

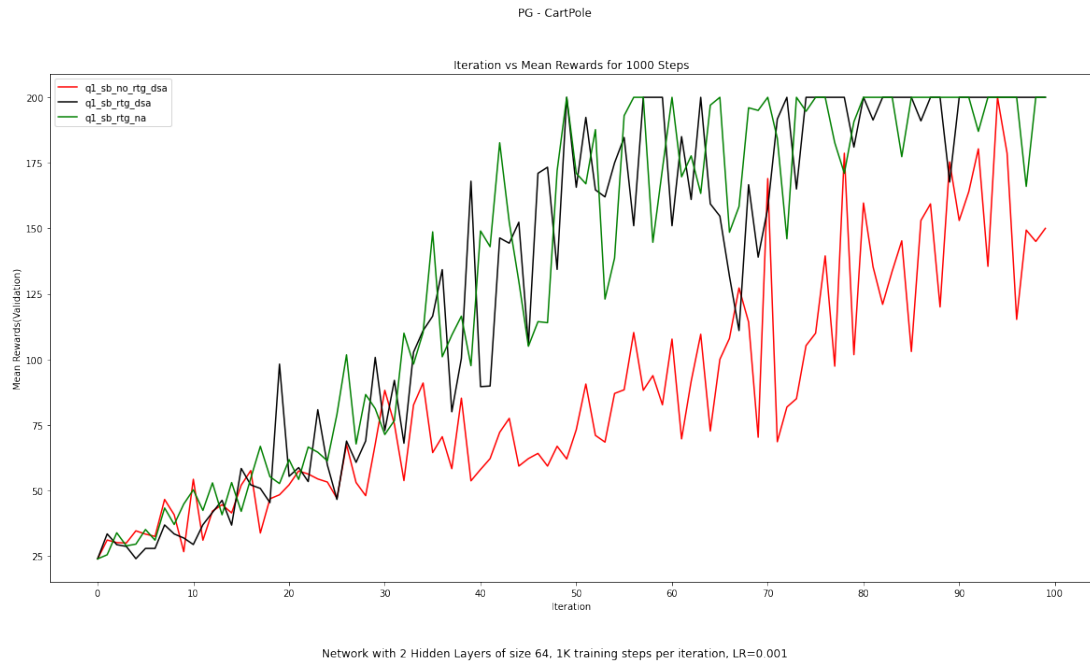
Assignment 2: Policy Gradients

huseyinabanox@gmail.com

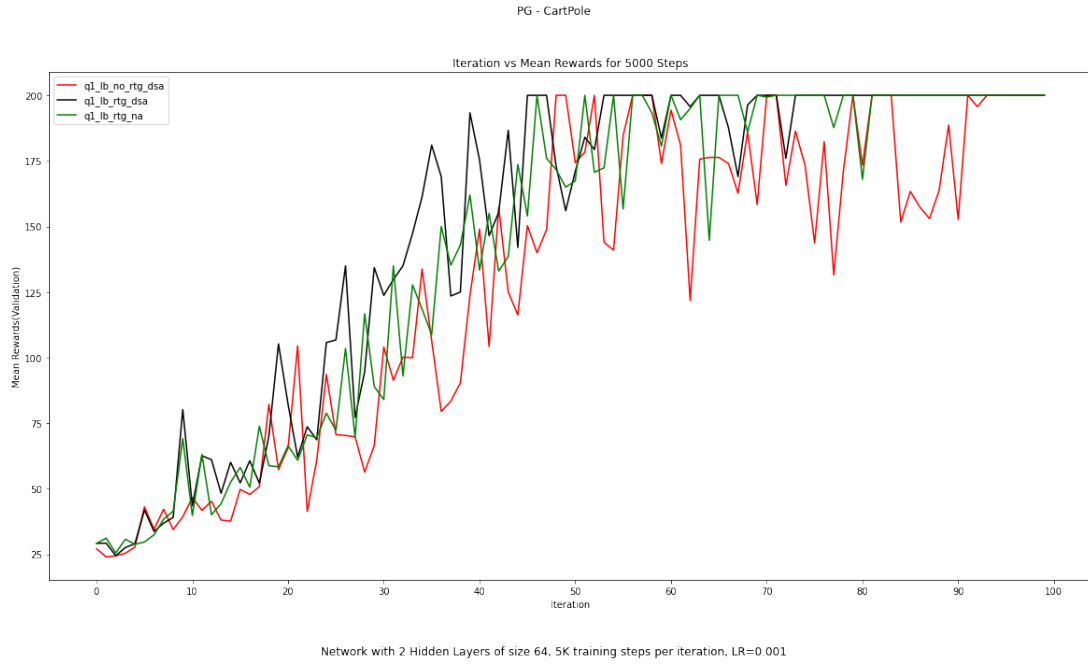
January 2023

1 Small-Scale Experiments

1.1 Experiment 1



Reward-to-go estimator performs better. Advantage standardization difference is not statistically significant.



Reward-to-go estimator performs better. Advantage standardization difference is not statistically significant. Increasing batch size helps the algorithm converge with fewer iterations.