

K-Means Clustering Algorithm :

Cluster the following eight point (with (x, y) representing locations) into three clusters :

$A_1(2, 10)$, $A_2(2, 5)$, $A_3(8, 4)$, $A_4(5, 8)$, $A_5(7, 5)$, $A_6(6, 4)$
 $A_7(1, 2)$, $A_8(4, 9)$

- Initial cluster centers are : $A_1(2, 10)$, $A_4(5, 8)$ & $A_7(1, 2)$

- The distance function betⁿ two points $a = (x_1, y_1)$ & $b = (x_2, y_2)$ is defined as -

$$P(a, b) = |x_2 - x_1| + |y_2 - y_1|$$

use K-means Algo. to find the three clusters centers after the second iteration.

Solⁿ :- we follow the above K-means clustering Algo:

I-1 :- * calculating distance betⁿ $A_1(2, 10)$ & $C_1(2, 10)$

$$P(A_1, C_1) = |x_2 - x_1| + |y_2 - y_1| \cong \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

$$= |2 - 2| + |10 - 10| = 0$$

* calculating distance betⁿ $A_1(2, 10)$ & $C_2(5, 8)$

$$P(A_1, C_2) = |x_2 - x_1| + |y_2 - y_1|$$

$$= |5 - 2| + |8 - 10|$$

$$= 3 + 2 = 5$$

* calculating distance betⁿ $A_1(2, 10)$ & $C_3(1, 2)$:

$$P(A_1, C_3) = |x_2 - x_1| + |y_2 - y_1|$$

$$= |1 - 2| + |2 - 10|$$

$$= 1 + 8 = 9$$

In the similar manner, we calculate the distance of other points from each of the center of the three clusters.

- * draw a table showing all results.
- * Using the table we decide with points belongs to which cluster
- * The given point belongs to that cluster whose center is nearest to it.

Given points	Dist from center (2,10) of C-1	Dist. from center (5,8) of C-2	Dist from center (1,2) of C-3.	points belongs to cluster
A ₁ (2,10)	0	5	9	C ₁
A ₂ (2,5)	5	6	4	C ₃
A ₃ (8,4)	12	7	9	C ₂
A ₄ (5,8)	5	0	10	C ₂
A ₅ (7,5)	10	5	9	C ₂
A ₆ (6,4)	10	5	7	C ₂
A ₇ (1,2)	9	10	0	C ₃
A ₈ (4,9)	3	2	10	C ₂

From here, New clusters are :-
 $C_1 = A_1(2,10)$, $C_2 = A_3(8,4), A_4(5,8), A_5(7,5), A_6(6,4), A_8(4,9)$
 $C_3 = A_2(2,5), A_7(1,2)$

* Now we recompute the new cluster.
 * The new cluster is computed by taking mean of all points contained in that cluster.

$$C_1 = (2,10) , C_2 = (8+5+7+6+4)/5 = (6,6) \\ (4+8+5+4+9)/5$$

$$C_3 = (2+1)/2, (5+2)/2 \\ = (1.5, 3.5)$$

Ex:- calculating distance betⁿ A₁(2,10) & C₁(2,10)

$$P(A_1, C_1) = |x_2 - x_1| + |y_2 - y_1| = |2 - 2| + |10 - 10| = 0$$

$$A_1(2,10) \& C_2(6,6) = |6 - 2| + |6 - 10| = 8$$

$$A_1(2,10) \& C_3(1.5, 3.5) = |1.5 - 2| + |3.5 - 10| = 0.5 + 6.5 = 7$$

In the similar manner, we calculate the distance of the other points from each of the center of the three cluster