Module 4

Q1) Differentiate between agglomerative & divisive clustering methods.

=>

Agglomerative clustering	Divisive clustering
(1) Bottom up approach	(1) Top - down approach
(2) Combines the closest cluster at each step	(2) Splits the most dissimilar cluster at each step
(3) Each data point is a single cluster	(3) All data points are in a single cluster
(4) More commonly used	(4) Used less often
(5) Low computational cost	(5) High computational cost
(6) Easy to implement	(6) Hard to implement
(7) Sensitive to noisy data and outliers	(7) Sensitive to initial split criteria
(8) Stops when all data points are in one cluster	(8) stops when each data point is in its own cluster.
(9) Builds the hierarchy from bottom to top (merging)	(9) Builds the hierarchy from top to bottom (splitting)

Page No.
Date

Q27 K-means clustering

=> (1) K-means clustering is an unsupervised machine learning algorithm used for clustering tasks.

(2) In K-means, unlabeled data is grouped into K-distinct clusters based on similarity.

(3) Each cluster is defined by its centroid, and the goal is to minimize the distance between data points and their respective cluster centroids.

(4) Steps in K-means

(i) Step 1: Initialize K
Choose the number of clusters (K)

(ii) Step 2: Select initial centroids
Randomly initialize K points as centroid

(iii) Step 3: Assign data points to nearest centroid
For each data point, compute its distance from all K centroids and assign it to the cluster with nearest centroid.

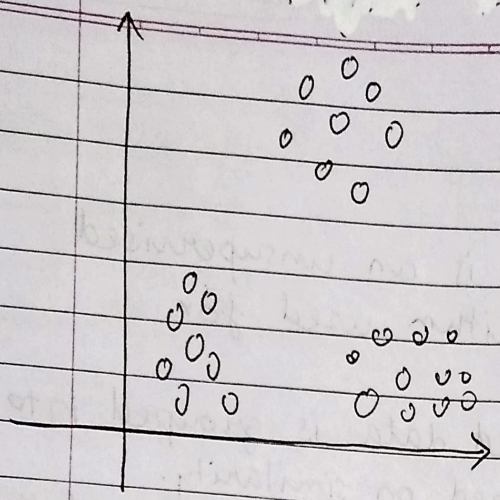
(iv) Step 4: Recalculate centroids
Compute the mean of all data points in each cluster and update the centroid positions

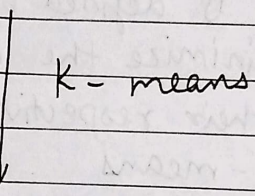
(v) Step 5: Repeat
Reassign all data points to the nearest centroid and recalculate the centroids

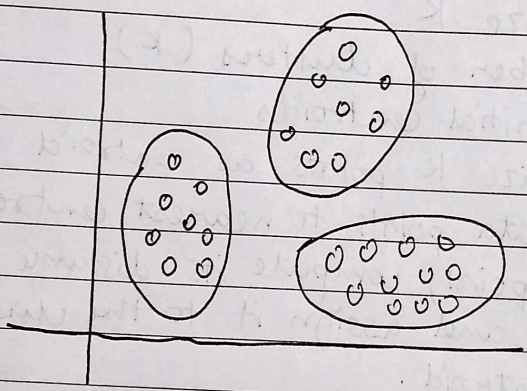
(vi) Step 6: Convergence check
Stop if:
(1) No data points change clusters
(2) The centroids do not move significantly.
(3) The max no of iterations is reached

(vii) Step 7: Output
The algorithm outputs K clusters and their corresponding centroids.

Page No.
Date







(5) Flowchart