Assignment A2 41403

- Title: Visualize the clusters using suitable tool.
- problem statement: consider a suitable dataset for clustering of data instances in different groups apply different clustering technique visualize the clusters using suitable tool
- learning objective: Understand clustering & different algorithm used for clustering data
- learning outcome: student will be able to understand different clustering methods & implement them.
- S/w H/w req: python, numpy, sklearn.
- Theory:

clustering Algorithm:

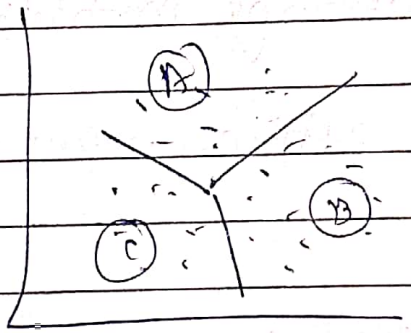
- clustering is a machine learning technique that involves the grouping of points
- It is used to analyze patterns & grouping
- It is a unsupervised learning algorithm.

- Types of clustering algorithm
  - K means clustering
  - Hierarchical clustering
  - mean shift clustering
  - fuzzy c clustering
  - spectral clustering

41403

clustering methods
- Density based
- Hierarchical based
- partioning based
- Grid based

a. k-means clustering:
- Simplest unsupervised clustering algorithm
- It partitions n observations into k clusters where each observation belongs to clusters nearest mean serving as a prototype cluster

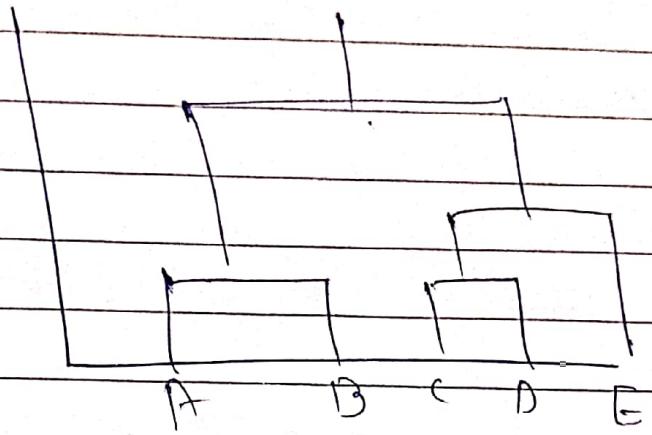


Application: marketing, biology, etc.

b Hierarchical clustering:
- It builds a hierarchy of clusters
- These are of 2 types:
  - Agglometric: bottom up approach
  - divisic: top down approch
- results are usually present in a dendrogram
- slow

41403

- linkage criteria: To compute the distance b/w two similar clusters many linkage criteria have developed



- Dataset used: K-means: iris dataset
  Hierarchical: mall customers

- conclusion: Thus I have understood different clustering algorithm & implemented K-means & hierarchical clustering algorithm