

# Hierarchical clustering (Agglomerative) Single Link

Data Points :

	$A_1$	$A_2$	$A_3$	$A_4$	$A_5$	$A_6$	$A_7$	$A_8$	$A_9$	$A_{10}$	$A_{11}$
$A_1$	0										
$A_2$	1.4	0									
$A_3$	5.9	4	0								
$A_4$	2.3	3.6	6.7	0							
$A_5$	2.8	3.16	4.2	3.0	0						
$A_6$	3.1	4	5.6	2.2	1.4	0					
$A_7$	5.8	6.3	6.3	5.0	3.1	2.8	0				
$A_8$	5.3	5	3	6.0	3.0	4.1	3.6	0			
$A_9$	5.6	5.8	5.09	5.3	2.8	3.1	1.4	2.2	0		
$A_{10}$	9.05	10.19	11.6	7.0	7.6	6.3	5.6	9.2	7.07	0	
$A_{11}$	7.07	8	8.94	5.3	5.09	4	2.8	6.4	4.2	2.8	0

Single linkage: This is the distance between the closest members of the two clusters

$\Rightarrow$  Euclidean distance :

$$D[(x, y), (a, b)] = \sqrt{(x-a)^2 + (y-b)^2}$$

$$A_1 = (3, 7)$$

$$A_2 = (2, 6)$$

$$A_3 = (2, 2)$$

$$A_4 = (5, 3)$$

$$A_5 = (5, 5)$$

$$A_6 = (6, 6)$$

$$A_7 = (8, 4)$$

$$A_8 = (5, 2)$$

$$A_9 = (7, 3)$$

$$A_{10} = (12, 8)$$

$$A_{11} = (10, 6)$$

- $\Rightarrow$  from the above data points. the closest points are  $A_1 = (3, 7)$  and  $A_2 = (2, 4)$   
 $\Rightarrow$  Now, add a new point  $(2.5, 6.5)$  to the list, which is formed by doing average of  $A_1$  and  $A_2$  i.e.  $\left(\frac{3+2}{2}, \frac{7+4}{2}\right) = (2.5, 6.5)$   
 $\Rightarrow$  Now, we will represent the pairwise distances between the updated data points in the form of table.

	$(2.5, 6.5)$	$(2, 2)$	$(5, 8)$	$(5, 5)$	$(6, 6)$	$(8, 4)$	$(5, 2)$	$(7, 3)$	$(12, 8)$	$(10, 6)$
$(2.5, 6.5)$	0									
$(2, 2)$	4.5	0								
$(5, 8)$	2.92	6.71	0							
$(5, 5)$	2.92	4.2	3	0						
$(6, 6)$	3.5	5.6	2.24	1.41	0					
$(8, 4)$	6.04	6.3	5	3.61	2.8	0				
$(5, 2)$	5.15	3	6	3	4.12	3.61	0			
$(7, 3)$	5.70	5.1	5.39	2.83	3.16	1.41	2.24	0		
$(12, 8)$	9.62	11.6	7.0	7.62	6.32	5.6	9.2	7.07	0	
$(10, 6)$	7.52	8.9	5.39	5.10	4	2.8	6.4	4.24	2.83	0

$\Rightarrow$  Again, repeat the same step

$\Rightarrow$  find the closest points

$\Rightarrow$  Add a new point by doing average of both the closest points

$\Rightarrow$  Update the table with new point

$\Rightarrow$  Repeat the above steps until all the data points are part of one cluster



let's draw the dendrogram

Single link hierarchical clustering diagram

