

# CAP 6629: Reinforcement Learning

Spring 2022

Course project 1

**Due: 02/16/2022 (Wednesday), 11:59PM**

Submission: A single PDF with your code (use any programming language), results and analysis.

**Your report should follow the template I showed in class.**

**Part 1:** Read chapter 2 and use *any programming language* to implement a multi-arm Bandit problem. You may follow the algorithm pseudo code (page 8 of lecture note). The reward distributions are provided on page 10 and you need to estimate the mean value of each action yourself. Please show 1) your average reward curves of different  $\epsilon$  values; and 2) percent of optimal actions (similar figures as we studied in the class).

*Hint: to get the percent of optimal actions, you need to get the optimal action from your simulation first. Then, you can either re-run the code to calculate the percent of optimal action, or use the previous recorded trial data.*

*Please use 2000 or 3000 for your iteration (instead of 1000 in the lecture note). I suggest to compare the trajectories of  $\epsilon = 0.1$  and  $\epsilon = 0.01$  over time.*

**Part 2:** Apply the algorithm in part 1 to a dataset below. The full reward distributions are provided here: <https://drive.google.com/file/d/1whkIlnL4FKeHg2IfdcbT1j18L26fg9aF/view>

Suppose an advertising company is running 10 different ads targeted towards a similar set of population on a webpage. Each column index represents a different ad. We have a 1 if the ad was clicked by a user, and 0 if it was not. A sample from the original dataset is shown below:

Ad 1	Ad 2	Ad 3	Ad 4	Ad 5	Ad 6	Ad 7	Ad 8	Ad 9	Ad 10
1	0	0	0	1	0	0	0	1	0
0	0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0
1	1	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0
0	0	0	0	0	0	0	1	0	0

Please provide the maximum reward you can achieve with this dataset, and what is the best ad from the algorithm.