

Algorithm K-MEANS

Input Data $\mathcal{D} = \{x_i\}_{i=1:n}$, number clusters K

Initialize *centers* $\mu_1, \mu_2, \dots, \mu_K \in \mathbb{R}^d$

Iterate until convergence

1. for $i = 1 : n$ (recalculate clustering)

$$k(i) = \underset{k}{\operatorname{argmin}} ||x_i - \mu_k||$$

2. for $k = 1 : K$ (recalculate centers)

$$\mu_k = \frac{1}{|C_k|} \sum_{i \in C_k} x_i$$