Reinforcement Learning

July 17-19, 2023

Day 1 (9:00am - 6:00pm)

- [9:00-9:30] **Welcome: Meet & Greet**

- [9:30-11:00] **Session 1 (1.5 hours): What is RL, why RL and basic RL**

    o Introduction to decision making

    o What is and isn't RL? How is RL Different from Supervised Learning.

    o The central challenge in RL: Exploration v/s Exploitation

- [11:00-11:30] break (0.5 hours)

- [11:30-12:45] **Session 2: Simple Sequential Decision Making**

    o The central challenge in RL: Exploration v/s Exploitation (continued)

    o Bandits

- [12:45-14:00] Lunch

- [14:00-15:30] **Session 3: Contextual Bandits and Policy Gradients**

    o Contextual Bandits

    o Basic Terminology: Markov decision process, what is an episode, etc.

- [15:30-16:00] break (0.5 hours)

- [16:00-17:00] **Session 4: How to use RL Algorithms? Walk through some applications**

    o Assumptions of MDPs

    o Recommendation systems

    o Balloon Localization

    o Manipulation

    o Urban Planning

    o Introduction to the Problem Clinic (15 min)

- [17:00-18:00] Reception (1 hour)

Day 2 (9:00a - 5pm)

- [9:00-10:30] **Session 1: Introduction to RL methods**

- Introduction to Policy Gradients

- REINFORCE

- **Value Based Reinforcement Learning**

    - Introduction to Policy Gradients: REINFORCE

    - Why value based RL?

    - Connection between dynamic programming and RL

    - Policy Iteration

    - Value Iteration

    - Q-Learning

    - What is off-policy learning?

    - Deep Q-Learning (DQN)

        - Target Network, Replay Buffer

    - Double Q-Learning (DDQN)

- **[10:30-10:45] Break**

- **[10:45-12:15] Session 2: Interactive session (Deep Q Networks)**

    - Setup an environment into a format amenable to RL algorithms

    - Hands-on Exercise on DQN, DDQN, interpreting outcomes

- **[12:15-12:30] Break**

- **[12:30-13:00] Session 3: Deep Q-Learning in Continuous Action Space**

    - Deep Deterministic Policy Gradients (DDPG)

    - Soft-Actor Critic (SAC)

    - Twin Delayed DDPG (TD3)

- **[13:00-14:00]** Lunch

- **[14:00-14:45] Group Session: Discuss and Formulate Problems into RL Framework**

- [**14:45-15:30] Session 4: Practical Perspectives**

    - The Reward Hacking Problem

    - What if my action space is large?

    - What is my RL problem is non-markov? How to design the state-space?

- **[15:30-16:00] Coffee Break**

- **[16:00-17:00] Group Session: Work on Problem Clinic**

Day 3 (9:00am - 5pm)

- **[9:00-11:30] Session 1 (2:30 hours): Problem Clinic Presentation & Discussion Part I**

- **[11:30-12:00]** Break

- **[12:00-13:00] Session 2 (1 hour): Problem Clinic Presentation & Discussion Part II**

- **[13:00-14:00]** Lunch (1 hours)

- **[14:00-15:30] Session 3: Discussion on Practice and Theory of RL**

    o Miscellaneous Assumptions and Roadmap for RL

        ▪ The reset requirement in RL

        ▪ Tradeoffs in using different RL Algorithms

        ▪ Evolutionary Methods and its relationship to reinforce

    o Theoretical concepts

        ▪ How to evaluate RL: asymptotic convergence, sample complexity, regret

- **[15:30-16:00] Coffee Break**

- **[16:00-17:00] Session 4: AMA and Discussion on Advanced Topics**

    o Discussion on limitation of RL techniques: Non-Stationarity, data inefficiency

    o Overview of Advanced Topics

        ▪ MCTS is one of them

        ▪ Based on class interest, we will delve into one of the advanced topics.

    o More Applications