Name: Tran Thi Thuy Kieu

Student ID: USTHBI7-088

Machine Learning

REPORT

K-MEANS CLUSTERING WITH EQUAL CLUSTERS

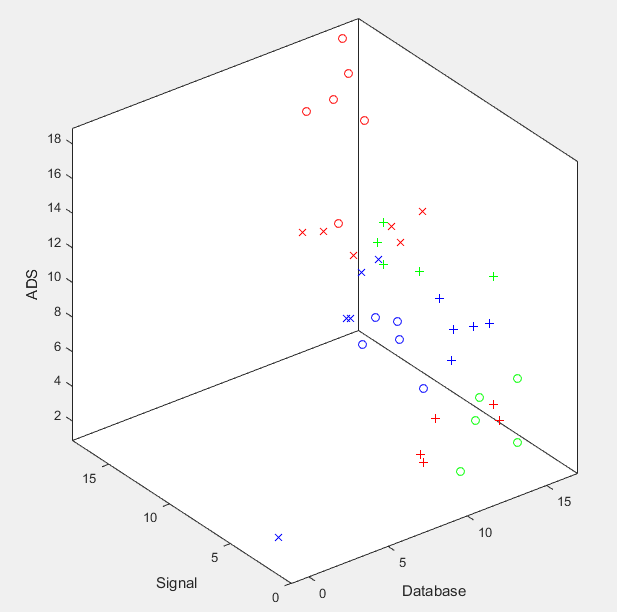
I- Introduction

* K-means (unsupervised learning/clustering algorithm) implemented in MATLAB
* “kmeans.m” is a script that generates uniformly distributed data points into equal clusters.
* “hungarian.m” is a script that contains Hungarian algorithm solving k-means phase.

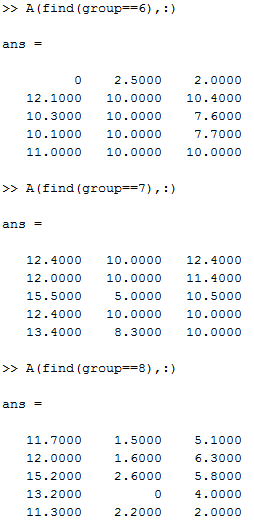
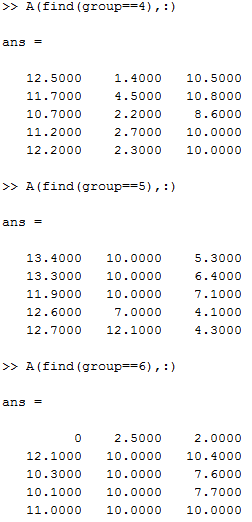
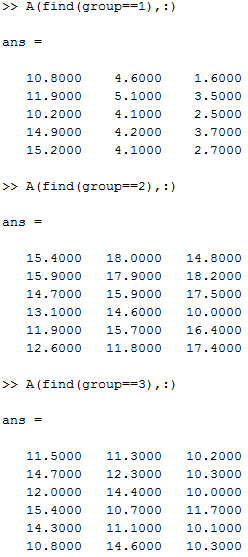
II- Algorithm details

* The program runs the entire process multiple times with different initializations of centroids
* Then, it uses the run which has the best mean squared error between points and their assigned centroids.
* For the k-means algorithm phase, applied Hungarian algorithm.

III- Result



*3D-Visualiztion of clusters*



*Points in each group*

IV- Assessment

* Quality: Nearly good.
* Need to improve code beverage.
* Time complexity: O(n^3).

V- Reference

* [Mikko Malinen : Balanced K-means for Clustering](https://link.springer.com/content/pdf/10.1007%2F978-3-662-44415-3_4.pdf)