

Mathematics behind Machine Learning

Week 3 Exercise Sheet: Unsupervised Learning and Reinforcement Learning

Brief recap

- **Unsupervised learning** works with data that has **no labels**.
- The goal is to discover structure, patterns, or groups in the data.
- **Clustering** groups similar data points together.
- A common clustering idea is **k-means**, which tries to place points near a cluster centre.
- **Reinforcement learning** is about choosing actions to maximise total reward.
- Sometimes the actions don't do exactly what we thought they would; this is randomness.
- In reinforcement learning, the main parts are:

agent, environment, state (s), action (a), reward (r)

- A **policy** describes how the agent behaves:

$$\pi(s) \rightarrow a$$

- We try to find the policy that maximises **expected** total reward, because the actual reward will vary depending on the scenario.
- An expectation can be calculated as the sum of all of the different (value) times (probability getting that value) pairs.

Exercises

K-means idea

A clustering method tries to minimise the distance between data points and their cluster centres. One simplified objective is

$$\sum_{i=1}^n (x_i - c_i)^2$$

where c_i is the centre of the cluster assigned to x_i .

In k-means, the cluster centres are often written as μ_1, μ_2, \dots as well.

We have the following data points on a number line:

2, 4, 5, 6, 15, 17, 19, 20

We want to split them into **2 clusters** ($K = 2$) using the K-means algorithm.

Start with these initial cluster centres (called *centroids*):

- Cluster A centroid = 4
- Cluster B centroid = 17

Follow these steps:

Step 1: Assign points

For each number, decide whether it is closer to centroid A or centroid B.

For example:

The point 5 is 1 away from 4 and 12 away from 17, so it goes into Cluster A.

Complete the table:

Point	Distance to 4	Distance to 17	Cluster
2			
4			
5			
6			
15			
17			
19			
20			

Step 2: Update the centroids

For each cluster, calculate the **mean (average)** of all points assigned to it.

New centroid formula:

$$\text{new centroid} = \frac{\text{sum of points in cluster}}{\text{number of points}}$$

Write your new centroids:

- New Cluster A centroid = _____
- New Cluster B centroid = _____

Step 3: Repeat

Using your new centroids:

1. Reassign every point to the nearest centroid.
2. Compute new averages again.

Step 4: Stop when nothing changes

K-means stops when the cluster memberships no longer change.

Questions

1. What are the final two clusters?
2. What are the final centroid values?
3. How many rounds did the algorithm take before it stopped?

Reinforcement Learning

Imagine a robot that chooses where to go. Each move down costs -1, in the other directions it costs 1, leaving the maze costs 0. Fill in the values of the white cells.

What if the maze was very large? What would be the most efficient way to find the correct way out?

What if instead of looking for one exit, we were looking to find loot throughout the maze?

Answering this question is really interesting and it is related to the so-called Travelling Salesman Problem and Dijkstra's algorithm.

					exit
	start				