ISOM3360 Assignment: K-means Clustering

Zhang Yichen

May 7, 2020

1 Answer

We suppose that the clusters with initial center on A_1, A_4, and A_7 are Cluster 1, Cluster 2 and Cluster 3 respectively.

- A1(2,10): $d(A_1, A_1)$ is the smallest, classified to Cluster 1

$$d(A_1, A_1) = \sqrt{(2-2)^2 + (10-10)^2} = 0$$

$$d(A_1, A_4) = \sqrt{(2-5)^2 + (10-8)^2} = \sqrt{13}$$

$$d(A_1, A_7) = \sqrt{(2-1)^2 + (10-2)^2} = \sqrt{65}$$

- A2(2,5): $d(A_2, A_7)$ is the smallest, classified to Cluster 3

$$d(A_2, A_1) = \sqrt{(2-2)^2 + (5-10)^2} = 5$$

$$d(A_2, A_4) = \sqrt{(2-5)^2 + (5-8)^2} = 3\sqrt{2}$$

$$d(A_2, A_7) = \sqrt{(2-1)^2 + (5-2)^2} = \sqrt{10}$$

- A3(8,4): $d(A_3, A_4)$ is the smallest, classified to Cluster 2

$$d(A_3, A_1) = \sqrt{(8-2)^2 + (4-10)^2} = 6\sqrt{2}$$

$$d(A_3, A_4) = \sqrt{(8-5)^2 + (4-8)^2} = 5$$

$$d(A_3, A_7) = \sqrt{(8-1)^2 + (4-2)^2} = \sqrt{53}$$

- A4(5,8): $d(A_4, A_4)$ is the smallest, classified to Cluster 2

$$d(A_4, A_1) = \sqrt{(5-2)^2 + (8-10)^2} = \sqrt{13}$$

$$d(A_4, A_4) = \sqrt{(5-5)^2 + (8-8)^2} = 0$$

$$d(A_4, A_7) = \sqrt{(5-1)^2 + (8-2)^2} = 2\sqrt{13}$$

- A5(7,5): $d(A_5, A_4)$ is the smallest, classified to Cluster 2

$$d(A_5, A_1) = \sqrt{(7-2)^2 + (5-10)^2} = 5\sqrt{2}$$
$$d(A_5, A_4) = \sqrt{(7-5)^2 + (5-8)^2} = \sqrt{13}$$
$$d(A_5, A_7) = \sqrt{(7-1)^2 + (5-2)^2} = 3\sqrt{5}$$

- A6(6,4): $d(A_6, A_4)$ is the smallest, classified to Cluster 2

$$d(A_6, A_1) = \sqrt{(6-2)^2 + (4-10)^2} = 2\sqrt{13}$$
$$d(A_6, A_4) = \sqrt{(6-5)^2 + (4-8)^2} = \sqrt{17}$$
$$d(A_6, A_7) = \sqrt{(6-1)^2 + (4-2)^2} = \sqrt{29}$$

- A7(1,2): $d(A_7, A_7)$ is the smallest, classified to Cluster 3

$$d(A_6, A_1) = \sqrt{(1-2)^2 + (2-10)^2} = \sqrt{65}$$
$$d(A_6, A_4) = \sqrt{(1-5)^2 + (2-8)^2} = 2\sqrt{13}$$
$$d(A_6, A_7) = \sqrt{(1-1)^2 + (2-2)^2} = \sqrt{0}$$

- A8(4,9): $d(A_8, A_4)$ is the smallest, classified to Cluster 2

$$d(A_8, A_1) = \sqrt{(4-2)^2 + (9-10)^2} = \sqrt{5}$$
$$d(A_8, A_4) = \sqrt{(4-5)^2 + (9-8)^2} = \sqrt{2}$$
$$d(A_8, A_7) = \sqrt{(4-1)^2 + (9-2)^2} = \sqrt{58}$$

After first iteration, there are three clusters with:

- Cluster 1: A_1
- Cluster 2: A_3,A_4,A_5,A_6,A_8
- Cluster 3: A_2,A_7

New centroids for next iteration is:

- From Cluster 1: $(2, 10)$
- From Cluster 2: $\frac{(8,4)+(5,8)+(7,5)+(6,4)+(4,9)}{5} = (6, 6)$
- From Cluster 3: $\frac{(2,5)+(1,2)}{2} = (1.5, 3.5)$