

Reinforcement Learning

Instructor

Adam P Tashman

Summary

Reinforcement Learning is a dynamic, active area in machine learning. It is fundamentally different from supervised and unsupervised learning, as it allows an agent to learn by interacting with its environment. This allows for learning when the ground truth is unavailable, which is often the case. Even when ground truth is available, it often becomes outdated in a dynamic environment (think predicting Netflix usage before and during a pandemic). This course will introduce the elements of RL, k-armed bandits, Markov Decision Processes, value functions, Policy Gradients, Q-Learning and deep Q-Learning. The focus will be on computation and not theorem-proof (e.g., convergence guarantees).

Books

Reinforcement Learning 2nd edition. RS Sutton & AG Barto

Mastering Reinforcement Learning with Python. E Bilgin

Topic List

- Foundations of Reinforcement Learning
- k-armed Bandit Problems
- Markov Decision Processes (MDPs)
- Approaches to solving MDPs:
 - > Dynamic Programming
 - > Monte Carlo
 - > Temporal-Difference Learning
- Q-Learning
- Deep Q-Learning (DQNs)
- Policy Gradients (PGs)

Learning Outcomes

- Understand the key elements of Reinforcement Learning
- Simulate a k-armed bandit problem with epsilon-greedy actions
- Understand the properties of Markov Decision Processes

- Understand, compute and improve value functions to optimize policy
- Study the Bellman optimality equations for estimating the optimal value functions
- Apply and contrast different methods for estimating value functions: dynamic programming, Monte Carlo simulation, temporal difference methods
- Explain how Q-Learning works and how it learns off policy
- Use Q-Learning to compute value functions
- Understand how to improve policy using Policy Gradients
- Compare the value-based approach to the Policy Gradient approach