

Lab-9

Build K-means algorithm to cluster a set of data
Stored in a CSV file.

To write

initial Data set

$$C_1 = (1, 0, 1.0)$$

$$C_2 = (5, 0, 2.0)$$

$$R_1 = (1, 0, 1.0)$$

$$R_2 = (1, 5, 2.0)$$

$$R_3 = (3, 0, 1.0)$$

$$R_4 = (5, 0, 2.0)$$

$$R_5 = (5, 5, 2.0)$$

$$R_6 = (4, 5, 2.0)$$

$$R_7 = (3, 5, 1.5)$$

cluster after iteration 1

$$C_1 = R_1, R_2, R_3$$

$$C_2 = R_4, R_5, R_6, R_7$$

new centroid

$$C_1 = (1.83, 2.33)$$

$$C_2 = (4.625, 5.375)$$

iteration 2

Final cluster after iteration 2

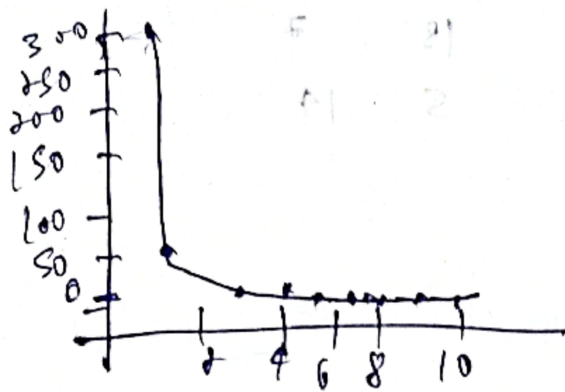
$$C_1 = R_1, R_2, R_3$$

$$C_2 = R_4, R_5, R_6, R_7$$

Tour 4

1. For (4, 10, 10, 10) data set

Draw the elbow plot that will show the optimal K value obtained



Optimal K value obtained was $K = 3.4$

1.2	2.0	2.4	1.8
2	0	4	1.4
2.2	2.5	2.1	2.6
			1.1

through lagging back the lagging (S, B)

$$\begin{bmatrix} 0.522.0 \\ 2.022.0 \end{bmatrix} = 10 \text{ show up to } 10$$

$$(0.522.0 - 1) \cdot (1.1 - 1) + 0.522.0 \cdot (1.1 - 1) = 1.1$$

what is it?

- 0.522.0 = 1.1
- 0.522.0 = 1.1
- 0.522.0 = 1.1
- 0.522.0 = 1.1