Assignment 2.

Date of Completion
27.8.2020

Title :- Visualize the clusters using suitable tool

Problem Statement :- Consider a suitable dataset. For clustering of data instances in different groups, apply different clustering techniques. Visualize the clusters using suitable tool

Learning Objective :- Understand clustering and different algorithms used for clustering data

Learning Outcomes :- Students will be able to understand different clustering methods and implement them.

Theor. Software / Hardware requirements :- Python, numpy, sklearn (packages)

Theory :-

Clustering Algorithms :-

1. Clustering is a machine Learning technique that involves the grouping of points.

2. It is used to analyse patterns and grouping.
3. It is a unsupervised learning algorithm.

Types of clustering algorithms:-
1) K Means clustering
2) Hierarchical clustering
3) Mean shift clustering
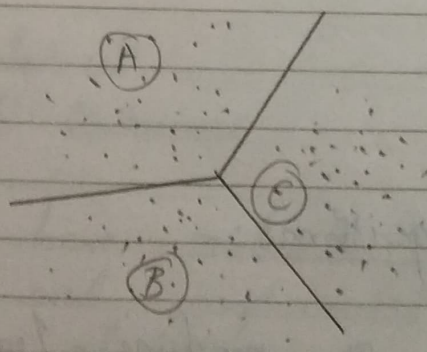4) Fuzzy C clustering
5) Spectral clustering etc.

Clustering Methods:-
1. Density Based
2. Hierarchical Based
3. Partitioning Based
4. Grid-based

K-Means algorithm:-
1. Simplest unsupervised clustering algorithm.
2. It partions n observation into k the clusters where each observation belongs to clusters nearest mean serving as a prototype cluster.
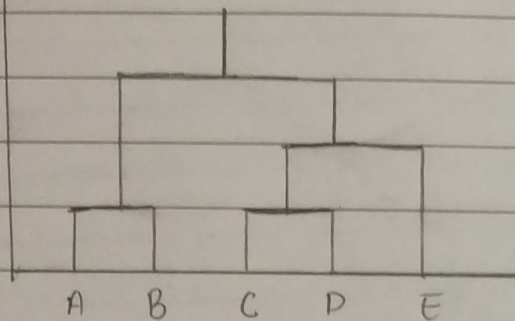
Applications:-
Marketing, Biology, Earthquake studies etc.

Hierarchical clustering:-

1. It builds a hierarchy of clusters.
2. 2 types :-
   a) Agglometric :- "Bottom up" approach
   b) Divisive :- "Top down" approach.
3. Results are usually presented in a dendogram.
4. Slow.

Linkage Criteria :-

To compute the distance between two similar clusters many linkage criteria have developed.

Dataset Used :-

K-means :- iris dataset
Hierarchical :- Mall Customers.

Conclusion :- Thus I have understood different clustering algorithms and implemented K-means & hierarchical clustering algorithm.