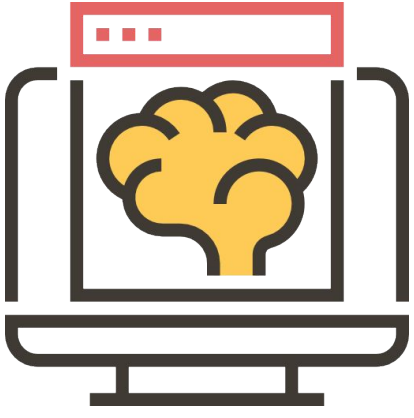


Unsupervised Learning

© All rights reserved by Fireblaze Technologies Pvt. Ltd.

Index

- Introduction To Hierarchical Clustering
- Hierarchical Clustering
- Dendrogram
- Working of Agglomerative Clustering
- Linkage Methods

© All rights reserved by Fireblaze Technologies Pvt. Ltd.

Introduction To Hierarchical Clustering

- Hierarchical clustering starts by assigning **all data** points as their own cluster.
- In this algorithm, we develop the hierarchy of clusters in the form of a tree, and this tree-shaped structure is known as the **dendrogram**.
- It combines the two nearest data point and merges it together to one cluster.

© All rights reserved by Fireblaze Technologies Pvt. Ltd.

Introduction To Hierarchical Clustering

1. Assign each data point to its own cluster.
2. Find closest pair of cluster using euclidean distance and merge them in to single cluster.
3. Calculate distance between two nearest clusters and combine until all items are clustered into a single cluster.
4. In this technique, you can decide the optimal number of clusters by noticing which vertical lines can be cut by horizontal line without intersecting a cluster and covers the **maximum** distance.

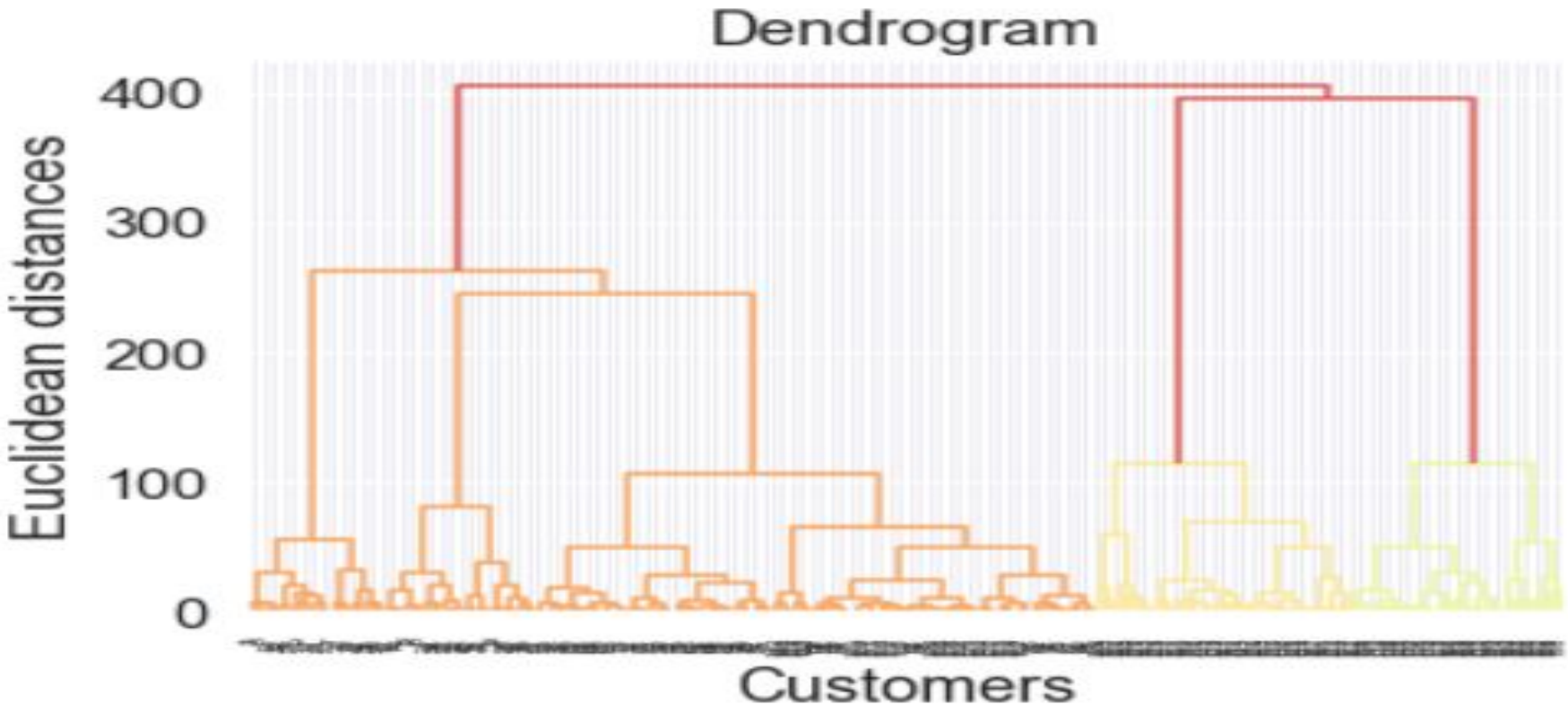
© All rights reserved by Fireblaze Technologies Pvt. Ltd.

Hierarchical Clustering

1. The hierarchical clustering technique has two approaches:
   a. **Agglomerative:** Agglomerative is a bottom-up approach.
   b. **Divisive:** Divisive algorithm is the reverse of the agglomerative algorithm as it is a top-down approach.

© All rights reserved by Fireblaze Technologies Pvt. Ltd.

Dendrogram



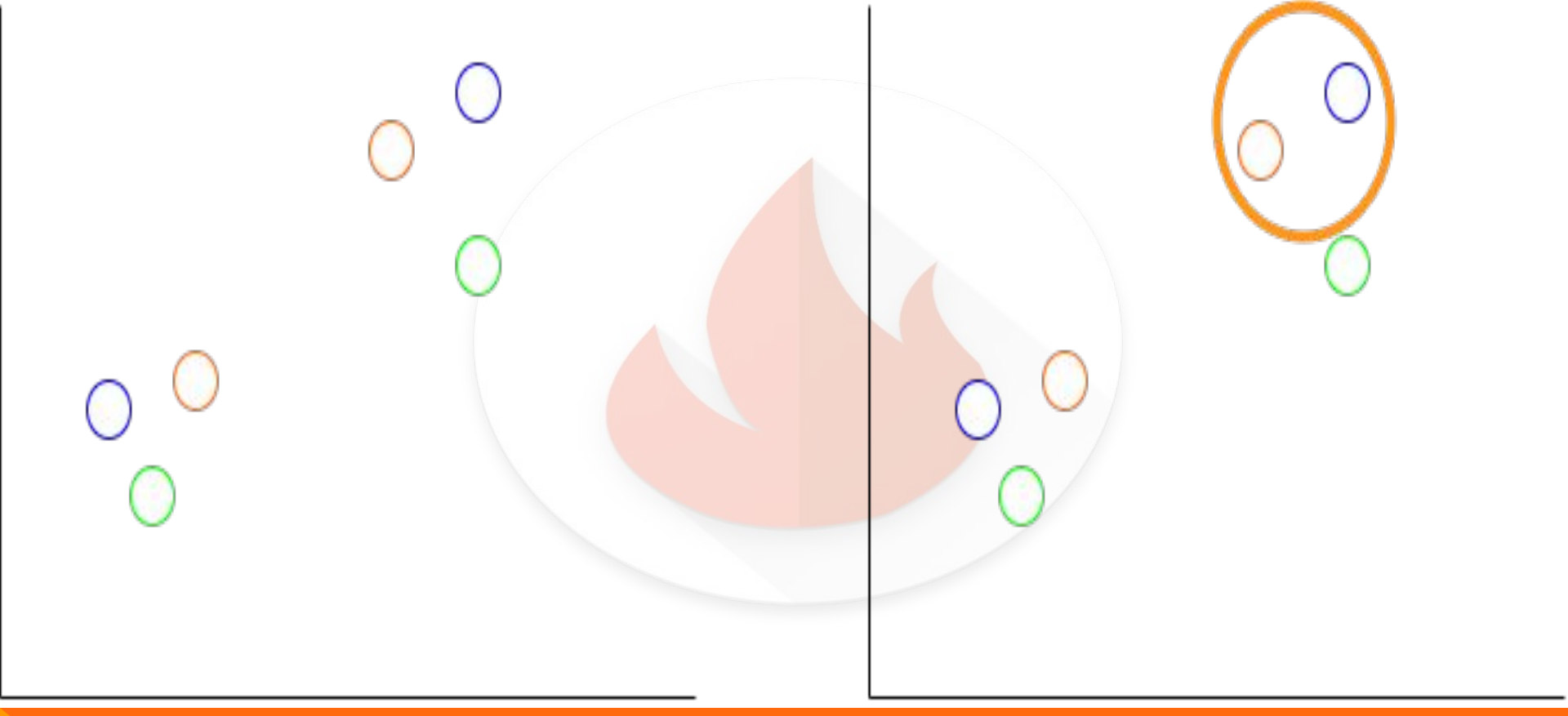
© All rights reserved by Fireblaze Technologies Pvt. Ltd.

Working of Agglomerative Clustering

1. Create each data point as a single cluster. Let's say there are N data points, so the number of clusters will also be N.
2. Take two closest data points or clusters and merge them to form one cluster.
3. Again, take the two closest clusters and merge them together to form one cluster. There will be N-2 clusters.
4. Once all the clusters are combined into one big cluster, develop the dendrogram to divide the clusters as per the problem.

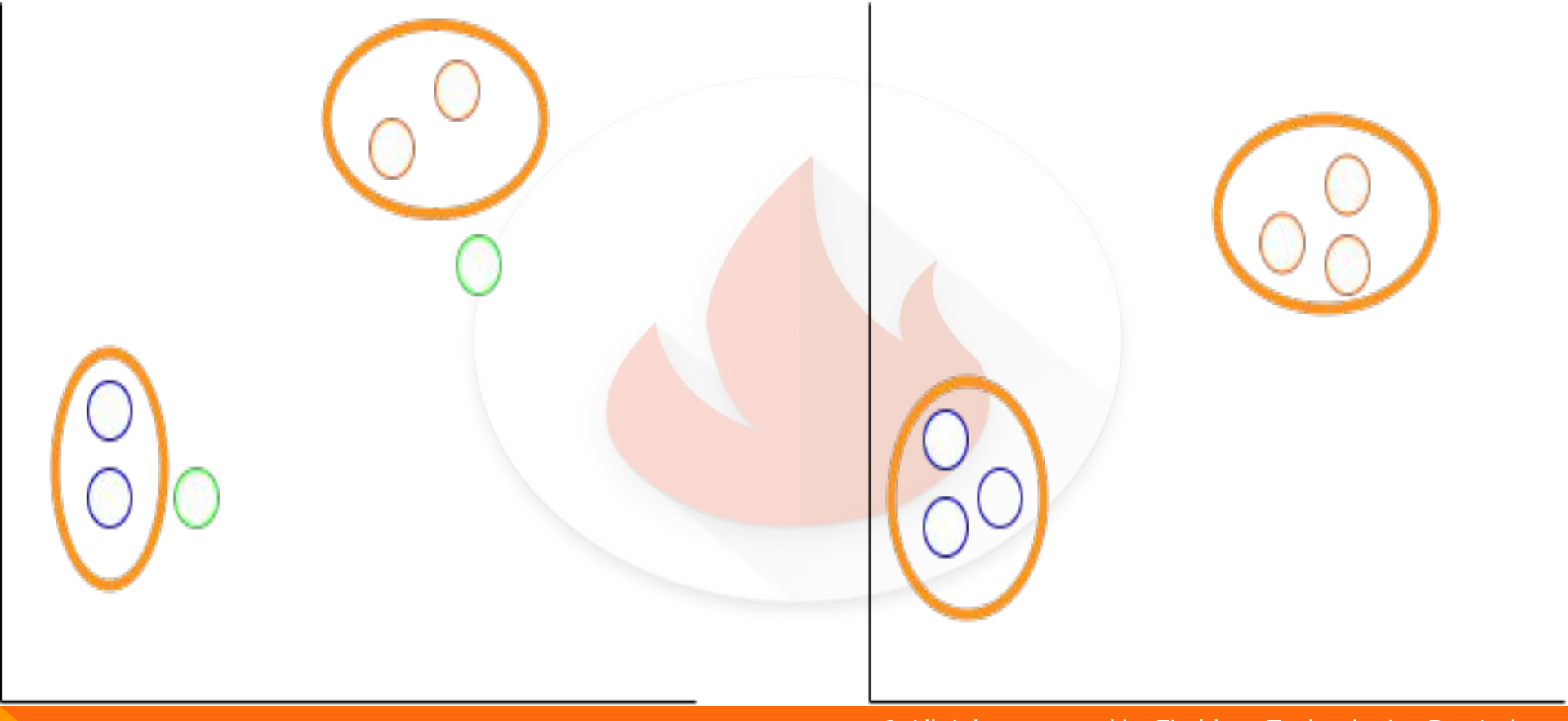
© All rights reserved by Fireblaze Technologies Pvt. Ltd.

Working of Agglomerative Clustering



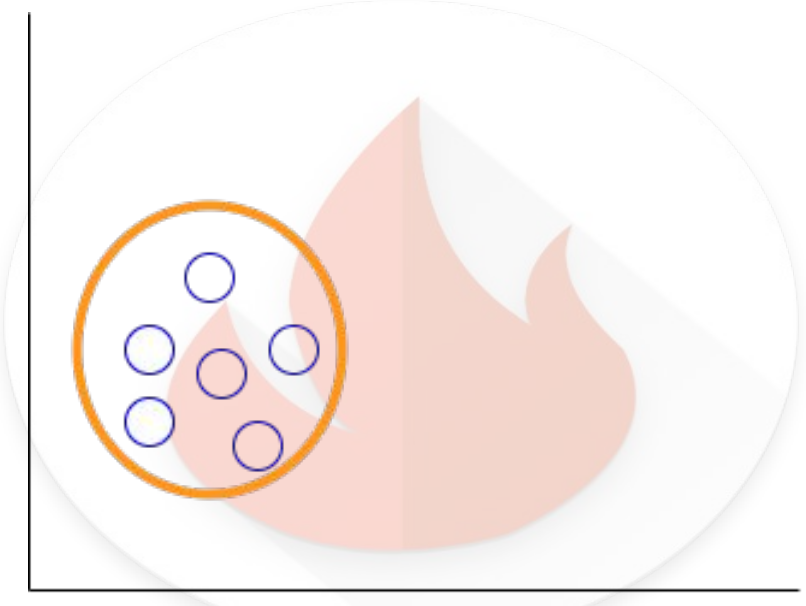
© All rights reserved by Fireblaze Technologies Pvt. Ltd.

Working of Agglomerative Clustering



© All rights reserved by Fireblaze Technologies Pvt. Ltd.

Working of Agglomerative Clustering



© All rights reserved by Fireblaze Technologies Pvt. Ltd.

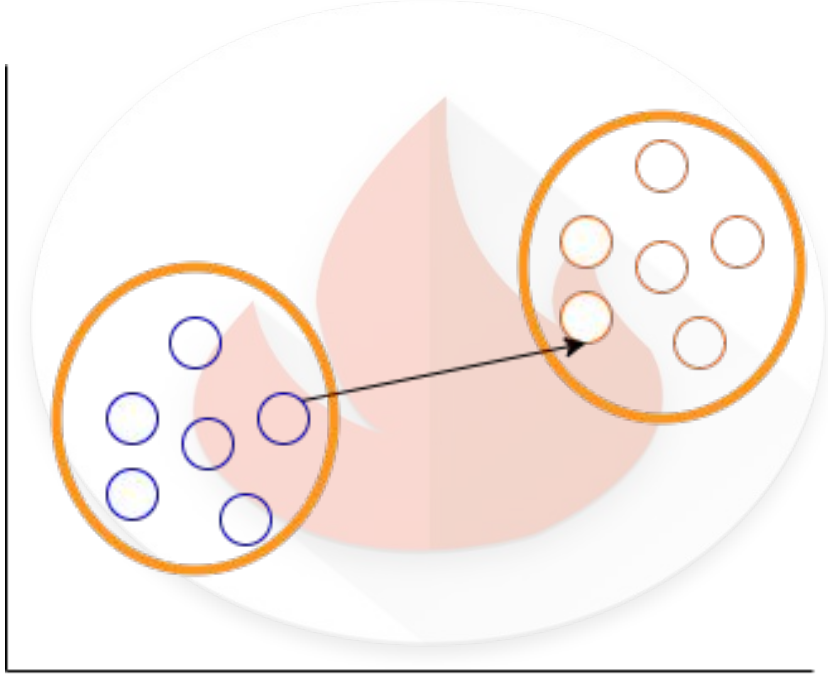
Linkage Methods

1. These measures are called Linkage methods. Some of the popular linkage methods are given below:
    a. Single Linkage
    b. Complete Linkage
    c. Average Linkage
    d. Centroid Linkage

© All rights reserved by Fireblaze Technologies Pvt. Ltd.

Linkage Methods

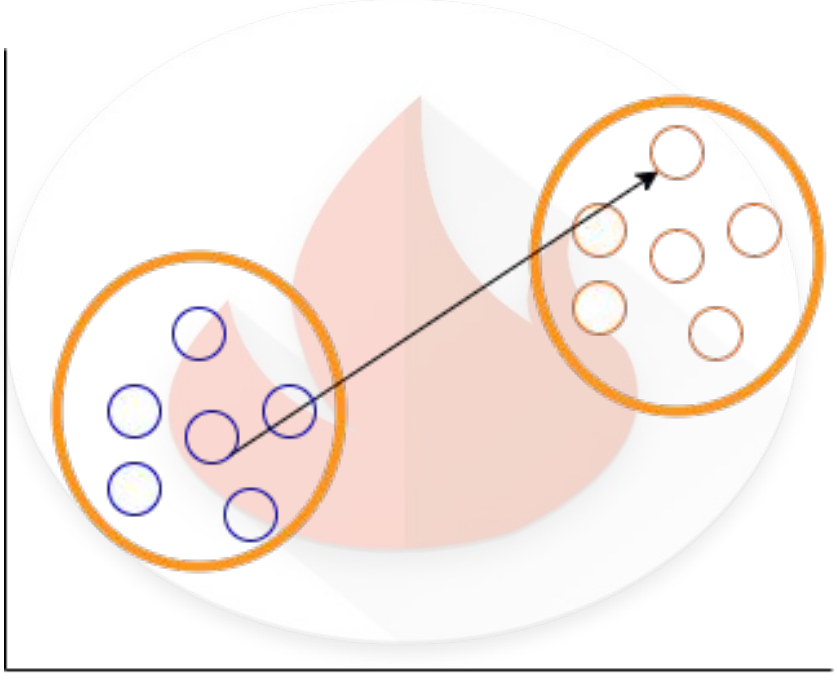
Single Linkage:- It is the Shortest Distance between the closest points of the clusters.



© All rights reserved by Fireblaze Technologies Pvt. Ltd.

Linkage Methods

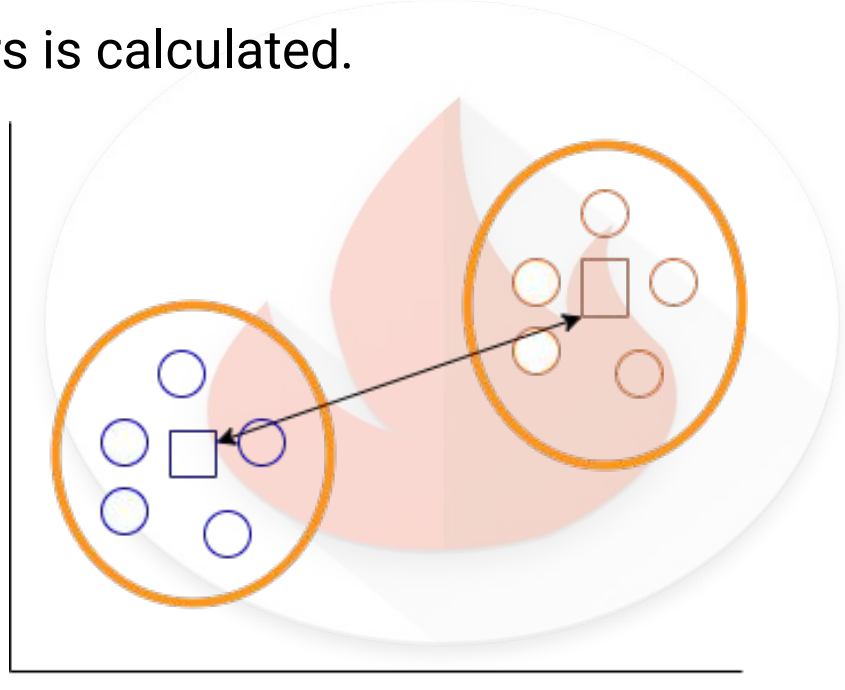
Complete Linkage:- It is the farthest distance between the two points of two different clusters.



© All rights reserved by Fireblaze Technologies Pvt. Ltd.

Linkage Methods

Centroid Linkage:- It is the linkage method in which the distance between the centroid of the clusters is calculated.



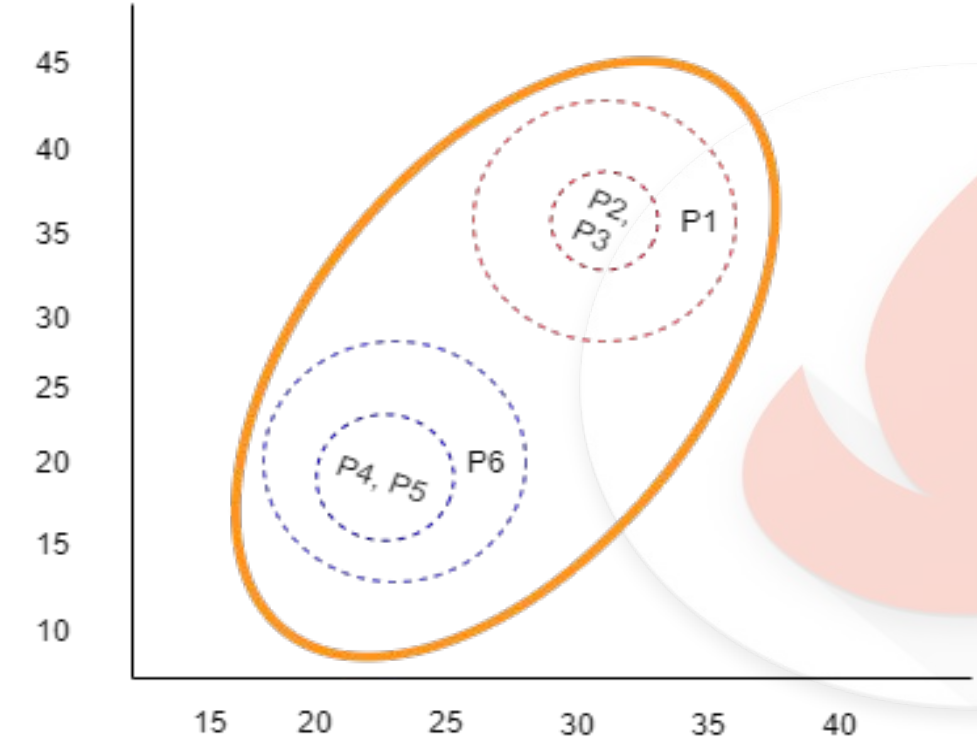
© All rights reserved by Fireblaze Technologies Pvt. Ltd.

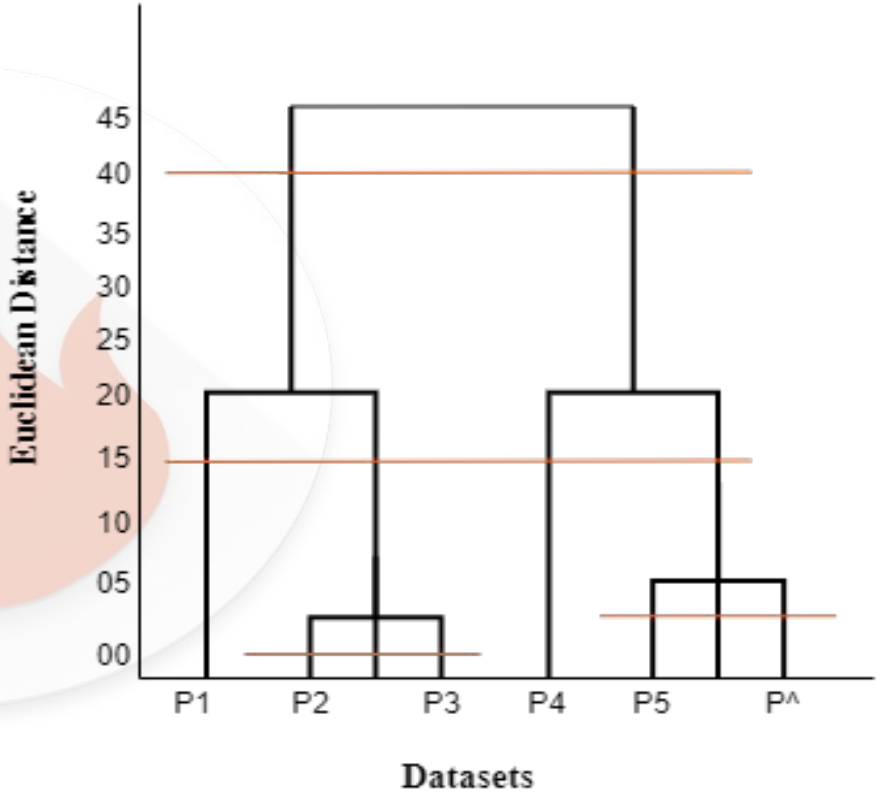
Linkage Methods

Average Linkage:- It is the linkage method in which the distance between each pair of datasets is added up and then divided by the total number of datasets to calculate the average distance between two clusters.

© All rights reserved by Fireblaze Technologies Pvt. Ltd.

Linkage Methods





© All rights reserved by Fireblaze Technologies Pvt. Ltd.

Thank you

© All rights reserved by Fireblaze Technologies Pvt. Ltd.