

Experiment 9

Student Name : Dhanuj Tomar

Branch: CSE

Semester: 5th

Subject Name: Machine learning lab

UID:20BCS1201

Section/Group:701B

Subject Code:20CSP-317

1. Aim:

Implement Principal Component Analysis.

2. Result and output:

➔ Importing PCA from sklearn and provide a dataset

```
import pandas as pd
from sklearn import svm
import numpy as np
from sklearn.decomposition import PCA

In [13]: #Get the IRIS dataset
url = "https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data"
data = pd.read_csv(url, names=['sepal length', 'sepal width', 'petal length', 'petal width', 'target'])
```

➔ Prepare the data with target point and PCA function

```
In [13]: #Get the IRIS dataset
url = "https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data"
data = pd.read_csv(url, names=['sepal length', 'sepal width', 'petal length', 'petal width', 'target'])

In [14]: #prepare the data
x = data.iloc[:,0:4]

In [15]: #prepare the target
target = data.iloc[:,4]

In [19]: pca = PCA(2)
mat_reduced = pca.fit_transform(x)
```

➔ Concat it with target variable to create a complete Dataset.

```
In [19]: pca = PCA(2)
mat_reduced = pca.fit_transform(x)

In [23]: #Creating a Pandas DataFrame of reduced Dataset
principal_df = pd.DataFrame(mat_reduced , columns = ['PC1','PC2'])

In [24]: #Concat it with target variable to create a complete Dataset
principal_df = pd.concat([principal_df , pd.DataFrame(target)] , axis = 1)

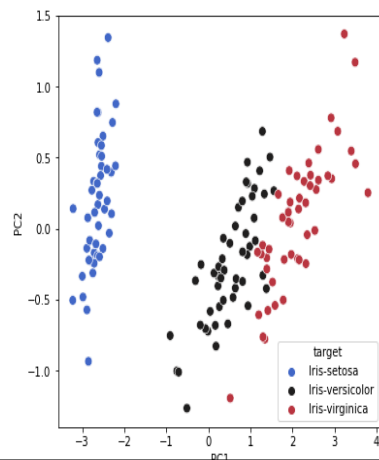
In [25]: import seaborn as sb
import matplotlib.pyplot as plt
```

➔Printing scatter plot for given PCA labels

```
In [25]: import seaborn as sb
import matplotlib.pyplot as plt

In [26]: plt.figure(figsize = (6,6))
sb.scatterplot(data = principal_df , x = 'PC1',y = 'PC2' , hue = 'target' , s = 60 , palette= 'icefire')

Out[26]: <AxesSubplot:xlabel='PC1', ylabel='PC2'>
```



- 1- Import PCA from sklearn.
- 2- Provide it with iris dataset.
- 3- Prepare the target and generate PCA() function.
- 4- Concat it with target variable to create a complete Dataset.
- 5- The Scatterplot for given columns PC1 And PC2 is generated.