EXPERIMENT NO. 6

Title :

Implementation Of any one Hierarchial Clustering method (Agglomerative Single Linkage)

Tool Used : Python

Theory :

Hierarchical Clustering is a method of cluster analysis which seeks to build a hierarchy of clusters. It is divided into two main types: agglomerative (bottom-up) and divisive (top-down). In this experiment, we focus on the agglomerative approach, specifically the Single Linkage method.

Agglomerative Hierarchical Clustering

Agglomerative hierarchical clustering starts with each data point as its own cluster. The algorithm then repeatedly merges the closest pair of clusters until all points are in a single cluster or a stopping criterion is met.

Single Linkage Method

The Single Linkage method, also known as the nearest neighbor method, defines the distance between two clusters as the minimum distance between any single data point in the first cluster and any single data point in the second cluster. This method tends to produce long, chain-like clusters.

Steps of the Single Linkage Method:

Initialization: Start with each data point as a separate cluster.

Distance Calculation: Compute the distance between all pairs of clusters.

Merge Clusters: Identify the pair of clusters with the smallest distance and merge them.

Update Distances: Recalculate the distances between the new cluster and all other clusters.

Repeat: Repeat steps 3 and 4 until all points are in a single cluster or the desired number of clusters is achieved.

Mathematical Representation:

Let (d(x, y)) be the distance between points (x) and (y). The distance between two clusters (A) and (B) in Single Linkage is defined as:

$$d(A,B)=\min\{d(x,y):x\in A,y\in B\}$$

This means that the distance between two clusters is the minimum distance between any pair of points from the two clusters.

Conclusion :

Here, we implemented one Hierarchial Clustering method (Agglomerative Single Linkage)