

GLS University
Faculty of Computer Applications & Information Technology
Machine Learning
Practical Notes
Unit4

K means of Clustering

Demo Code:

```
import numpy as np
import matplotlib.pyplot as plt
from sklearn.cluster import KMeans

# Given data: Height and Weight
data = np.array([
    [185, 72],
    [170, 56],
    [168, 60],
    [179, 68],
    [182, 72],
    [188, 77],
    [180, 71],
    [180, 70]
])

# Applying K-Means Clustering with K=2
kmeans = KMeans(n_clusters=2, random_state=42, n_init=10)
kmeans.fit(data)
labels = kmeans.labels_
centroids = kmeans.cluster_centers_

# Output the cluster labels
print("Cluster Labels:", labels)

# Plot the clusters
plt.figure(figsize=(8,6))
plt.scatter(data[:, 0], data[:, 1], c=labels, marker='o', label="Data Points")
plt.scatter(centroids[:, 0], centroids[:, 1], c='red', marker='X', s=200, label="Centroids")
plt.xlabel("Height")
plt.ylabel("Weight")
plt.title("K-Means Clustering (K=2)")
plt.legend()
plt.grid(True)
plt.show()
```

Demo Code CSV File:

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

# Load data from CSV file
data = pd.read_csv("data.csv") # Ensure the CSV file has 'Height' and 'Weight' columns

# Extract relevant features
X = data[['Height', 'Weight']].values

# Applying K-Means Clustering with K=2
kmeans = KMeans(n_clusters=2, random_state=42, n_init=10)
kmeans.fit(X)
labels = kmeans.labels_
centroids = kmeans.cluster_centers_

# Output the cluster labels
print("Cluster Labels:", labels)

# Plot the clusters
plt.figure(figsize=(8,6))
plt.scatter(X[:, 0], X[:, 1], c=labels, marker='o', label="Data Points")
plt.scatter(centroids[:, 0], centroids[:, 1], c='red', marker='X', s=200, label="Centroids")
plt.xlabel("Height")
plt.ylabel("Weight")
plt.title("K-Means Clustering (K=2)")
plt.legend()
plt.grid(True)
plt.show()
```