

Algorithms Comparison.

Minh Hoang Nguyen

minhhoangnguyen0911@gmail.com

July 2023

1 Testing Results

This section presents the results, which have been accumulated from testing the three trained agents over 250 testing seeds of CartPole environment. The results show that REINFORCE is the algorithm with the highest average, however, DQN is the algorithm with the lowest standard deviation.

Algorithms	Average (μ)	Standard Deviation (σ)
REINFORCE	558	341
Actor-Critic (A2C)	340	162
Deep Q Network (DQN)	173	75

Table 1: Comparing the results of 3 algorithms: REINFORCE, A2C and DQN.

2 Training Comparison

This section presents the training process of the three algorithms; each algorithm was trained over 200 episodes 5 times. The reason I have recorded rewards over episodes instead of steps like in other Reinforcement Learning environments was that the CartPole environment only has 2 types of reward (1 when the pole still stands and 0 when it falls), so recording the reward every step would result in redundancy. For the above-mentioned reason, a higher reward in an episode also indicates a better performance.

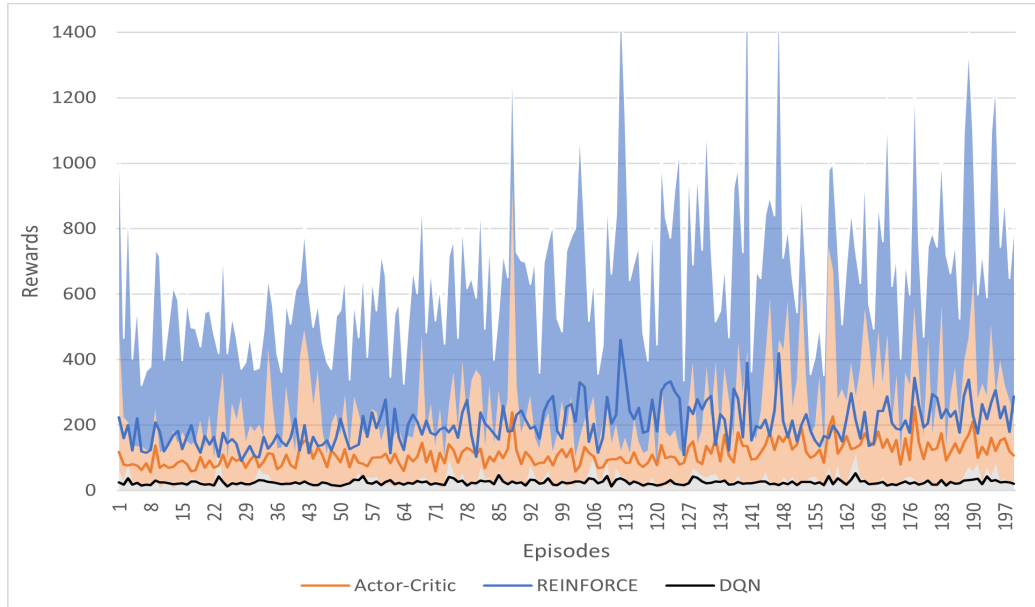


Figure 1: Comparison between the three algorithms.

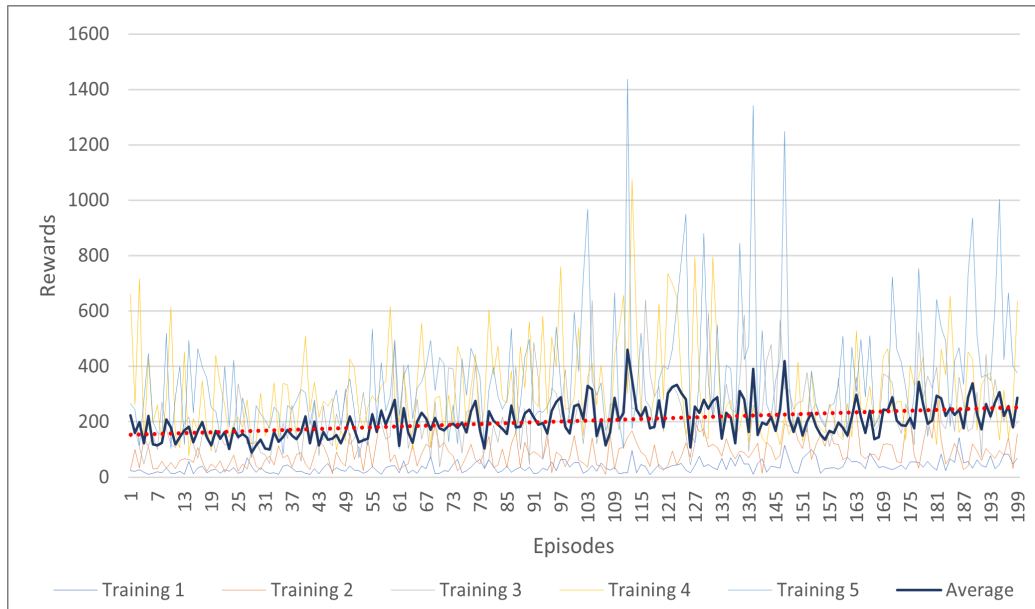


Figure 2: CartPole training using REINFORCE Algorithm. (Number of training episodes: 200)

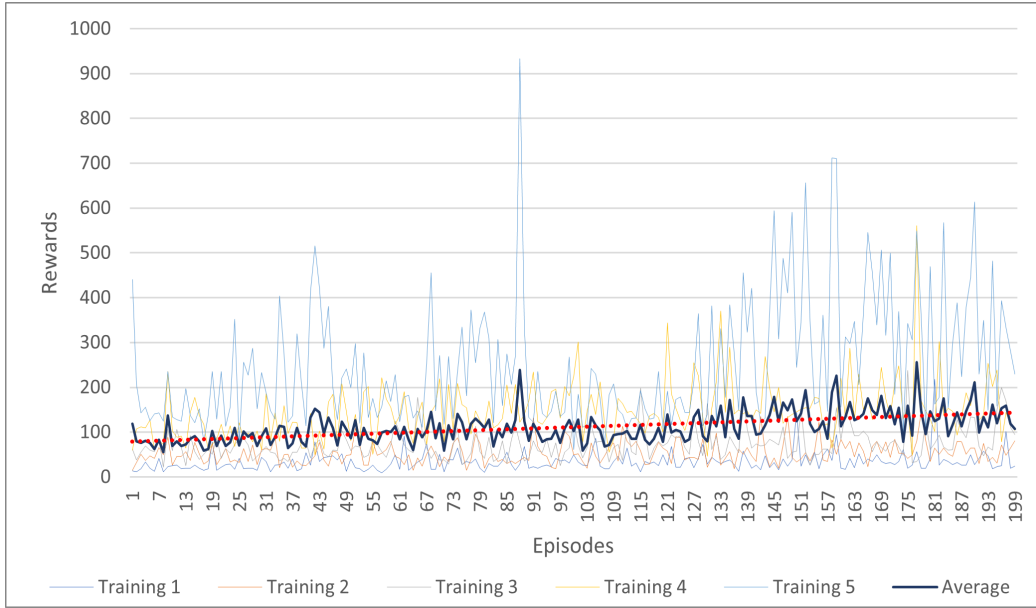


Figure 3: CartPole training using Actor-Critic Algorithm. (Number of training episodes: 200)

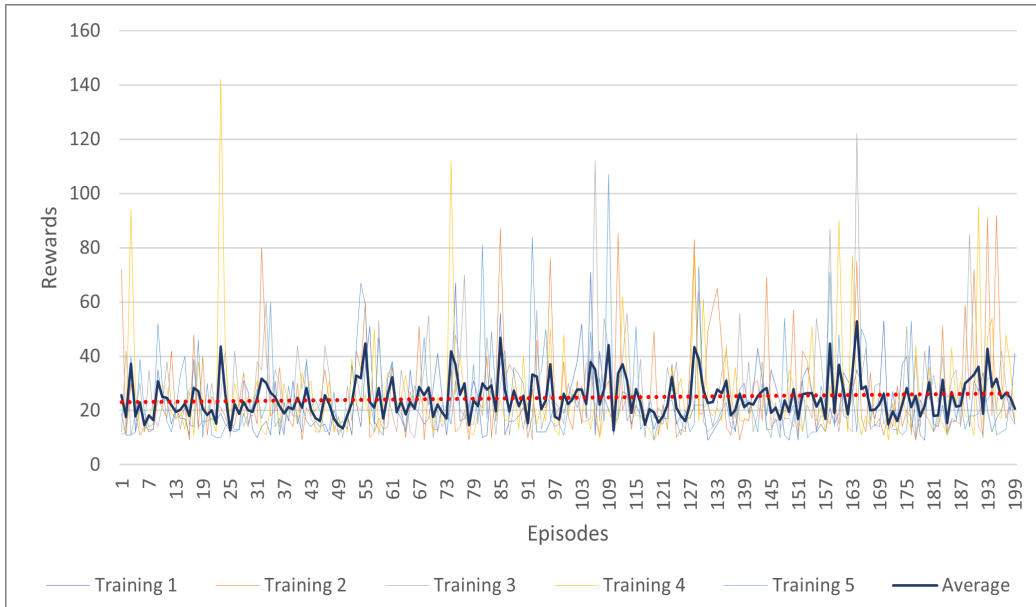


Figure 4: CartPole training using DQN Algorithm. (Number of training episodes: 200)