C1: Intro to RL

▼ Table of Content

What is RL

- Branch of ML where learning occurs through interaction with environment
- Goal-oriented learning where learner not thought what action to take, albeit, learn from the consequences of actions

High-Level RL Step

1. Agent interact with env by performing action
2. Agent perform action and move from one state to another
3. Agent receive reward based on the action performed
4. Agent understand the quality of an action based on the reward

How is RL different from other ML Paradigms

- Supervised Learning
  - RL does not have a human defined label or instruction to instruct the agent moving towards specific goal

- Unsupervised Learning
    - Unsupervised only have the input data and does not have any reward function/environment to interact with

Elements of RL

Agent

- Bot/User/Machine that interacts with the environment based on its policy

Policy Function

- Defines the agent's behaviour and action to take given the current state
- Often denoted by π

Value Function V

- Quality of the current state towards the agent (Long Term)
    - expected long-term return with discount γ
- Often denoted by $v(s)$

Reward R

- an immediate signal received in a given state (Short-Term)

> **Value function** , rather than immediate rewards, is what reinforcement learning seeks to predict and control.

Model

- Agent's representation of the environment
- Learning can be either model-based or model-free training:
    - Model-based agent exploits previously learned information to accomplish task
    - Model-free agent relies on trial-and-error experience for performing right action

Agent Environment Interface

Agent perform action A_t at time t from one state, S_t to another state S_{t+1}. And based on the action and state, a reward R is given.

Ultimately, RL is about finding the optimal actions that will increase the reward R.