Machine learning assignment 5

# Sai Kiran Reddy Kotha

Banner ID: 700746206

Githublink: <https://github.com/reddysahib/assignment_ml>

VideoLink: https://drive.google.com/file/d/1On6PVFHNafAvif4Q74HsvuEVQ4CqClP\_/view?usp=sharing

Graphical user interface, text, application

Description automatically generated

Table

Description automatically generated

Table

Description automatically generated

filling null values with mean for pca

using kmeans on data before pca

Graphical user interface, text, application, email

Description automatically generated

Table

Description automatically generated

Table

Description automatically generated

Text, table

Description automatically generated with medium confidence

Table

Description automatically generated

Table

Description automatically generated

Chart, scatter chart

Description automatically generated

Briefly identify the difference between PCA and LDA

Pca is unsupervised and lda is supervised

If you’re working with labeled data and your goal is to find a low-dimensional representation then use lda otherwise pca

if we want to retain as much data as possible but want to decrease the number of columns then pca is better

This means that PCA finds directions of maximum variance regardless of class labels while LDA finds directions of maximum class separability.

PCA performs better in case where number of samples per class is less. Whereas LDA works better with large dataset

having multiple classes; class separability is an important factor while reducing dimensionality