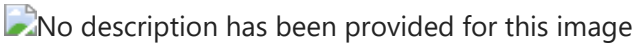
55. Hierarchical Clustering

It is applied for linearly separable data

- It is used to group the unlabelled datasets into a cluster and aslo known as hierarchical cluster analysis or HCA.
- In the algorithm, we develop the hierarchy of clusters in the form of a tree, and this tree-shaped structure is known as the **dendrogam**.

Dendrogram

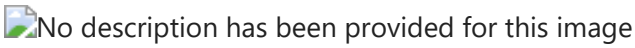
- It is a tree like structure that is mainly used to store each step as a memory that the HC algorithm performs.
- The dendrogram plot, the Y-axis shows the **Euclidean distances** b/w the data points, and the x-axis shows all the data points of the given dataset.



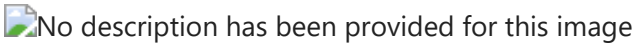
Hierarchical clustering technique has two approaches:

1. **Agglomerate:** Agglomerative is a bottom-up approach, in which the algorithm starts with taking all data points as single clusters and merging them until one cluster is left. This is popular algorithem and **bottom-up approach**.
2. **Divisive:** Divisive algorithm is the reverse of the agglomerative algorithm as it is **top-down approach**.

Agglomerate Clustering:



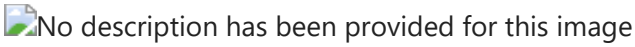
Divisive Clustering:

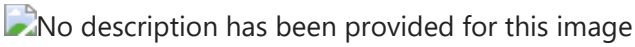


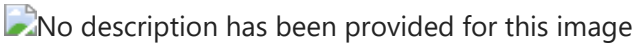
Measure for the distance between two clusters

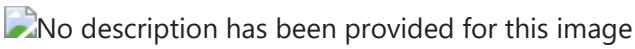
- The closest distance b/w the two clusters is crucial for the hierarchical clustering.
- There are various ways to calculate the distance b/w two clusters, and these ways decided the rule for clustering. These measures are called **Linkage methods:**

- **Single Linkage** - We take minimum distance b/w two clusters
- **Complete Linkage** - We take maximum distance b/w two clusters
- **Average Linkage** - We take average distance b/w two clusters
- **Centroid Linkage** - We take central point and then calculate distance b/w two clusters

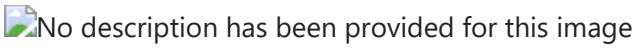








To desing best number of clusters



In [ ]: