# 

**CSE523 : Machine Learning**

Winter 2021 - 2022

**Weekly Report - 11**

Dt : 20-04-2022

**Group Name : Discover Decipher**

**Group Members**

Nimisha Patel - AU1940146 - Btech Computer Science and Engineering.

Sakshi Shah - AU1940213 - Btech Computer Science and Engineering.

Astha Patel - AU1940312 - Btech Computer Science and Engineering.

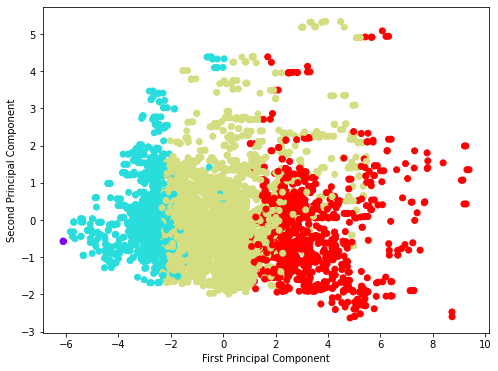
Kareena Matwani - AU1940314 - Btech Computer Science and Engineering.

**Task performed this week.**

* **Dimensionality reduction**

We performed LDA, and the model accuracy for random forest classifiers decreased which suggested that the decreased dimensions from 11 dimensions to 2 dimensions are not optimal for modeling. Thus, we decided to tune an important hyper parameter of the Linear Discriminant Analysis method for our dataset that is ‘solver’. After evaluation configurations by using repeated cross-validation, svd solver performed better compared to eigen and lsqr. The shrinkage value consideration as 0.1 also shows mean accuracy 0.90 which is lesser than the previous dataset without dimensionality reduction.

Tried PCA for dimensionality reduction, the dimensions of the dataset reduced from 11 features to 2 dimensional columns. Modeled Randomforest classifier on the reduced dataset



**PCA components**

Accuracy of Random Forest classifier on training set: 0.79.

Accuracy of Random classifier on test set: 0.77.

The reduced dimensions by PCA are not optimal for classification through random forest.

* **Worked on the final analysis and conclusions.**

**Things to be done**

* **Documentation : Preparing final presentation, code files and Final Report.**