Yash's plan for RL study

Week 1:

1. Dynamic Programming for RL: Policy Iteration, Value Iteration.
2. Monte Carlo methods.
3. Temporal Difference (TD) Learning: TD(0), TD(λ).

Hands on Experience: Implement Policy Iteration and Value Iteration, and use Monte Carlo methods to estimate state-value functions. Apply TD(0) for value updates in CartPole environment.

Week 2:

4. Value-based methods: Q-learning and SARSA.
5. Policy gradient methods: Advantages over value-based methods.
6. Introduction to Function Approximation (Linear, DNN-based).

Hands on Experience: Using the same CartPole from last week, Write a tabular Q-learning agent for balancing the pole. Modify it to use SARSA and compare results. Explore policy gradients using a linear function approximator for policy learning.

Week 3:

7. Basics of Deep Q-Networks (DQN): Experience replay, target networks.
8. Look at OpenAI Gym.

Hands on Experience: Train a Deep Q-Network (DQN) with experience replay and target networks. Evaluate the model on the same CartPole and plot performance.

Week 4:

9. Proximal Policy Optimization (PPO): Clipped objective, advantages.
10. Actor-Critic Methods: Combining policy and value functions.

Hands on Experience: Implement PPO for continuous action control on LunarLander environment. Use an Actor-Critic method to combine policy and value functions.

Week 5-6:

11. Trust Region Policy Optimization (TRPO): Concept and differences from PPO.
12. Advantage Actor-Critic (A2C) and Asynchronous Advantage Actor-Critic (A3C).
13. Soft Actor-Critic (SAC): Handling continuous action spaces.

Hands on Experience: On the same LunarLander environment, Implement TRPO and compare its stability and performance against PPO. Train a Soft Actor-Critic (SAC) agent for more robust control in continuous spaces.

Week 7-8:

14. Cooperative and competitive Multi agent RL environments.
15. Independent Q-learning in multi-agent setups.
16. Multi-agent PPO.

Some useful links:

You can find different RL environment in OpenAI's Gym library: https://gymnasium.farama.org/

CartPole: https://gymnasium.farama.org/environments/classic_control/cart_pole/

LunarLander: https://gymnasium.farama.org/environments/box2d/lunar_lander/

Multi Agent Unity : https://docs.unity3d.com/Packages/com.unity.ml-agents@3.0/manual/index.html

Another Cool Course by Google DeepMind: https://youtu.be/2pWv7GOvuf0?si=AAfl62qb6zFGNvMP