Predictions:", new_labels)
```

```
In [9]: new_data = np.random.rand(4,2)
        print('input_data for testing model',new_data)
        new_labels = kmeans.predict(new_data)
        print("New Data Predictions:", new_labels)

        input_data for testing model [[0.86320386 0.44022361]
         [0.3076229  0.7753791 ]
         [0.56598163 0.66831945]
         [0.74356132 0.21351112]]
        New Data Predictions: [0 0 0 1]
```