Quiz No. 6

Artificial Intelligence in the Built Environment (CE-451)

Student ID:

Question

Given the dataset in table below.

Points	x_1	x_2
1	1	2
2	1	4
3	1	0
4	10	2
5	10	4
6	10	0

Selecting Points 3 and 5 ($k = 2$) as the initial centroids, make data clusters.

Formulas

1. **Euclidean distance** between two points $(p = (x_1, y_1))$ and $(q = (x_2, y_2))$:

$$d(p, q) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.$$

When comparing distances you can compare squared distances (no need to take square roots).

2. **Mean (average)** — centroid for a cluster of (n) points $\left((x_i, y_i)\right)$ is

$$\bar{x} = \frac{1}{n} \sum_i x_i, \quad \bar{y} = \frac{1}{n} \sum_i y_i.$$

3. **Objective (SSE)**: sum of squared Euclidean distances of points to their centroids:

$$SSE = \sum_{\text{clusters}} \sum_{p \in \text{cluster}} |p - \text{centroid}|^2.$$