Homework 2 Report

Reinforcement Learning

Name: Kai-Jie Lin
Student ID: 110652019
April 30, 2023

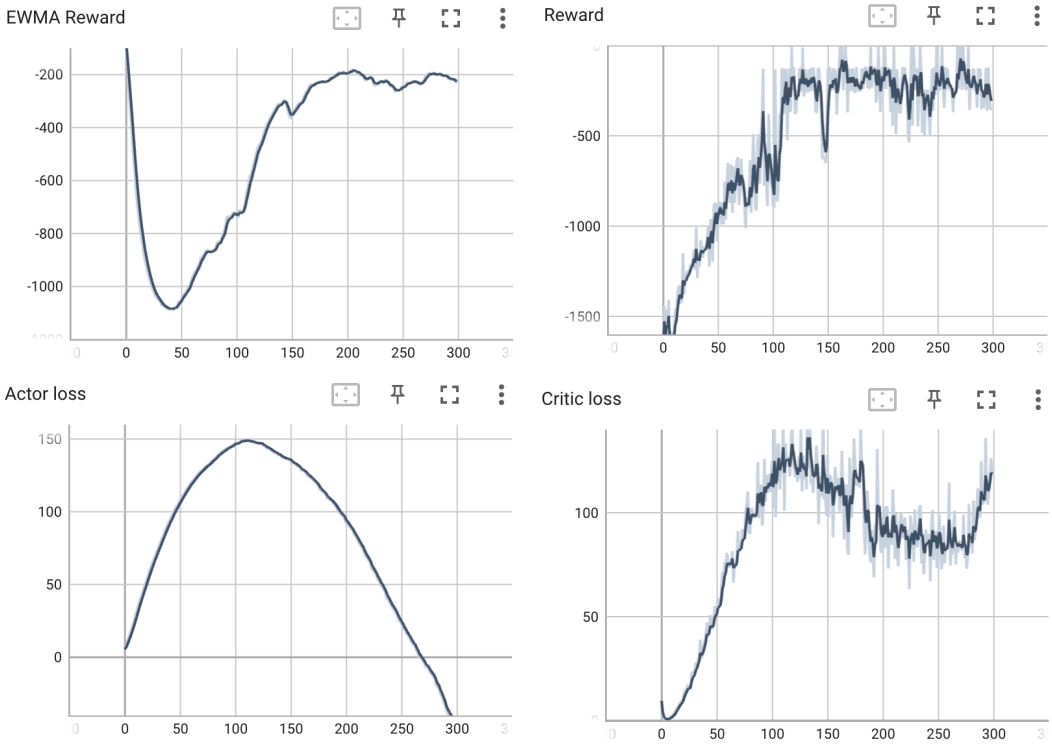
1 Experiment of DDPG

1.1 Pendulum-v1

Learning Rate(Actor)	0.0001
Learning Rate(Critic)	0.001
Batch Size	128
Hidden Size	128
Layer Number	1

Table 1: Hyperparameters

Result: Reach well policy in near 300 steps.



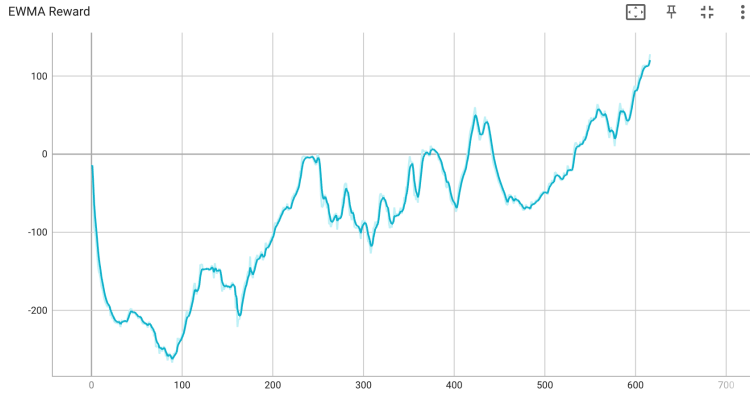
1.2 LunarLanderContinuous-v2

Use bayesian optimization to tune the hyper parameters. Reach EWMA reward=120 in 616 steps.

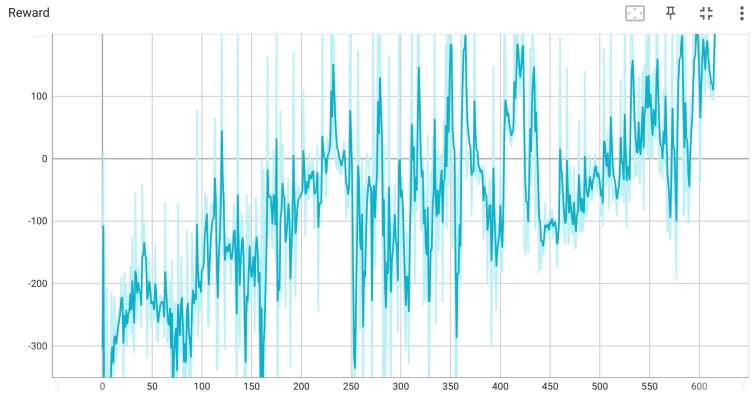
Learning Rate(Actor)	0.008
Learning Rate(Critic)	0.001
Learning Rate decay rate(Actor)	0.85
Learning Rate decay rate(Critic)	0.81
Batch Size	211
Hidden Size	128
Layer Number	2
Noise Scale	0.28

Table 2: Hyperparameters

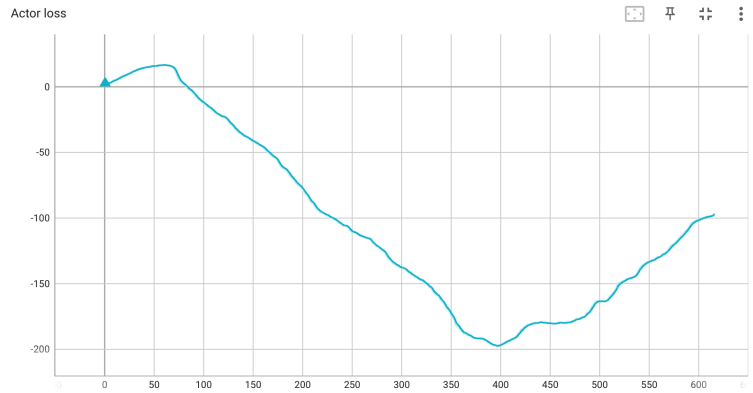
EWMA Reward:



Reward:



Loss of Actor:



Loss of Critic: