FRA 503: Deep Reinforcement Learning

Homework 1

Part 1: Setting up Multi-armed Bandit.

For this first part, you will implement a multi-armed bandit framework from scratch as explained in the class. The component should include, but is not limited to,

1. A bandit class.

   This class should include:

   a. A constructor which initializes n bandits, each with its own hidden reward distribution.
   b. A function which returns a reward signal
2. An agent class.

   This class should include:

   a. A constructor which initializes an agent with learnable parameters, and steps for each action.
   b. An update function that updates the agent's learnable parameters and steps.
3. A simulation script for running experiments.

Part 2: Implementing epsilon-greedy algorithm.

You must implement an epsilon-greedy algorithm in the simulation script, analyze the result, and plot a graph on timesteps vs reward for each bandit.

Part 3: Implementing UCB

You must implement a UCB algorithm in the simulation script, analyze the result, and plot a graph on timesteps vs reward for each bandit.