Practice Problem 9

Ian Chen

March 21, 2024

Problem 1 — Given the above data points, perform a k-means clustering on this dataset using Euclidean distance as the distance function. Here, k is chosen to be 3. The initial centroids are randomly selected as (1,1), (1,2), (3,1). Show the steps of the algorithm until convergence. What are the final clusters and their final centroids?

Answer

Iteration	*Centroids*	*Associated Points*	*New Centroids*
0	*(1, 1)*	*(1, 1)*	*(1, 1)*
	(1, 2)	*(1, 2), (2, 5)*	*(1.5, 3.5)*
	(3, 1)	*(3, 1), (3, 2), (4, 1), (4, 4)*	*(3.5, 2)*
1	*(1, 1)*	*(1, 1), (1, 2)*	*(1, 1.5)*
	(1.5, 3.5)	*(2, 5)*	*(2, 5)*
	(3.5, 2)	*(3, 1), (3, 2), (4, 1), (4, 4)*	*(3, 2)*
2	*(1, 1.5)*	*(1, 1), (1, 2)*	*(1, 1.5)*
	(2, 5)	*(2, 5)*	*(2, 5)*
	(3, 2)	*(3, 1), (3, 2), (4, 1), (4, 4)*	*(3, 2)*

Final centroids and associated points:

Centroid	*Points*
(1, 1.5)	*(1, 1), (1, 2)*
(2, 5)	*(2, 5)*
(3, 2)	*(3, 1), (3, 2), (4, 1), (4, 4)*