Kmeans_22_6_25

Sunday, June 22, 2025 8:01 AM

K-means Algorithm (EM)

Steps

Example

unsupervised clustering

K = 3

points	Coordinates
A_1	(2, 10)
A_2	(2, 6)
A_3	(11, 11)
A_4	(6, 9)
A_5	(6, 4)
A_6	(1, 2)
A_7	(5, 10)
A_8	(4, 9)
A_9	(10, 12)
A_{10}	(7, 5)
A_{11}	(9, 11)
A_{12}	(4, 6)
A_{13}	(3, 10)
A_{14}	(3, 8)
A_{15}	(6, 11)

2 dimension data

Centroid 1 - (2, 6)

Centroid 2 - (5, 10)

Centroid 3 - (6, 11)

l_2 norm

Euclidean = $\sqrt{(x - x_1)^2 + (y - y_1)^2}$

Step 1

Step 1 → Euclidean = $\sqrt{(x-x_1)^2+(y-y_1)^2}$

Point	Distance from C1 (2,6)	Distance from C2 (5,6)	Distance from C3 (6,11)	Assignment
$A_1(2,10)$	4	3	4.12	C_2
$A_2(2,6)$	0	–	–	C_1
$A_3(11,11)$	– $\sqrt{9^2+5^2}$	$\sqrt{6^2+1^2}$	$\sqrt{5^2+0^2}$	C_3
$A_4(6,9)$	–	–	–	

Cluster 1 – $A_2(2,6)$, $A_5(6,4)$, $A_6(1,2)$, $A_{10}(7,5)$, $A_{12}(4,6)$, $A_{14}(3,8)$ ⟹ $\left(\frac{2+6+1+7+4+3}{6}, \dots\right)$ (3.8, 5.16)

Cluster 2 – $A_1(2,10)$, $A_4(6,9)$, $A_7(5,10)$, $A_8(4,9)$, $A_{13}(3,10)$ ⟹ (4, 9.6)

Cluster 3 – $A_3(11,11)$, $A_9(10,12)$, $A_{11}(9,11)$, $A_{15}(6,11)$ ⟹ (9, 11.25)

Step 2

Points	Dist from C_1 (3.8, 5.16)	Dist from C_2 (4, 9.6)	Dist from C_3 (9, 11.25)
$A_1(2,10)$			

A_1 (2, 1.5)

A_2 (2, 6)

Centroid 1 = (4, 4.6) − [A_2, A_5, A_6, A_9, A_{11}, A_{10}, A_{12}]

Centroid 2 = (4.143, 9.57) − [A_1, A_4, A_7, A_8, A_{13}, A_{14}, A_{15}]

Centroid 3 = (10, 11.33) − [A_3, A_9, A_{11}]

Step 3

Points	Dist from C_1 (4, 4.6)	Dist from C_2 (4.143, 9.57)	Dist from C_3 (10, 11.33)	Assignment
A_1				
A_2				

Hierarchical Clustering

Example

	A_1	A_2	A_3	A_4	A_5
A_1	0	0.10	0.41	0.55	0.35
A_2	0.10	0	0.64	0.47	0.98
A_3	0.41	0.64	0	0.44	0.85
A_4	0.55	0.47	0.44	0	0.76
A_5	0.35	0.98	0.85	0.76	0

Combine A_1 & A_2

A5

Combine A_1 & A_2

distance $\{(A_1, A_2), A_3\} = \min(dist(A_1, A_3), dist(A_2, A_3))$

$= \min(0.41, 0.64)$

$= 0.41$

distance $\{(A_1, A_2), A_4\} = \min\{0.55, 0.47\}$

$= 0.47$

distance $\{(A_1, A_2), A_5\} =$

$= 0.35$

	A_{12}	A_3	A_4	A_5
A_{12}	0	0.41	0.47	0.35
A_3	0.41	0	0.44	0.85
A_4	0.47	0.44	0	0.76
A_5	0.35	0.85	0.76	0

dist $\{(A_{12}\ A_5), A_3\} = 0.41$

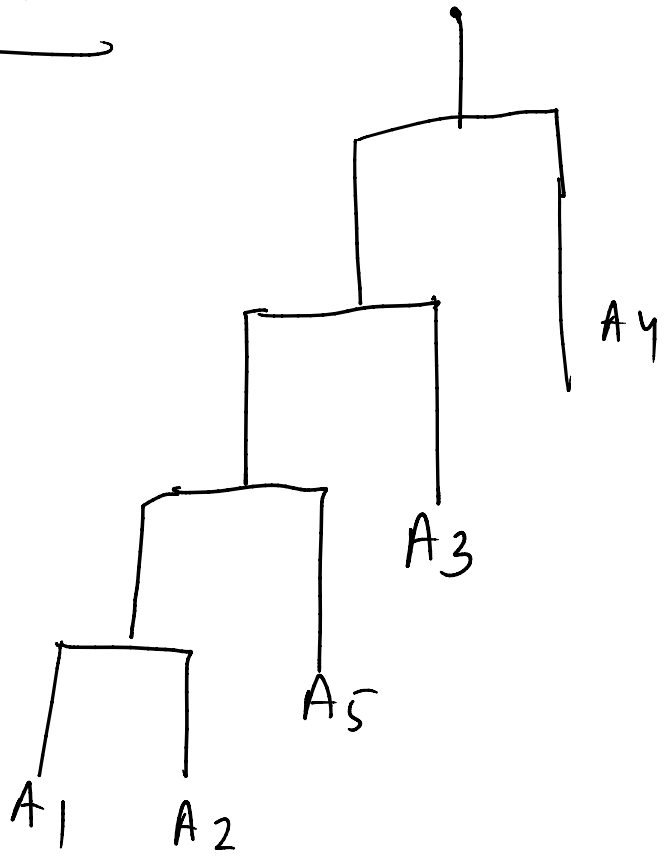
dist $\{(A_{12}\ A_5), A_4\} = 0.47$

	A_{125}	A_3	A_4
A_{125}	0	0.41	0.47
A_3	0.41	0	0.44
A_4	0.47	0.44	0

distance$\left((A_{125}, A_3), A_4\right) =$ 0.44

	A_{1253}	A_4
A_{1253}	0	0.44
A_4	0.44	0

Dendogram



Q.

	A_1	A_2	A_3	A_4	A_5	A_6
A_1	0	0.24	0.22	0.37	0.34	0.23
A_2	0.24	0	0.15	0.2	0.14	0.25
				0.16	0.28	0.11

A_2	0.24	0	0.15	[illegible]	[illegible]	0.25
A_3	0.22	0.15	0	0.16	0.28	0.11
A_4	0.37	0.2	0.16	0	0.29	0.22
A_5	0.34	0.14	0.28	0.29	0	0.39
A_6	0.23	0.25	0.11	0.22	0.39	0