Quiz

Q1 → True

Q2 → True

Q3 → True

ML (Clustering)

- AI → simulate the human
  - it can sense (take data)
    reason (learn from data)
    act (decision)
    adapt

  eg. Chat GPT, self-driving car

- ML ≠ AI
  - ML is part of AI (ML is brain) — learn from data

  * define a set of rules *
    - algorithm

now ML can be used in detect pattern?

Types of ML

- unsupervised (no training step) → group data
- supervised (has training step)
  - training data w/known target variable (discrete / continuous)

Q1 → False ? (can be classification or clustering)

Q2 → True

Q3 → True

Unsupervised Learning

- identify groups of elements
- suitable when we have unlabelled data

K-means

- minimize the distance bet. centroid (mean of the cluster) and data

types of distance

- Euclidean
- Manhattan
- Mahalobnis

* Initialization
  - randomly select the centroid of clusters

choosing optimum value of k

- Silhouette method
- Elbow method
- EDA ?

Optimize the performance

- reduce data dimension (ex. PCA)