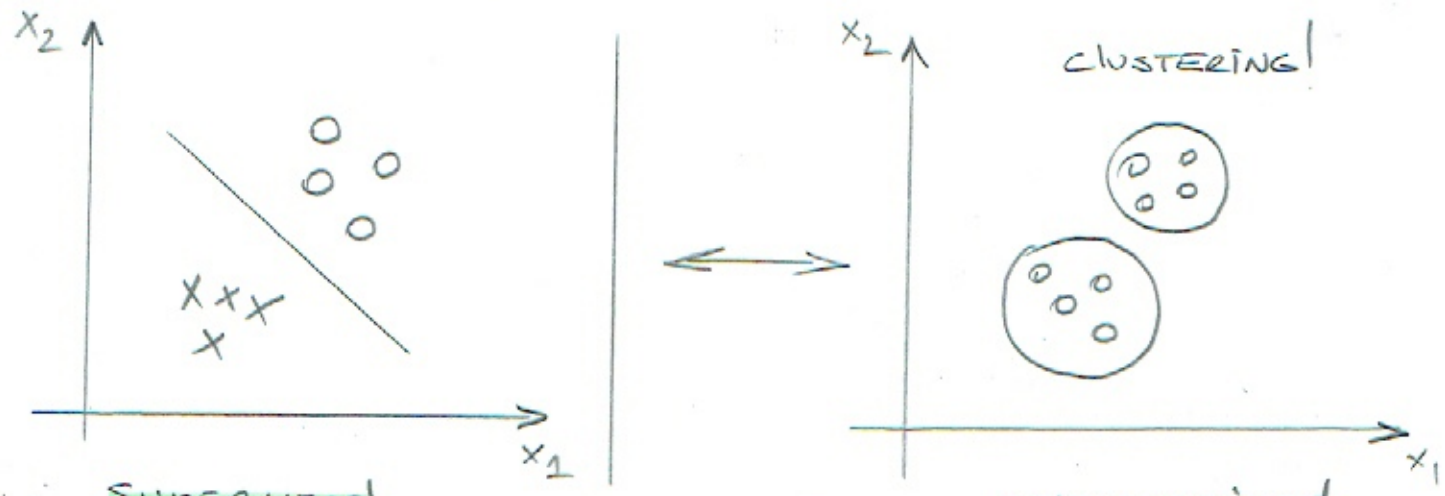
CLUSTERING k-MEANS



SUPERVISED LEARNING

$D = \{(x^1, y^1), (x^2, y^2) \dots (x^m, y^m)\}$

UNSUPERVISED LEARNING.

$D = \{x^1, x^2, \dots, x^M\}$

CLUSTERING

→ objects IN a CLUSTER should have similar charact^s but charact^s should be diffent from other CLUSTER

→ FIND structure IN unlabeled data

Apps: segmentation of CUSTOMERS OR T-shirt sizing

CLUSTERING
- exclusive (1 point, 1 CLUSTER) → K-MEANS
- overlapping (1 point → none cluster) → FUZZY / C-MEANS
- HIERARCHICAL
  ↳ 2 CLUSTER have A parent-child relationship

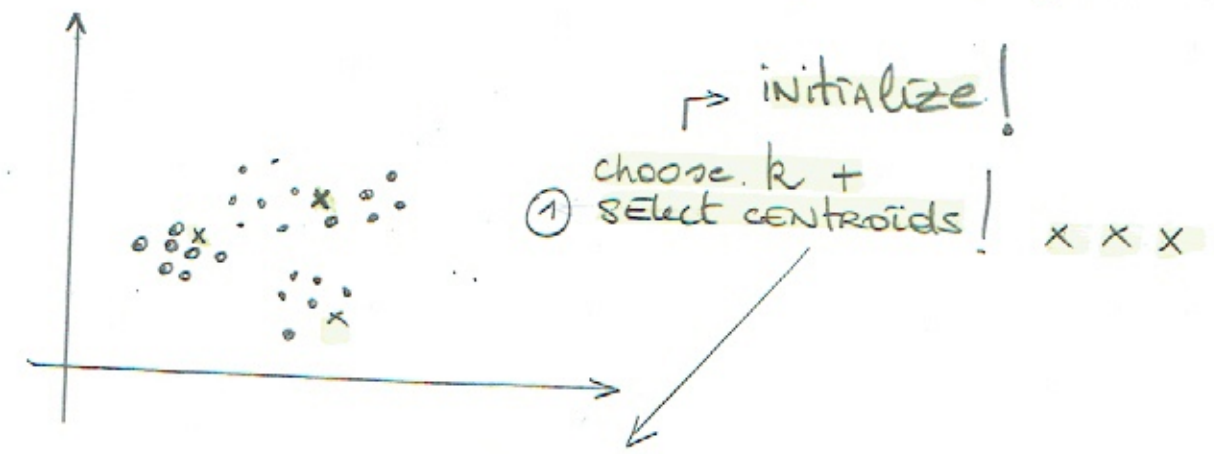
K-MEANS → k = hyperparameter !

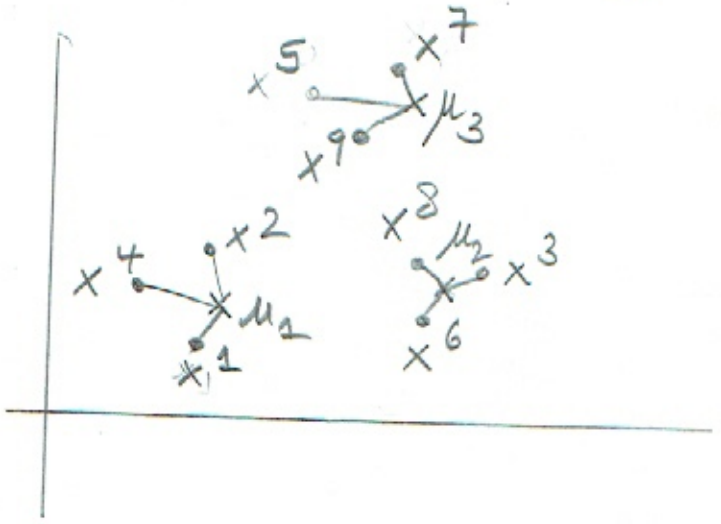
↓

CLUSTERS

$$D = \{x^1, x^2, \dots, x^m\}$$

example : Google News → clustering of groups of news (per topic)





x : centroids μ_1, μ_2, μ_3

$$c^{(i)} = \min_k \| x^i - \mu_k \|^2$$

② identify the centroid closest to $x^{(i)}$

→ Assign to centroid μ_i → $x^{(i)}$

③ calculate average of points per cluster

new $\mu_1 = \frac{1}{3}\left[x^1 + x^2 + x^4\right]$

$\mu_2 = \dots$

$\mu_3 = \dots$

→ move centroids!

↓

Repeat until centroids dont move anymore

optimization function

$c^{(i)}$: index of cluster to which x^i is assigned

μ_k : cluster centroid k

$\mu_{c^{(i)}}$ cluster centroid to which x^i is assigned

$$C = \min_k \| x^i - \mu_{c^{(i)}} \|^2$$

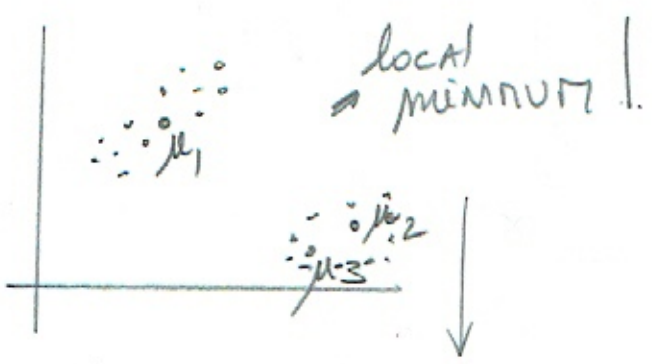
$(c^{(i)}, \mu_k)$

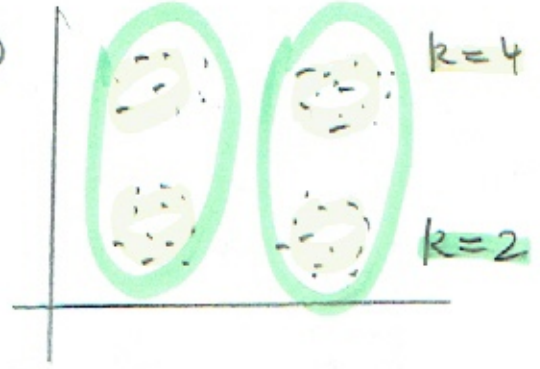
↳ select $c^{(i)}, \mu$ as to minimize C

WCSS = within cluster sum of squares

initialise

K ? → $k < M$

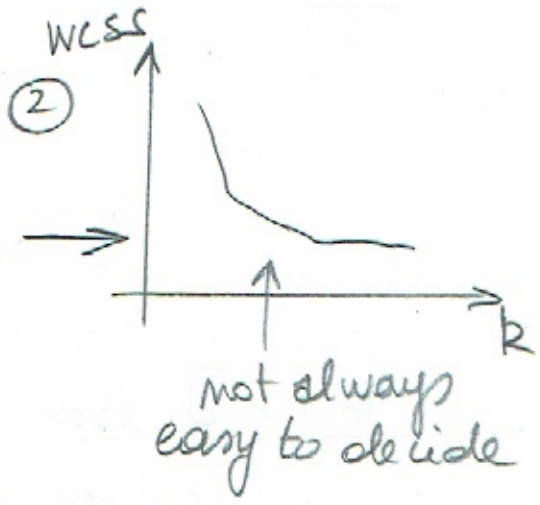




How to counter

① take samples as μ

② Do different init runs and take the one with lowest wcss



not always easy to decide

③ intuition a functional knowledge