

# Practice Problem 9

Ian Chen

March 24, 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**

<i>Iteration</i>	<i>Centroids</i>	<i>Associated Points</i>	<i>New Centroids</i>
<i>0</i>	$(1, 1)$ $(1, 2)$ $(3, 1)$	$(1, 1)$ $(1, 2), (2, 5)$ $(3, 1), (3, 2), (4, 1), (4, 4)$	$(1, 1)$ $(1.5, 3.5)$ $(3.5, 2)$
<i>1</i>	$(1, 1)$ $(1.5, 3.5)$ $(3.5, 2)$	$(1, 1), (1, 2)$ $(2, 5)$ $(3, 1), (3, 2), (4, 1), (4, 4)$	$(1, 1.5)$ $(2, 5)$ $(3.5, 2)$
<i>2</i>	$(1, 1.5)$ $(2, 5)$ $(3.5, 2)$	$(1, 1), (1, 2)$ $(2, 5)$ $(3, 1), (3, 2), (4, 1), (4, 4)$	$(1, 1.5)$ $(2, 5)$ $(3.5, 2)$

*Final centroids and associated points after convergence:*

<i>Centroid</i>	<i>Points</i>
$(1, 1.5)$	$(1, 1), (1, 2)$
$(2, 5)$	$(2, 5)$
$(3.5, 2)$	$(3, 1), (3, 2), (4, 1), (4, 4)$