**Ex. No.: 10**

**Date: 10/11/24**

**A PYTHON PROGRAM TO IMPLEMENT DIMENSIONALITY REDUCTION USING PCA**

**Aim:**

To implement Dimensionality Reduction using PCA in a python program.

**Algorithm:**

Step 1: Import Libraries

Import necessary libraries, including pandas, numpy, matplotlib.pyplot, and sklearn.decomposition.PCA.

Step 2: Load the Dataset (iris dataset)

Load your dataset into a pandas DataFrame.

Step 3: Standardize the Data

Standardize the features of the dataset using StandardScaler from sklearn.preprocessing.

Step 4: Apply PCA

* Create an instance of PCA with the desired number of components.
* Fit PCA to the standardized data.
* Transform the data to its principal components using transform.

Step 5: Explained Variance Ratio

* Calculate the explained variance ratio for each principal component.
* Plot a scree plot to visualize the explained variance ratio.

Step 6: Choose the Number of Components

Based on the scree plot, choose the number of principal components that explain a significant amount of variance.

Step 7: Apply PCA with Chosen Components

Apply PCA again with the chosen number of components.

Step 8: Visualize the Reduced Data

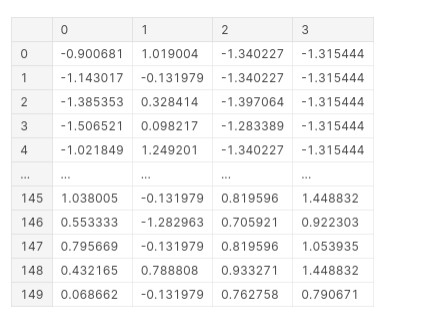
* Transform the original data to the reduced dimension using the fitted PCA.
* Visualize the reduced data using a scatter plot.

Step 9: Interpretation

Interpret the results, considering the trade-offs between dimensionality reduction and information loss.

**PROGRAM:**

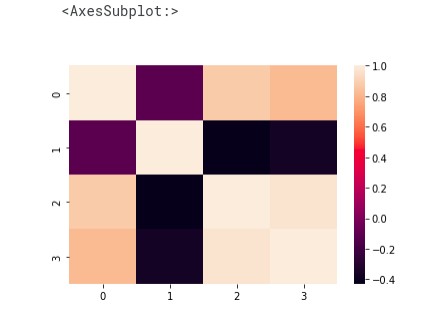
from sklearn import datasets import pandas as pd from sklearn.preprocessing import StandardScaler from sklearn.decomposition import PCA import seaborn as sns iris = datasets.load\_iris() df = pd.DataFrame(iris['data'], columns = iris['feature\_names']) df.head()



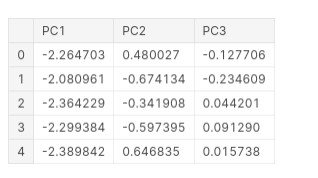


scalar = StandardScaler()

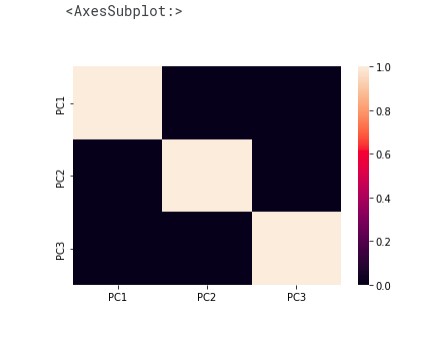
scaled\_data = pd.DataFrame(scalar.fit \_transform(df)) #scaling the data scaled\_datasns.heatmap(scaled\_data.corr())



pca = PCA(n\_components = 3) pca.fit(scaled\_data) data\_pca = pca.transform(scaled\_data) data\_pca = pd.DataFrame(data\_pca,columns=['PC1','PC2','PC3']) data\_pca.head()



sns.heatmap(data\_pca.corr())



**RESULT:-**

Thus Dimensionality Reduction has been implemented using PCA in a python program successfully and the results have been analyzed.