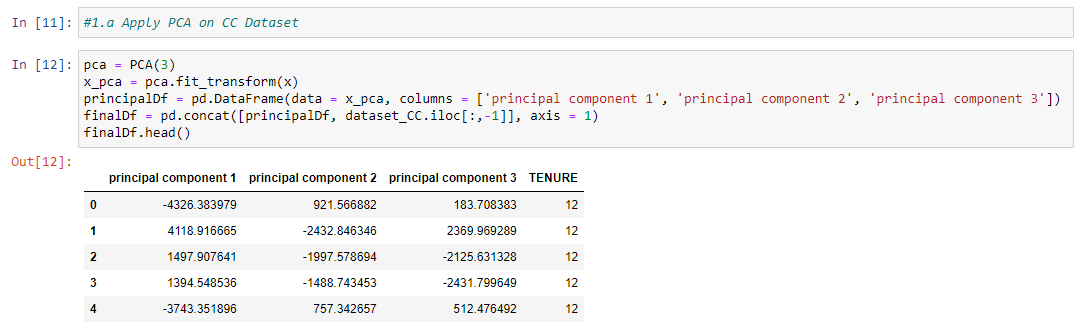
ASSIGNMENT 5

1. Principal Component Analysis

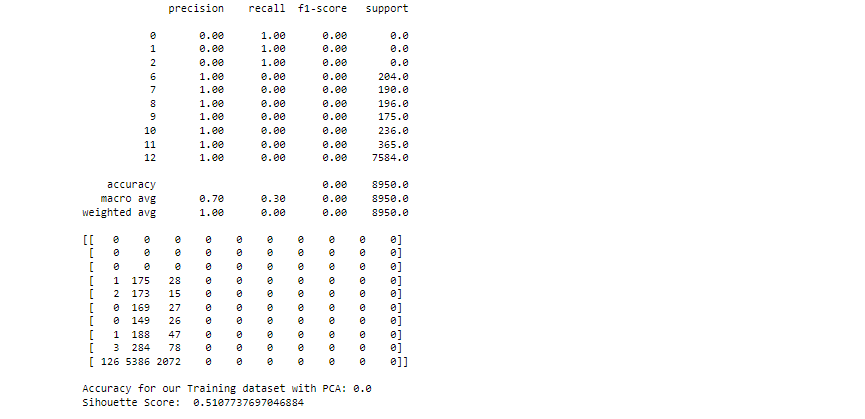
a. Apply PCA on CC dataset.

Use PCA to reduce the dimensionality of our data. Select an appropriate number of components and analyze total variance explained. Interpret to make sense of the principal components.

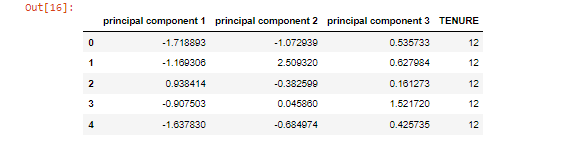


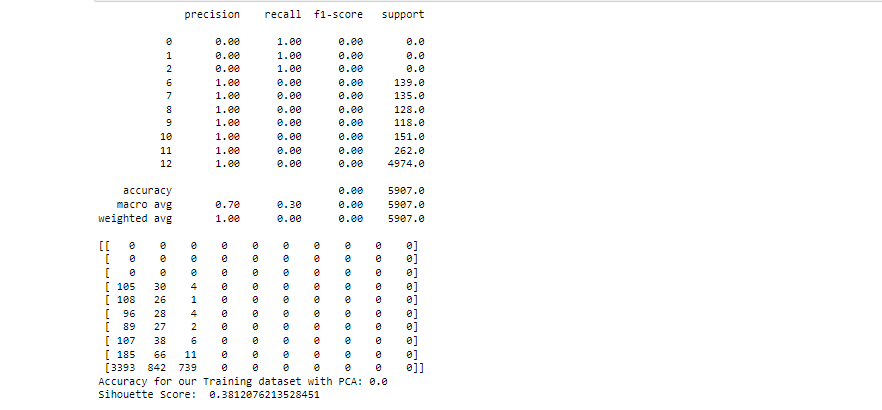
b. Apply k-means algorithm on the PCA result and report your observation if the silhouette score has improved or not?

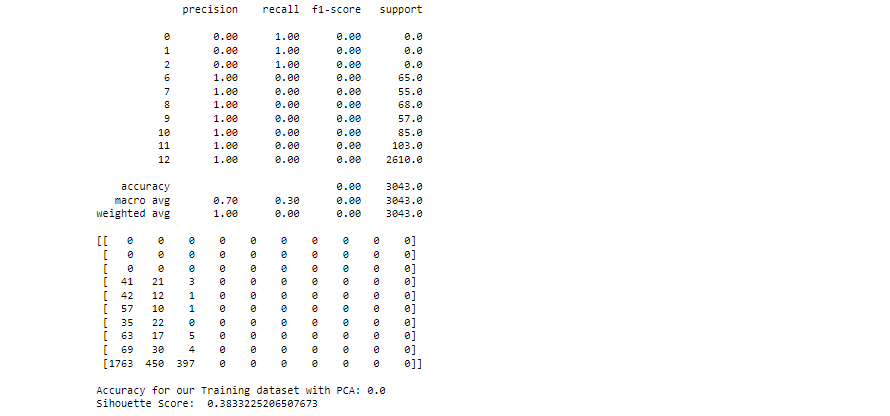
The silhouette value is a measure of how similar an object is to its own cluster (cohesion) compared to other clusters (separation). The silhouette ranges from −1 to +1, where a high value indicates that the object is well matched to its own cluster and poorly matched to neighbouring clusters.



c. Perform Scaling+PCA+K-Means and report performance.







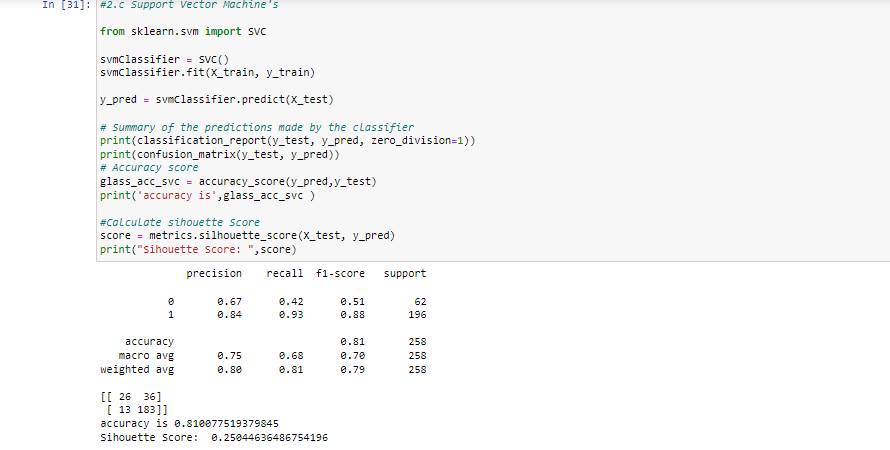
2. Use pd\_speech\_features.csv

a. Perform Scaling

b. Apply PCA (k=3)



c. Use SVM to report performance



3.Apply Linear Discriminant Analysis (LDA) on Iris.csv dataset to reduce dimensionality of data to k=2.

'LDA finds the linear discriminants in order to maximize the variance between the different categories while minimizing the variance within the class.'

