COMP47590

Advanced Machine Learning

Lab Task 5: Multi-armed Bandits

Introduction

Multi-armed bandit problems display most of the characteristics of a reinforcement learning problem. In this task you well implement and compare policies for multi-armed bandit problem.

Tasks

1. Download the k-armed bandit demo python notebook.
2. Design an experiment to compare the performance of a **random** arm selection policy, a **greedy** arm selection policy, and an ***ε*-greedy** arm selection policy
    - Consider 4-armed bandit problems.
    - Assume rewards come from fixed Guassian distributions.
3. Perform this experiment.
4. An ***ε*-first** policy uses a random policy for the first N trials and then a greedy policy after that. Implement this policy.
5. An ***ε*-decreasing** policy basically uses an ***ε*-greedy** strategy except the value of ε decreases over time. Implement this policy.
6. Compare the two new policies to those implemented already.
7. There are many scenarios in which the distribution of rewards changes over time. Change the experimental scenario so that the distribution rewards for each arm changes every 100 pulls.
8. Re-run your experiment to test whether the ranking of policies changes when this is the case.