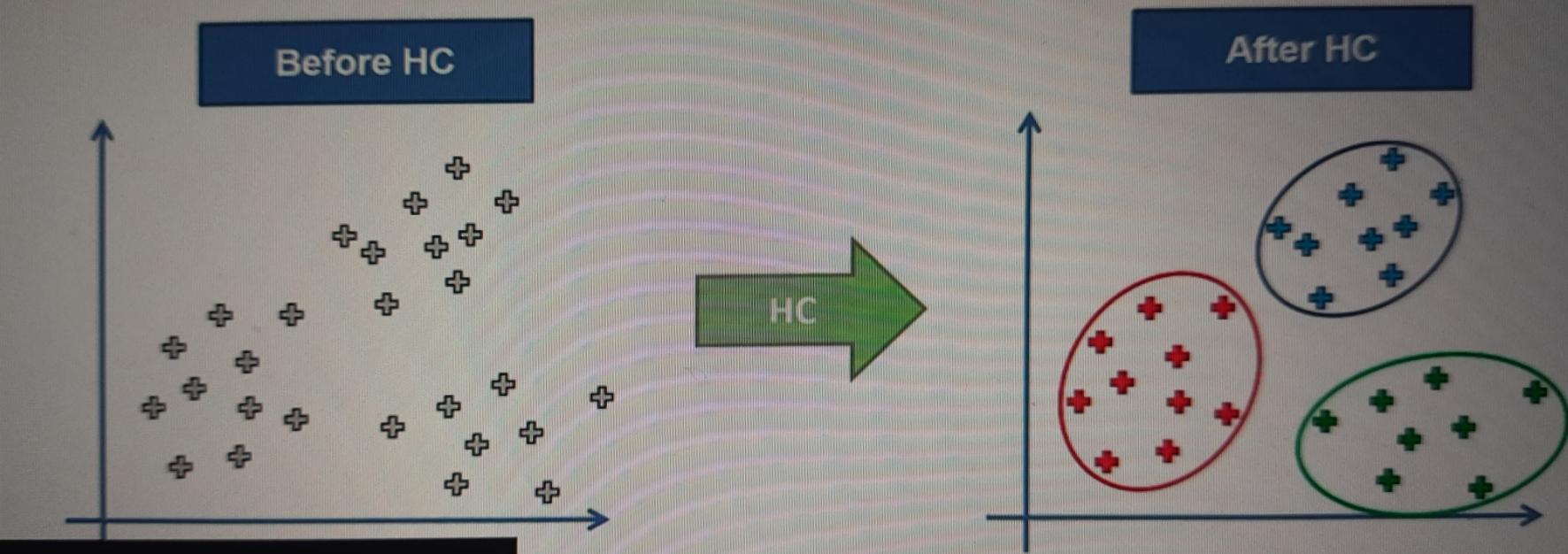


HC Intuition: Understanding HC



What HC does for you



written a note here.

Same as K-Means but different process

NOTE:

Agglomerative

&

Divisive

written a note here.

Agglomerative HC

STEP 1: Make each data point a single-point cluster → That forms N clusters



STEP 2: Take the two closest data points and make them one cluster → That forms N-1 clusters



STEP 3: Take the two closest clusters and make them one cluster → That forms N - 2 clusters

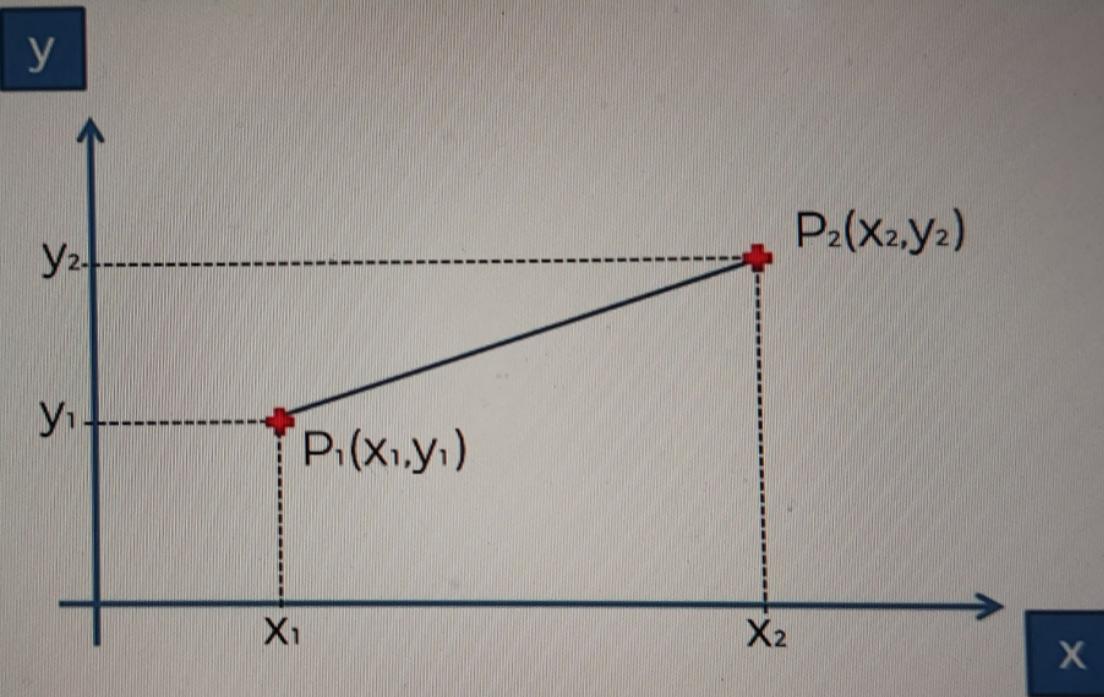


STEP 4: Repeat STEP 3 until there is only one cluster



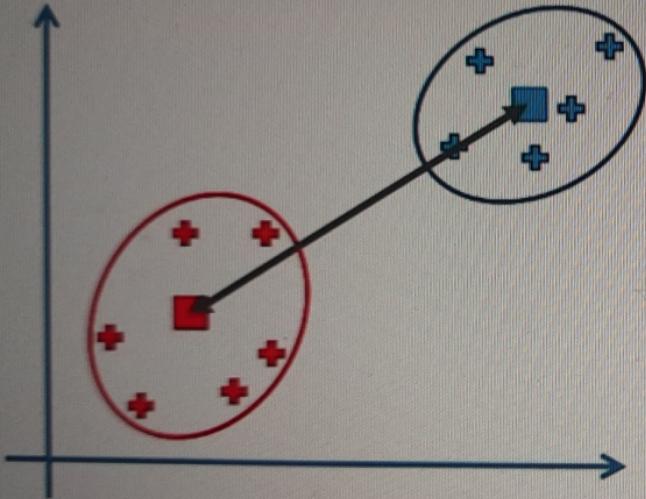
FIN

Euclidean Distance



$$\text{Euclidean Distance between } P_1 \text{ and } P_2 = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Distance Between Clusters

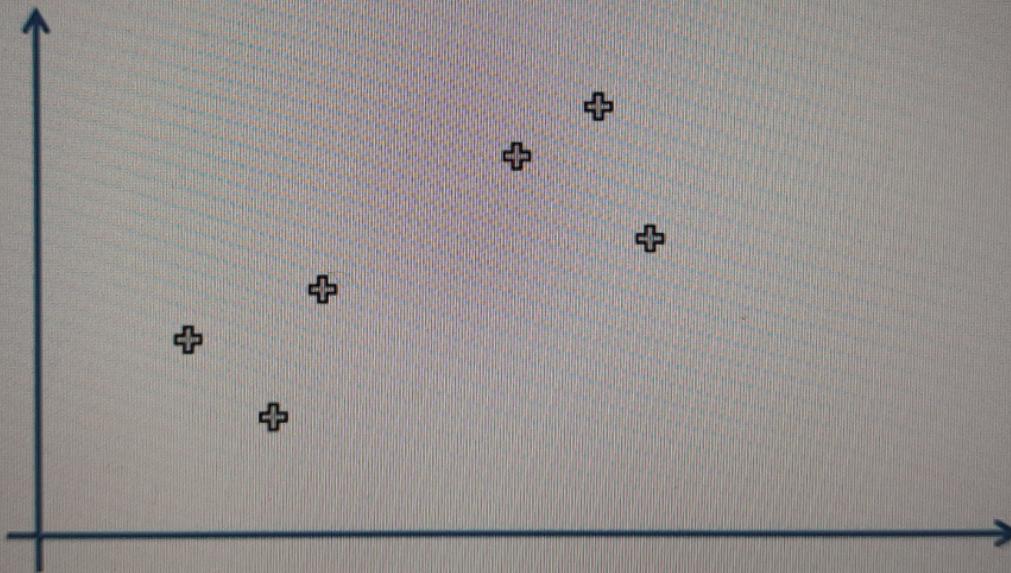


Distance Between Two Clusters:

- Option 1: Closest Points
- Option 2: Furthest Points
- Option 3: Average Distance
- Option 4: Distance Between Centroids

Agglomerative HC

Consider the following dataset of $N = 6$ data points



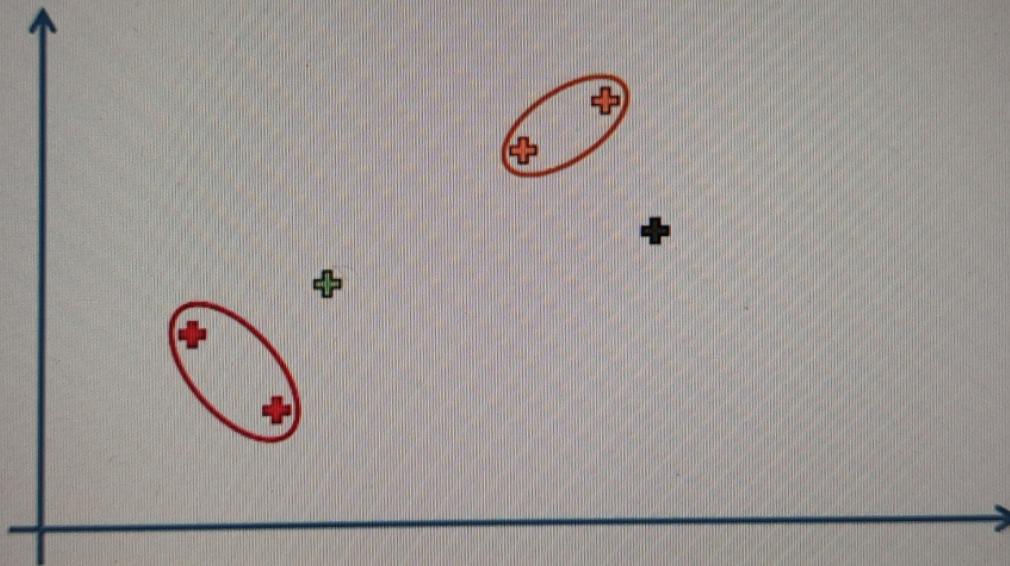
Agglomerative HC

STEP 1: Make each data point a single-point cluster → That forms 6 clusters



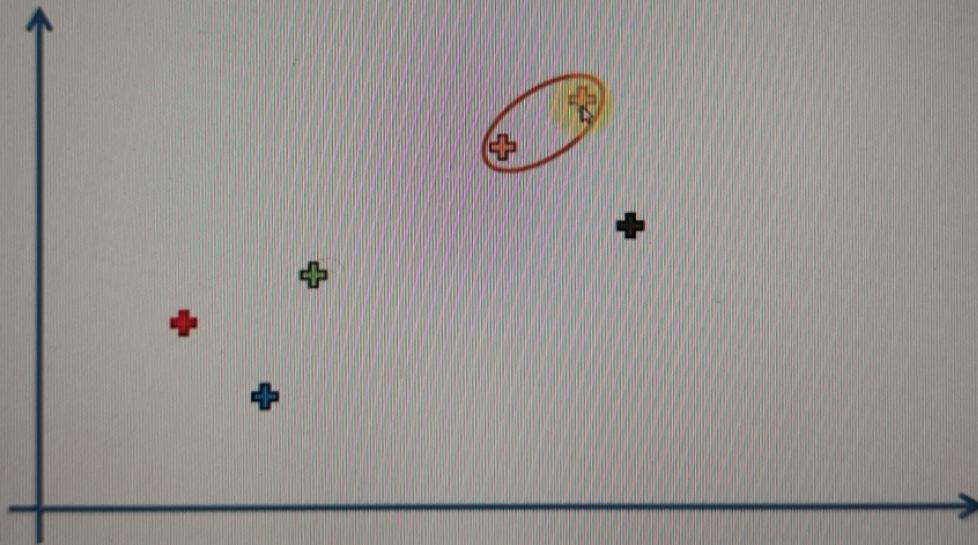
Agglomerative HC

STEP 3: Take the two closest clusters and make them one cluster
→ That forms 4 clusters



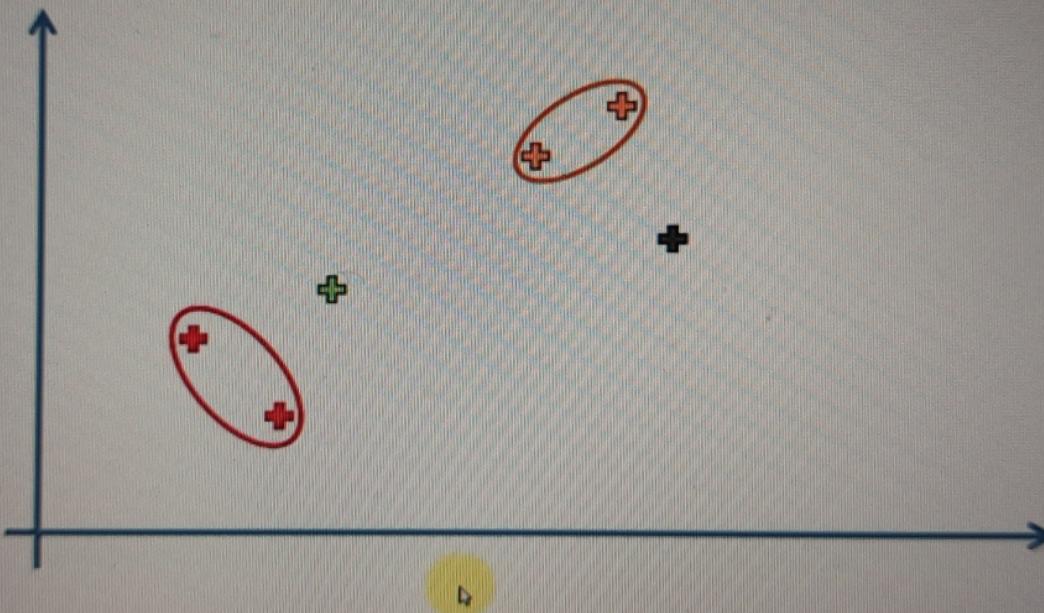
Agglomerative HC

STEP 2: Take the two closest data points and make them one cluster
→ That forms 5 clusters



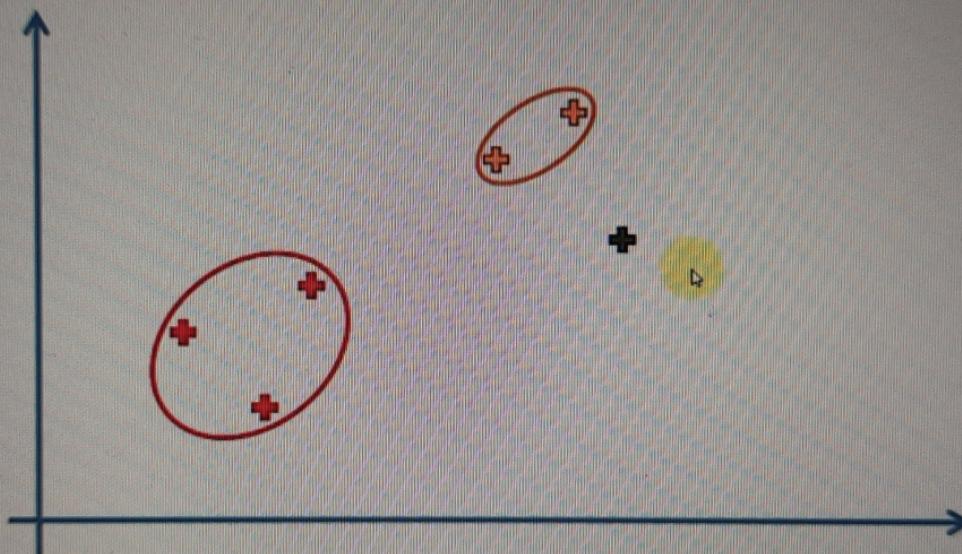
Agglomerative HC

STEP 3: Take the two closest clusters and make them one cluster
→ That forms 4 clusters



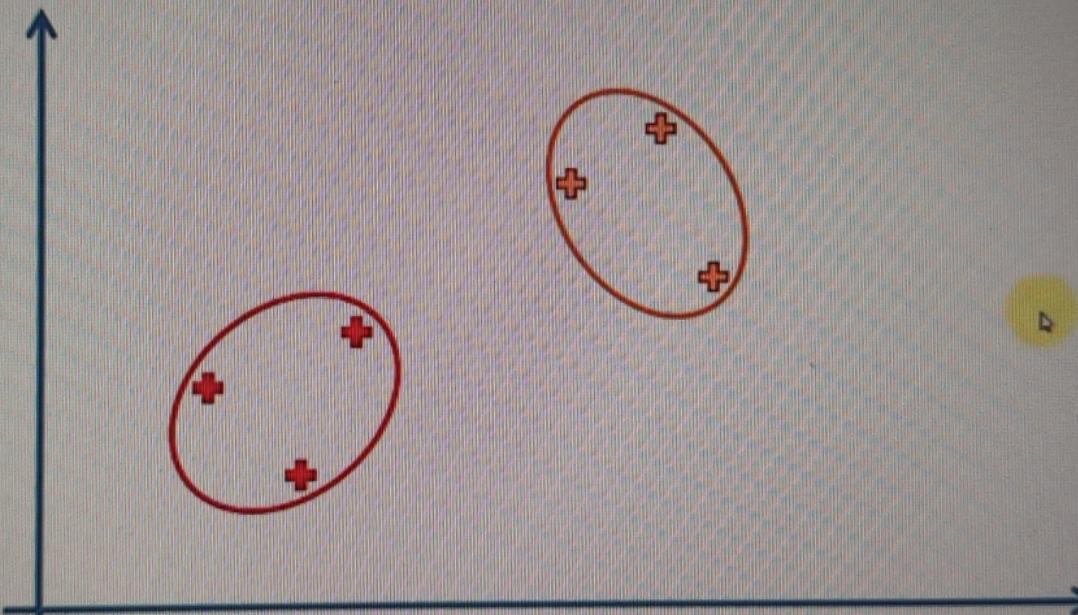
Agglomerative HC

STEP 4: Repeat STEP 3 until there is only one cluster



Agglomerative HC

STEP 4: Repeat STEP 3 until there is only one cluster



Agglomerative HC

STEP 4: Repeat STEP 3 until there is only one cluster

