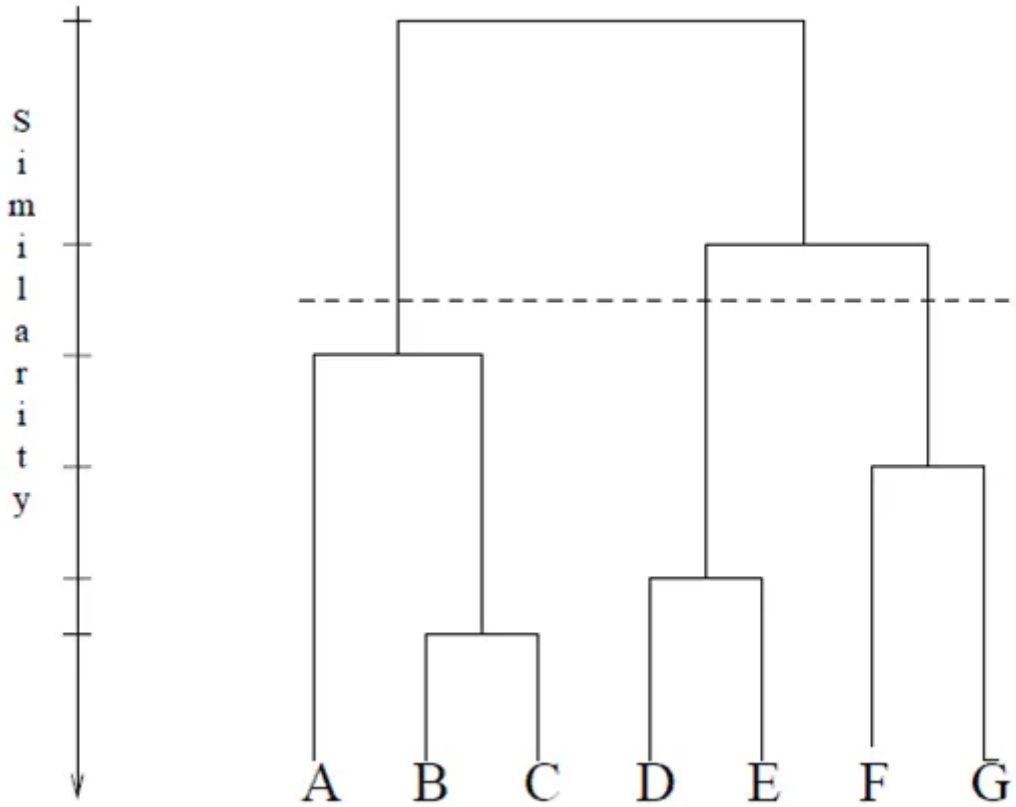
Learn with an example : Hierarchical Clustering

Rohan Joseph · Follow
2 min read · Jun 18, 2018

What is hierarchical clustering (agglomerative) ?

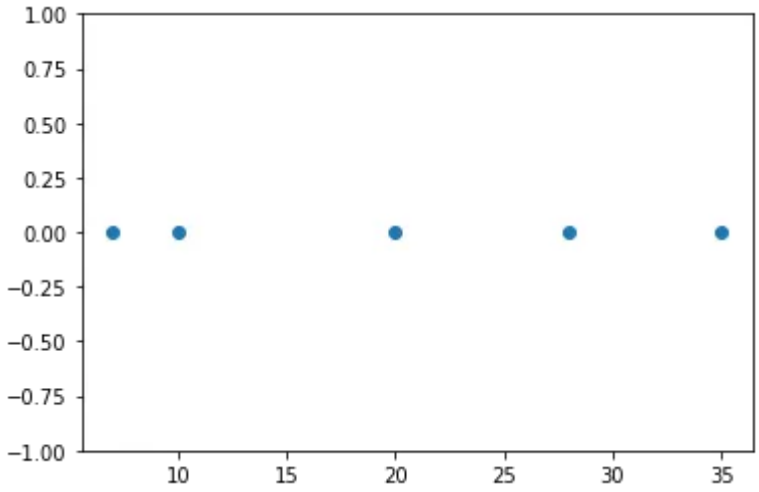
Clustering is a data mining technique to group a set of objects in a way such that objects in the same cluster are more similar to each other than to those in other clusters.

In hierarchical clustering, we assign each object (data point) to a separate cluster. Then compute the distance (similarity) between each of the clusters and join the two most similar clusters. Let's understand further by solving an example.



Objective : For the one dimensional data set **{7,10,20,28,35}**, perform hierarchical clustering and plot the dendogram to visualize it.

Solution : First, let's the visualize the data.

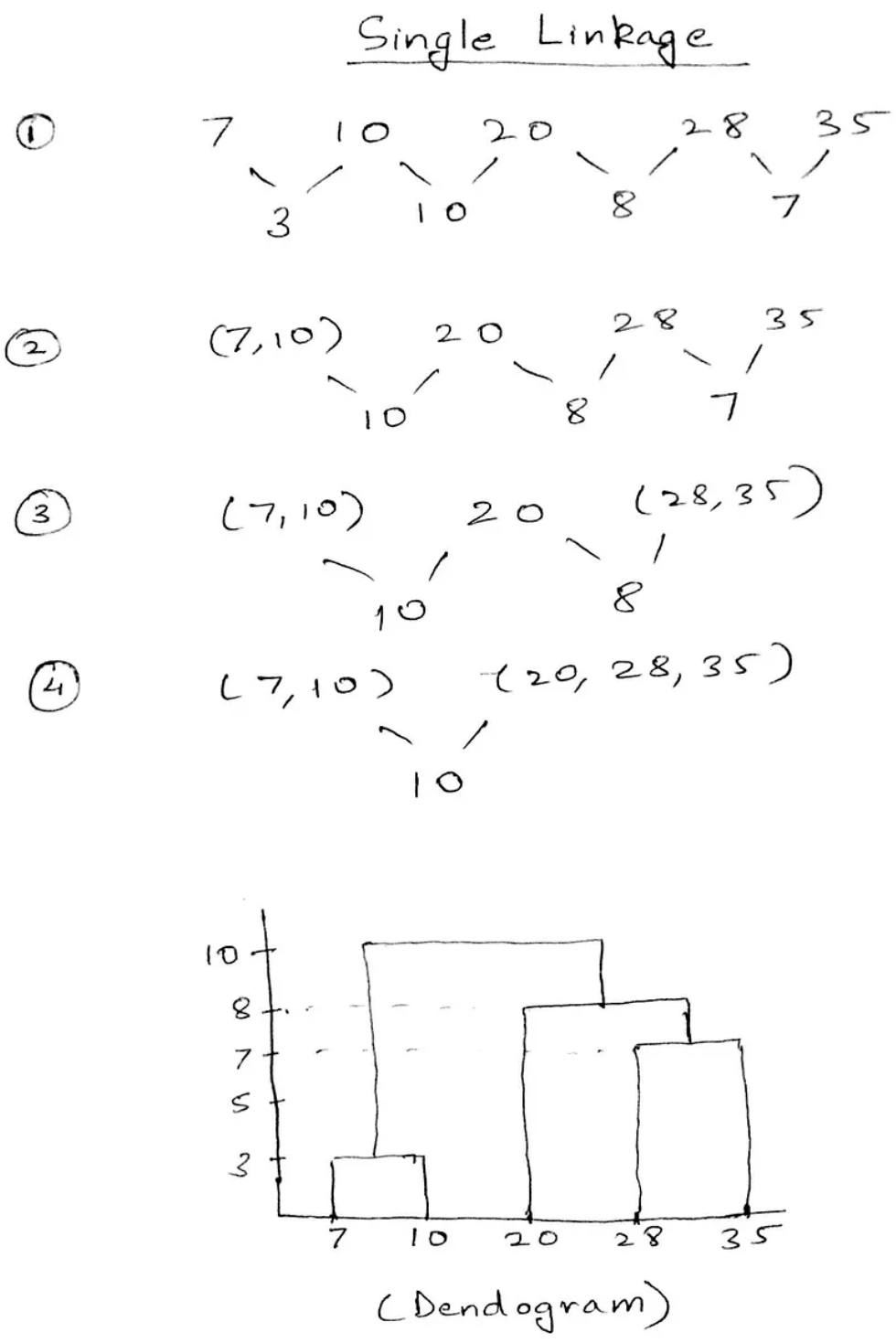


Observing the plot above, we can intuitively conclude that:

1. The first two points (7 and 10) are close to each other and should be in the same cluster
2. Also, the last two points (28 and 35) are close to each other and should be in the same cluster
3. Cluster of the center point (20) is not easy to conclude

Let's solve the problem by hand using both the types of agglomerative hierarchical clustering :

1. **Single Linkage :** In single link hierarchical clustering, we merge in each step the two clusters, whose two closest members have the smallest distance.

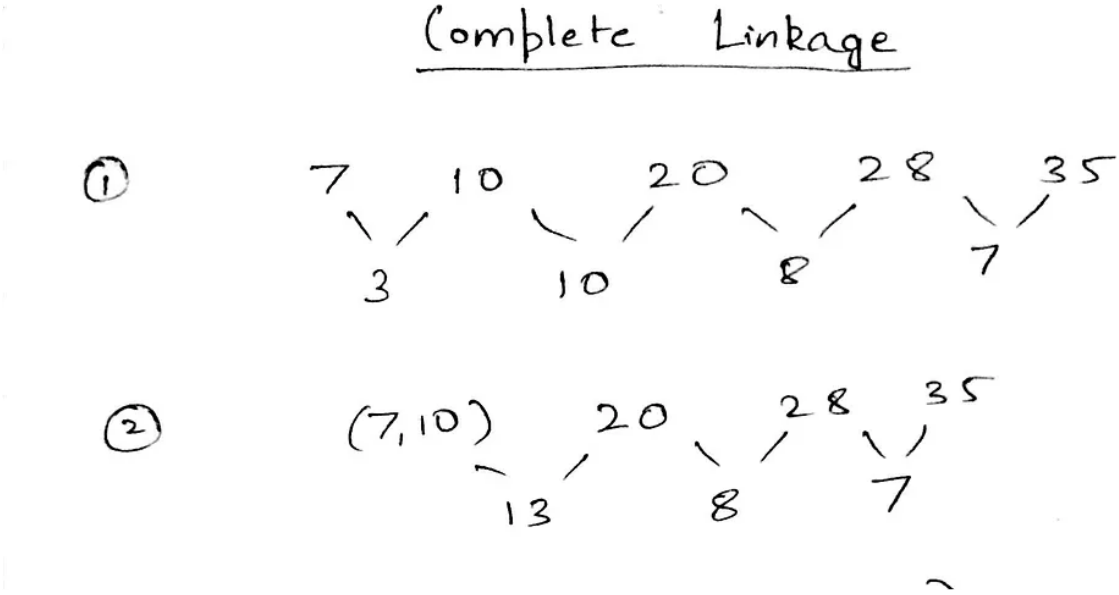


Using single linkage two clusters are formed :

Cluster 1 : (7,10)

Cluster 2 : (20,28,35)

2. **Complete Linkage :** In complete link hierarchical clustering, we merge in the members of the clusters in each step, which provide the smallest maximum pairwise distance.



Using complete linkage two clusters are formed :

Cluster 1 : (7,10,20)

Cluster 2 : (28,35)

Conclusion : Hierarchical clustering is mostly used when the application requires a hierarchy, e.g creation of a taxonomy. However, they are expensive in terms of their

computational and storage requirements.

— —

Connect on LinkedIn

Data Science | Analytics | Machine Learning | Deep Learning | Data

Follow

Written by Rohan Joseph

1.4K Followers

Data Scientist @ Apple https://www.linkedin.com/in/rohan-joseph-b39a86aa/

More from Rohan Joseph