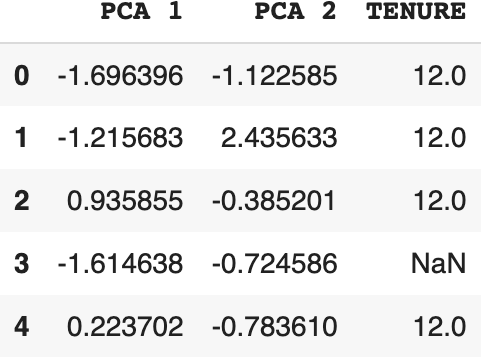
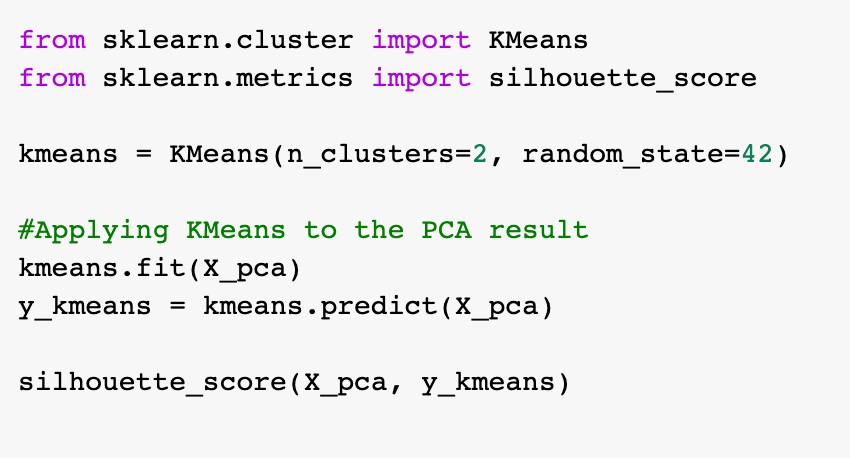


## Output:



**Explana,on:**

Firstly dropped the rows with null values, then the features are standardized using StandardScaler() function. PCA has applied to the scaled data, then the finalData is printed by concatenating the principal data and tenure.



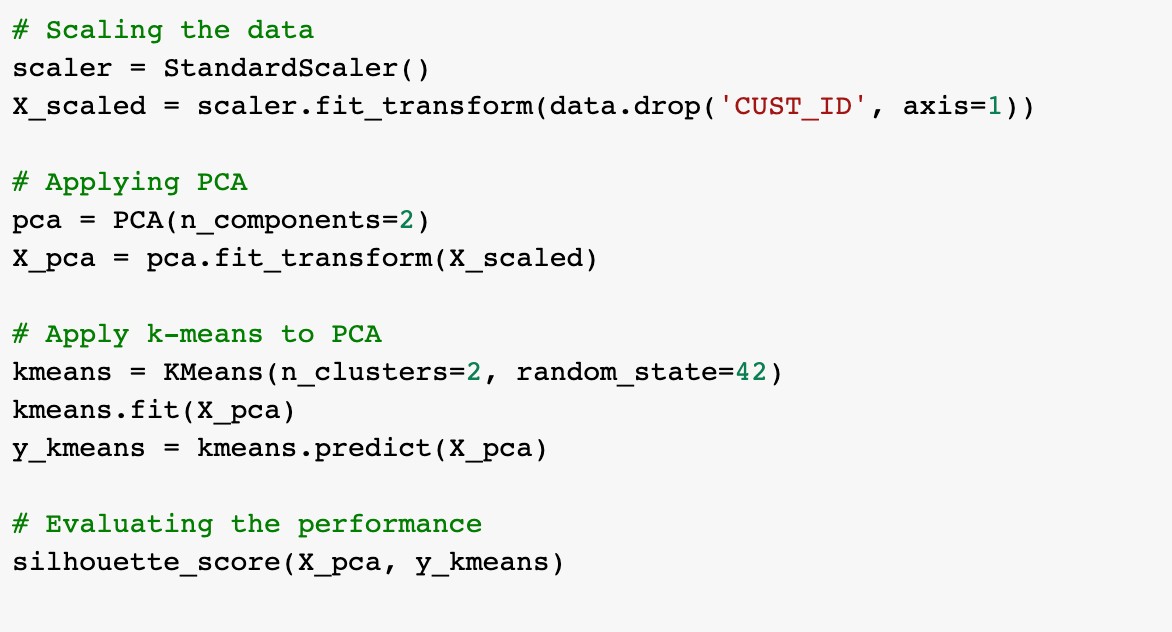
## Output:



**Explana,on:**

Applied k-means() to the PCA result using the silhoue8e score to evaluate the performance.

# 1(c):

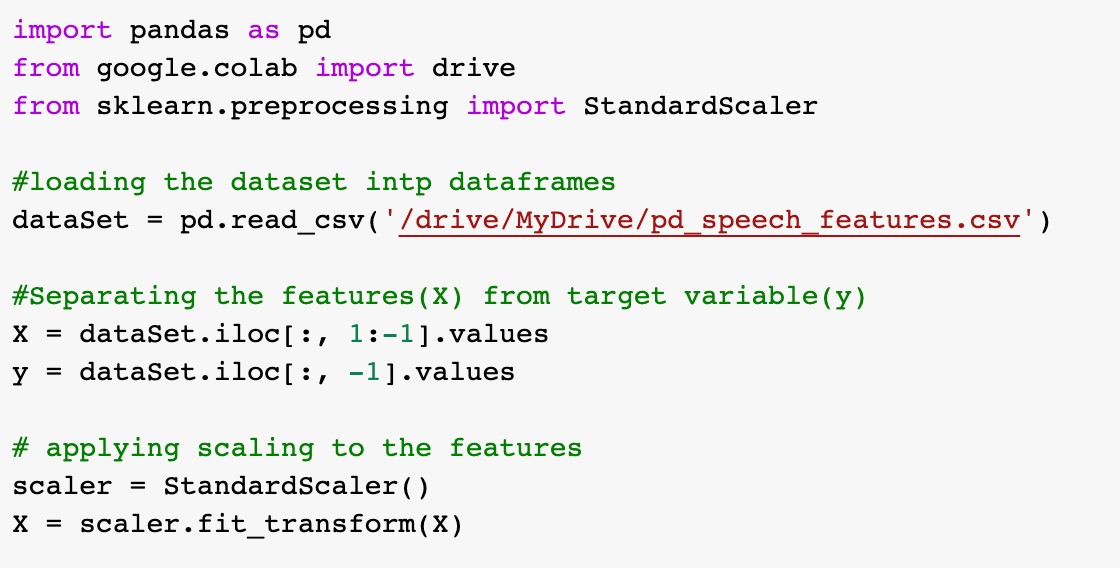




## Explana,on:

The data has been scaled by using the StandardScaler() funcAon, then aCer applied the PCA for the above result. And applied k-means() to PCA result. Finally evaluated the performance by using the silhoue8e score.

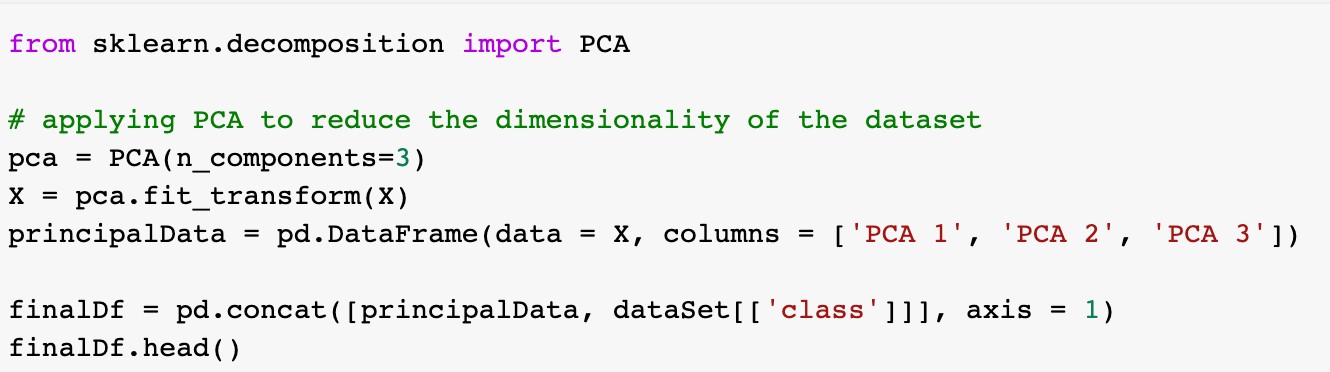
# Solu%on 2(a):



## Explana,on:

ACer loading the dataset separated the features(X) from target variable(y) then applied the scaling to the features using StandardScaler() funcAon.

# 2(b):

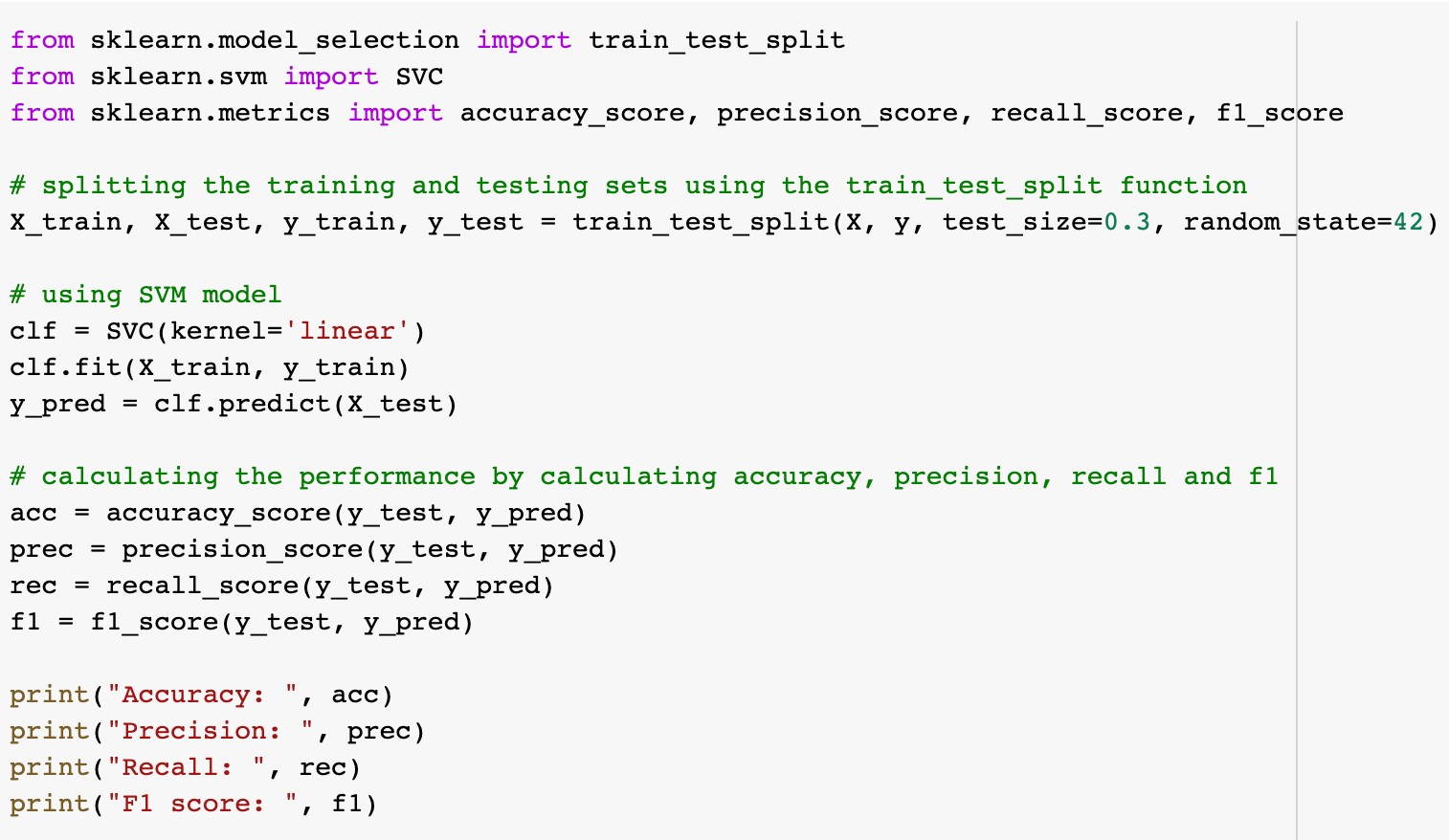


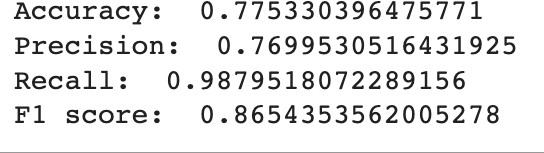


## Explana,on:

Now applying the PCA to reduce the dimensionality of the dataset. Then the final data is printed by concatenating all the principal data with class.

# 2(c):

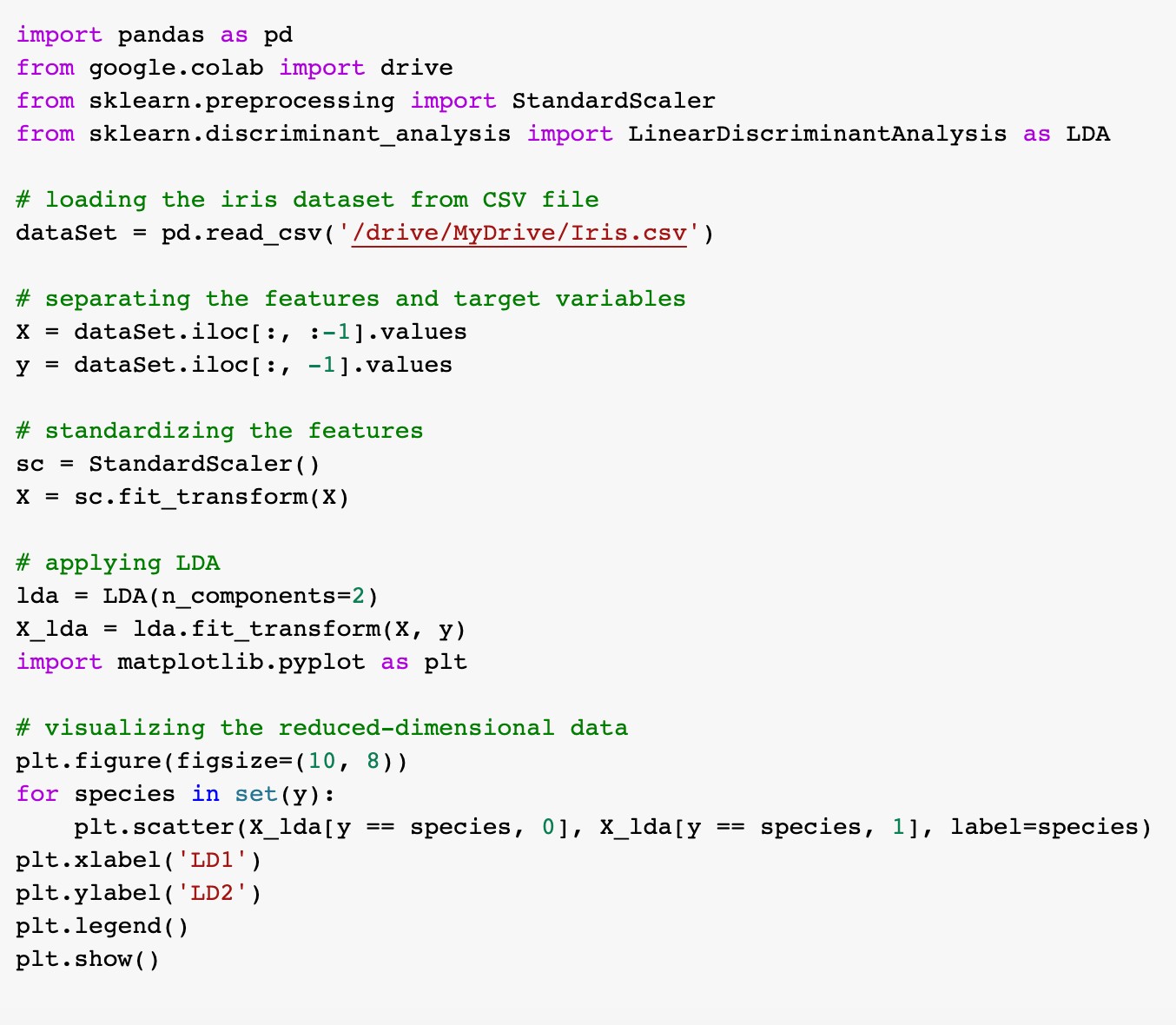


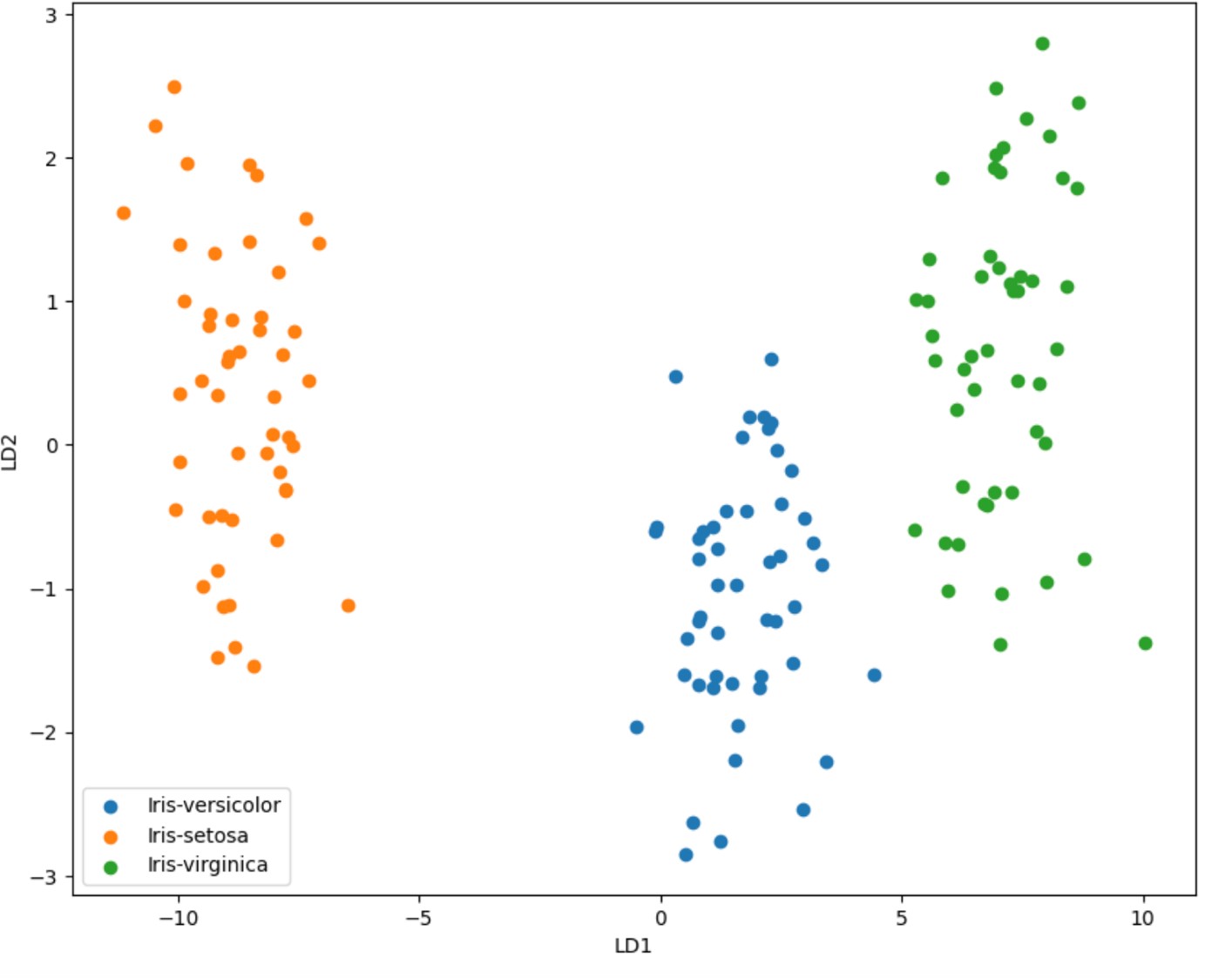


## Explana,on:

Used SVM model to calculate the performance of the model. Firstly split the dataset into training and tesAng sets using the train\_test\_split funcAon. Then used the trained model to predict the target variable for tesAng set and ﬁnally evaluated the performance of the model using Accuracy, Precision, Recall and F1 score.

# Solu%on 3:





## Explana,on:

Firstly, loaded the Iris dataset from the drive. Then separated the features(X) from target variable(y) then aCer applied the scaling to the features using StandardScaler() funcAon. Applied LDA to report the features onto a lower-dimensional space. Then visualized the data.

# Solu%on 4:

The aim of LDA and PCA is to maximize the variance in a lower dimension using linear transformaAons.

PCA is an unsupervised learning method that looks for a new set of orthogonal variables known as Principal components, that can best accommodate for the data's variaAon. The ﬁrst principal component accounts for the majority of the variance in the data, and the principal components are arranged in order of their capacity to explain variance. PCA is frequently used for exploratory research and data visualizaAon.

The LDA method is of supervised learning which looks for a linear combinaAon of features that best disAnguishes between the classes in the data. Maximizing the variance between classes while minimizing the variance within classes is the aim of LDA. LDA is frequently applied to classiﬁcaAon issues, where the goal is to predict, from a new observaAon's features, what class it belongs to.

PCA is an unsupervised technique for reducing the dimensionality of the data by idenAfying the principal components that explain the most variance, whereas LDA is a supervised technique for reducing the dimensionality of the data by idenAfying the linear combinaAon of features that best separates the classes in the data.

# GitHub Link:

# https://github.com/Sowmyaala/Assignment-5