

**Answer:** In question, it has been mentioned that A and C are two centers. That means, we have to create two clusters. Since, we will use K-mean algorithm, our centroid will be eventually changed. Following Steps need to be followed according to K-mean algorithm.

Step-1: decide number of cluster, in our case, we will create two clusters

Step-2: select random centers. In question it was mentioned A and C.

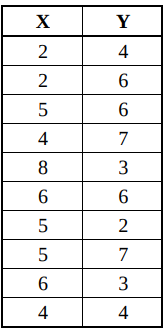
Step-3 : calculate distance between all centroids and all data points by applying D=|u1-v1|+|u2-v2|+…..+|Uq-vq| or D=SQRT((u1-v1)2 + (u2-v2)2) . So in our case , A and C are two centers. So, distance will be calculated B to A, B to c, D to A, D to C, E to A , E to C, A to A , A to C, C to C and C to A.

Step-4: all data points will be assigned to its nearest cluster (a centroid is considered as nearest cluster based on minimum distance which is calculated between each specific data point and all centroids)

Step-5: recalculate the center for each cluster

Step-6: continue from step-3 to step-5 until cluster allocation for each data point is stabilized.

Practice Question:



I am selecting two random centroid (2, 4) (6, 6) from data point. So I will create 2 clusters

|  |  |  |  |
| --- | --- | --- | --- |
| Data point | Dist-1 | Dist-2 | Cluster |
| (2, 4) | (6,6) |
| (2,4) | 0 | 4.472 | 1 |
| (2,6) | 2 | 4 | 1 |
| (5,6) | 3.605 | 1 | 2 |
| (4,7) | 3.605 | 2.236 | 2 |
| (8,3) | 6.082 | 3.605 | 2 |
| (6,6) | 4.472 | 0 | 2 |
| (5,2) | 3.605 | 4.123 | 1 |
| (5,7) | 4.242 | 1.414 | 2 |
| (6,3) | 4.123 | 3 | 2 |
| (4,4) | 2 | 2.828 | 1 |

New centroid:

For Cluster-1 : (13/4, 16/4) = (3.25, 4),

For Cluster-2 : (34/6,32/6)=( 5.66, 5.33)

|  |  |  |  |
| --- | --- | --- | --- |
| Data point | Dist-1 | Dist-2 | Cluster |
| (3.25, 4) | (5.66, 5.33) |
| (2,4) | 1.25 | 3.894 | 1 |
| (2,6) | 2.358 | 3.720 | 1 |
| (5,6) | 2.657 | 0.940 | 2 |
| (4,7) | 3.092 | 2.354 | 2 |
| (8,3) | 4.854 | 3.302 | 2 |
| (6,6) | 3.400 | 0.751 | 2 |
| (5,2) | 2.657 | 3.394 | 1 |
| (5,7) | 3.473 | 1.795 | 2 |
| (6,3) | 2.926 | 2.354 | 2 |
| (4,4) | 0.75 | 2.127 | 1 |

Cluster-1: (2, 4), (2, 6), (5, 2), (4, 4)

Cluster-2: (5, 6), (4, 7), (8, 3), (6, 6), (5, 7), (6, 3)