

3CS452 : Data Warehousing and Data Mining Lab

Assignment 5 : Design and implementation of following clustering algorithms

1. Hierarchical clustering - AGNES & DIANA. Plot Dendrogram.
2. *k*-Means
3. *k*-Medoids (PAM)
4. DBSCAN

Note :

- Use IRIS, Breast Cancer Wisconsin data sets from UCI Machine Learning Repository (<http://archive.ics.uci.edu/ml/>).
- For DBSCAN, use *US Census Data (1990) Data Set* [https://archive.ics.uci.edu/ml/datasets/US+Census+Data+\(1990\)](https://archive.ics.uci.edu/ml/datasets/US+Census+Data+(1990))
- Extract different training and test data sets from original data set.
- Tabulate the results with cluster validation accuracy.

Dr. B. F. Momin
Course Teacher