Day 5 In-Class Exercises: Unsupervised and Reinforcement

Student guide. Run each lab step in order and predict before you run. Formative, not graded.

What to run in the labs

Morning: K-Means and PCA notebook

- Run the toy K-Means cell; watch the centroids settle.
- PREDICT the elbow: at which k does inertia stop dropping sharply? Then plot it.
- WRONG WAY: run K-Means with no fixed random seed; note how clusters change every run. Then fix it.
- WRONG WAY: force k = 2 without checking the elbow. Then choose k properly.
- Name each cluster; identify the high-risk bucket.
- PREDICT how many components explain more than 80% of the variance; then plot explained variance.
- Read the loadings: which component captures volatility?

Afternoon: reinforcement-learning bandit lab

- Run the toy cell that pulls each of the four assets a few times; see the noise.
- PREDICT: can a few noisy pulls reveal the best asset?
- Write the explore-or-exploit choice yourself, then run the agent.
- WRONG WAY: set the exploration rate to 0 (pure greed); see it lock onto a worse asset. Then add a little exploration.
- Tweak the exploration-rate sweep; find the value that earns the most.
- Q-learning gridworld: find the value-update (Bellman) line and explain each piece in words.
- PREDICT: which action has the higher value in every cell, left or right?

Discussion exercises

EX-D5-1. Which distance, straight-line (Euclidean) or angle (cosine)? Choose one for each, with a reason.

- (a) two companies by absolute size (revenue and total assets in euros), to group similarly sized firms.
- (b) two earnings snippets by topic, regardless of length.
- (c) two customers by total euros spent.

EX-D5-2. Would you trade this bandit tomorrow? Give three concrete reasons NOT to wire the simulated bandit to a real account.

EX-D5-3. Name the paradigm (supervised, unsupervised, or reinforcement) and justify in a few words:

- (a) predict default from labeled past loans.
- (b) group 500 stocks with no categories.
- (c) learn a trade-execution policy from reward feedback.
- (d) find the three hidden factors driving a portfolio.

EX-D5-4. Read the silhouette. Clustering A scores 0.62; clustering B scores 0.28.

- (a) Which is cleaner, and what does the score measure?
- (b) Could a high silhouette still be the wrong clustering for your purpose?