DS-GA 3001.005
NYU Center for Data Science

Reinforcement Learning

Homework 03

Exercise 1 (30 points)

Function Approximation in Reinforcement Learning:

- **1.1 (5 points)**: What is the difference between tabular Q-learning and Q-learning with function approximation?

- **1.2 (5 points)**: Why are TD updates by Gradient Descent called *Semi-Gradient* updates?

- **1.3 (10 points)**: What challenges may tabular Q-learning meet when learning to play the game of Chess? Provide at least 2 answers.

- **1.4 (10 points)**: What challenges may value function approximation meet in a Reinforcement Learning context compared to offline regression in a Supervised Learning context? Provide at least 3 answers.

Exercise 2 (20 points)

Deep Q-Network (DQN):

- **2.1 (10 points)**: What is Experience Replay? How does it benefit DQN?

- **2.2 (10 points)**: Provide a complete pseudo-code for the DQN algorithm.

Exercise 3 (50 points)

For 3.1 and 3.2, please produce the exact code (not pseudo-code):

- **3.1 (15 points)**: Write a custom function that implements Experience Replay updates for an arbitrary minibatch size (use the TD(0) target as in the original DQN algorithm). The required input to this function must be listed as arguments and/or defined in the docstring.

- **3.2 (20 points)**: Implement both, the Gym environment and an agent to learn to control a cart pole. The agent should be the PPO algorithm from the *stable-baselines* Python library. It should start from scratch, train for 50,000 steps, and be evaluated on 50 episodes.

- **3.3 (15 points)**: List at least 3 examples of RL agent components that you can import ready-to-use from the *keras-rl* Python library. For each component, provide both the exact method name (list of arguments is not necessary) and a description of what it is.