Solution-Hierarchical Clustering

**Summary:**

As you know, Hierarchical Clustering is an unsupervised machine learning algorithm, which is used to group the unlabeled datasets into a cluster.

In this algorithm, we develop the hierarchy of clusters in the form of a tree.

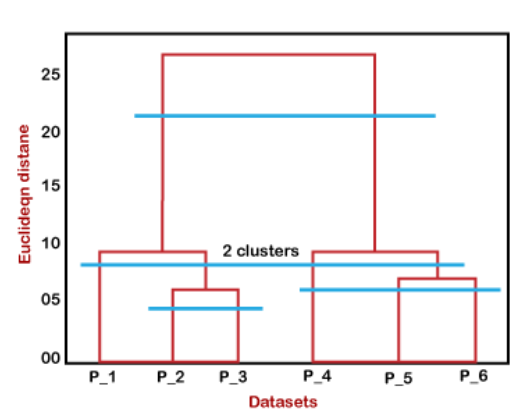
The hierarchical clustering technique has two approaches: Agglomerative and Divisive.

For more information, see <https://www.javatpoint.com/hierarchical-clustering-in-machine-learning>.

**Suggestions:**

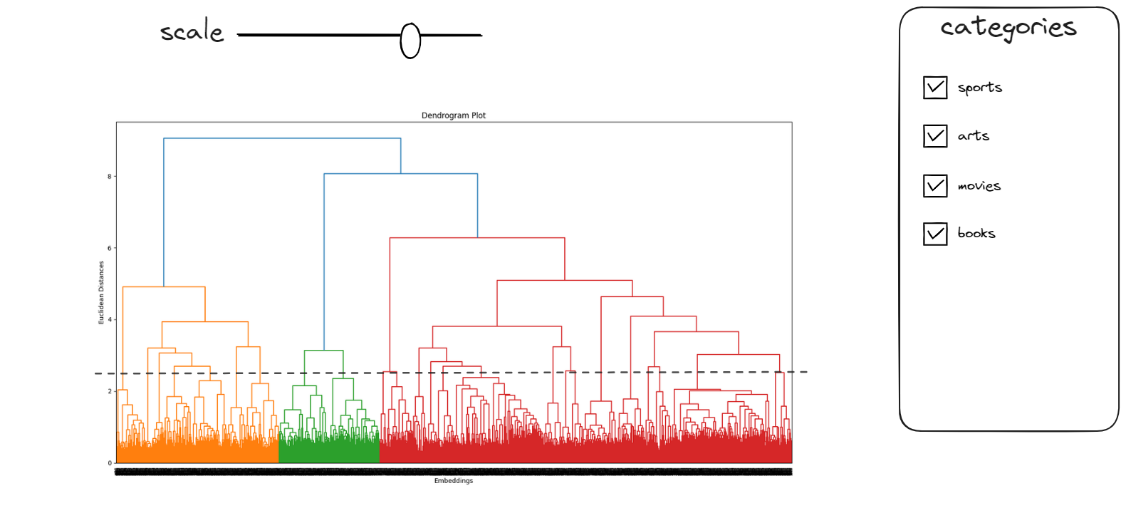
1. Working of Dendrogram

In the dendrogram plot, the Y-axis shows the Euclidean distances between the data points, and the x-axis shows all the data points of the given dataset.



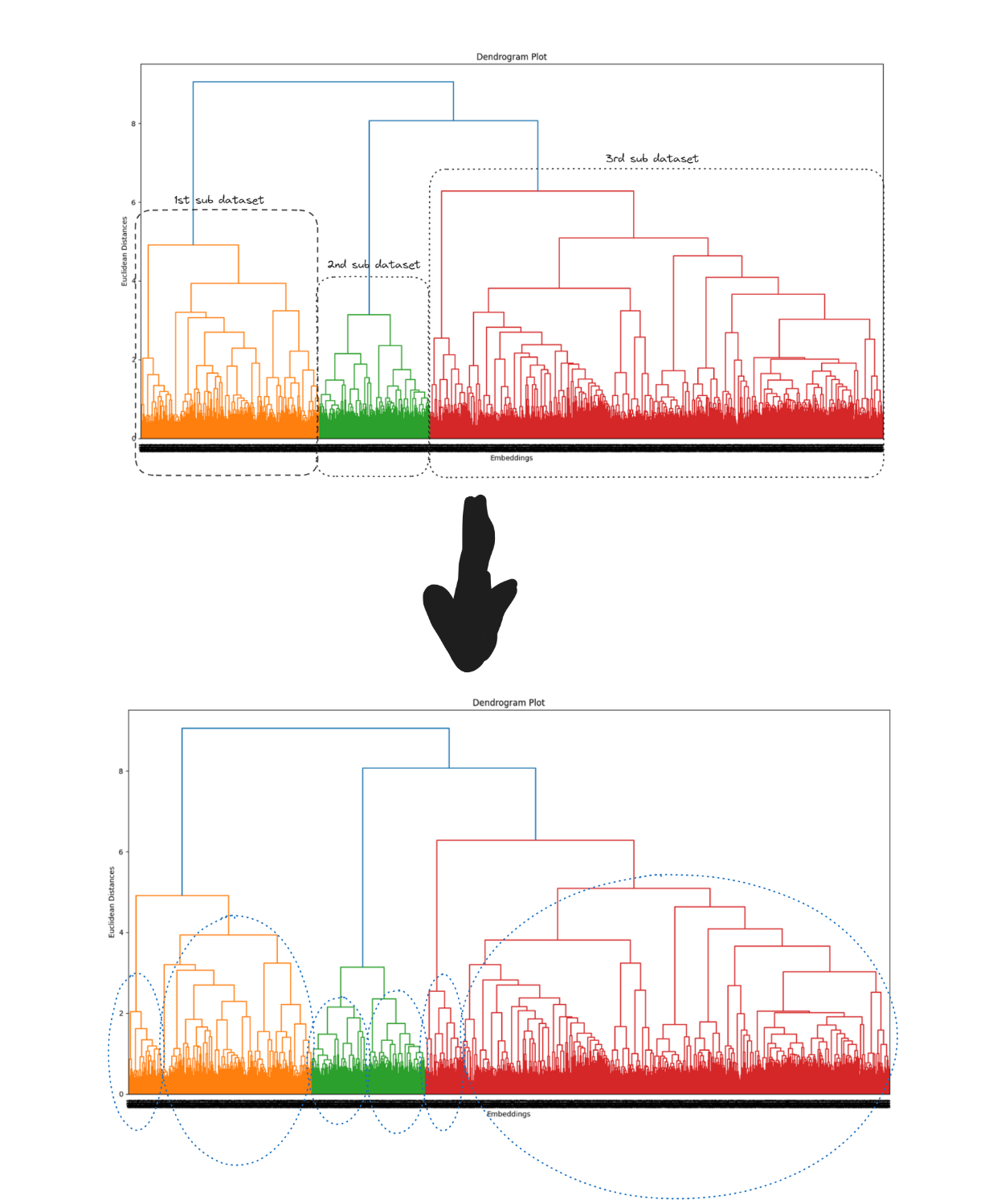
We can cut the dendrogram tree structure at any level as per our requirement, ie input the value of scaling as like similarity, and cut the dendrogram and get the clusters.

Users can control the scale value from 0 to 1, until they obtain the desired number of clusters.



1. Re-clustering sub tree

Hierarchical Clustering can provide the optimal number of clusters. It may be small number 2-4.  
We will re-cluster the sub datasets.



We repeat the re-cluster 3-5 times, depending on amout of dataset. Then, it automatically gets the clustering.

**Conclusion:**

I think that 2nd method is automatic and 1st method is more user-friendly.