# **PRINCIPAL COMPONENT ANALYSIS**

### **Aim**

To implement Principal Component analysis

### **Problem description**

This project involves implementing Principal Component analysis on a wine dataset for dimensionality reduction and visualization.

**Algorithm**

#### **1) Import the required libraries**

Import the python libraries numpy,pandas,matplotlib,sklearn

**2) Load the dataset**

Load the dataset **“**winedata.csv**”**

#### **3) Preprocessing the dataset**

#### **4)Standardize the features using** StandardScaler **to ensure that all features have the same scale**

#### **5) Apply PCA to reduce the dimensionality of the dataset**

#### **6) Visualize the data in the reduced-dimensional space**

### **Program code/ Pseudocode**

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

%matplotlib inline

from sklearn.preprocessing import StandardScaler

from sklearn.decomposition import PCA

df = pd.read\_csv(r'C:\ardhra\PCA\wine.data.csv')

df.head(10)

X = df.drop('Class',axis=1)

y = df['Class']

scaler = StandardScaler()

X = scaler.fit\_transform(X)

dfx = pd.DataFrame(data=X,columns=df.columns[1:])

pca = PCA(n\_components=None)

dfx\_pca = pca.fit(dfx)

dfx\_trans = pca.transform(dfx)

dfx\_trans = pd.DataFrame(data=dfx\_trans)

plt.figure(figsize=(10,6))

plt.scatter(dfx\_trans[0],dfx\_trans[1],c=df['Class'],edgecolors='k',alpha=0.75,s=150)

plt.grid(True)

plt.title("Class separation using first two principal components\n",fontsize=20)

plt.xlabel("Principal component-1",fontsize=15)

plt.ylabel("Principal component-2",fontsize=15)

plt.show()

### 

### 

### **Result**

