Name : Devkumar Biswas

Class: BE(AI&DS)

Div: B

Subject : ML(CL-I Lab)

Roll no. : BEAD21267

**Assignment No. - 1**

**Problem Statement** : Feature Transformation (Any one)

1. To use PCA Algorithm for dimensionality reduction. You have a dataset that includes measurements for different variables on wine (alcohol, ash, magnesium, and so on). Apply PCA algorithm & transform this data so that most variations in the measurements of the variables are captured by a small number of principal components so that it is easier to distinguish between red and white wine by inspecting these principal components. Dataset Link: <https://media.geeksforgeeks.org/wp-content/uploads/Wine.csv>

Code :

import pandas as pd

from sklearn.decomposition import PCA

from sklearn.preprocessing import StandardScaler

import matplotlib.pyplot as plt

df = pd.read\_csv('Wine.csv')

df

df.keys()

print(df.isnull().sum())

X = df.drop('Customer\_Segment', axis=1) # Features

y = df['Customer\_Segment']

sc = StandardScaler() #Standardize features by removing the mean and scaling to

pca = PCA()

X\_pca = pca.fit\_transform(X)

explained\_variance\_ratio = pca.explained\_variance\_ratio\_

plt.plot(range(1, len(explained\_variance\_ratio) + 1), explained\_variance\_ratio.cumsum(), marker='o',

linestyle='--')

plt.xlabel('Number of Principal Components')

plt.ylabel('Cumulative Explained Variance')

plt.title('Explained Variance Ratio')

plt.show()

n\_components = 12 # Choose the desired number of principal components you want to reduce a dimention to

pca = PCA(n\_components=n\_components)

X\_pca = pca.fit\_transform(X)

X\_pca.shape

X.shape

red\_indices = y[y == 1].index

white\_indices = y[y == 2].index

plt.scatter(X\_pca[red\_indices, 0], X\_pca[red\_indices, 1], c='red', label='Red Wine')

plt.scatter(X\_pca[white\_indices, 0], X\_pca[white\_indices, 1], c='blue', label='White Wine')

plt.xlabel('Principal Component 1')

plt.ylabel('Principal Component 2')

plt.legend()

plt.title('PCA: Red Wine vs. White Wine')

plt.show()

Output :







