UDACITY DRLND

P1-BANANA-NAVIGATION

Description :

The challenge requires to train an agent to navigate and collect bananas in a large square World. A reward of +1 is provided for collecting a yellow banana, and a reward of -1 is provided for collecting a blue banana. The agent must get an average score of +13 over 100 consecutive episodes.

Learning Algorithm : DQN

The learning algorithm is Vanilla Deep Q Learning which is used to solve the Space Invader Gym environment in the previous session. The agent is unaware of the entire state and tries to get reward through function approximations of TD-learning. Pixels are passed as input and a neural network is employed to process the state. The architecture of the network is as follows:

Fully connected layer| Input: 37- State size, Output : 128

Fully connected layer| Input: 37 , Output : 64

Fully connected layer| Input:64, Output : 4 - Action size

Parameters:

Replay buffer size : 100000, Minibatch size:64, Learning Rate: 5e-4

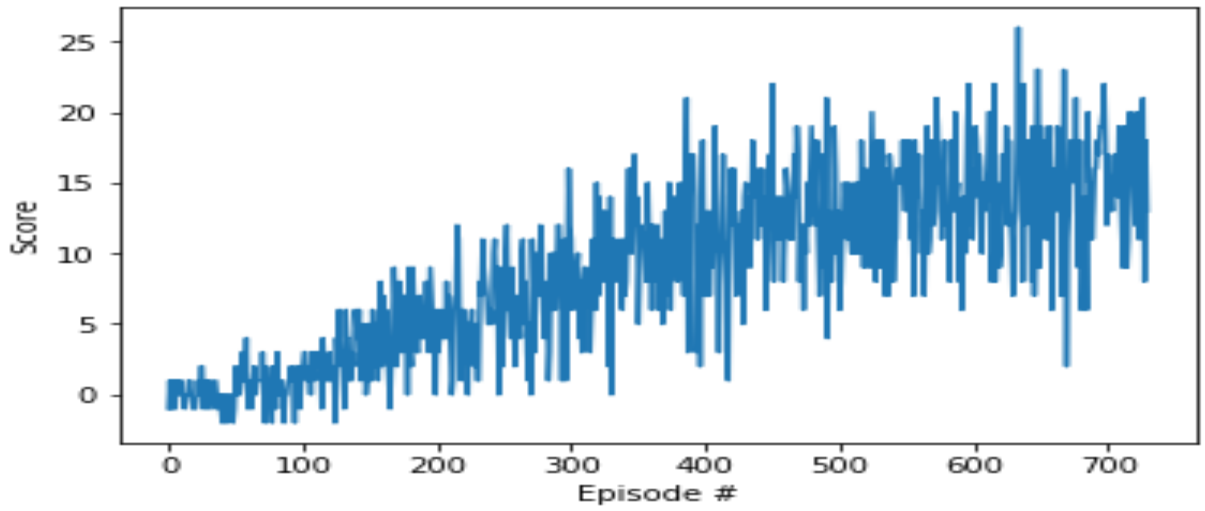
Discount factor: 0.99, Epsilon start:1, Epsilon end:0.01

The Algorithm utilizes Replay buffer and the fixed Q-targets with soft updates for fixed intervals.

---

*Please note that the deadline is Jan 15th

Agent and Environment: The environment is solved in 630 episodes.



Future Work:

In the DQN tutorial, other variants of DQN are discussed such as Prioritized experience replay, Double DQN, Rainbow. I will understand and try to implement few other variants such as C51, QRNs with regular and prioritized replay buffers with double learning by following the following github repository. :)

https://github.com/qfettes/DeepRL-Tutorials

---

*Please note that the deadline is Jan 15th