

Project 1 - Banana Collector

Quentin Viville
Udacity
Deep Reinforcement Learning Nanodegree

1 Goal of the project

The goal for this project is to train an agent to solve the Banana Collector environment.

2 Algorithm and neural network architecture

The algorithm used to train the agent is a Deep Q-Network (DQN). The neural network architecture retained is based on three feed-forward layers with ReLu activation functions. This is the original architecture given in the course for the Lunar Lander environment. After trying different architectures (more layers, less layers, other activation functions), it seems that this simple architecture gives good results. Besides, it shows that the DQN algorithm, with no modification to the neural network part, is able here to solve efficiently two very different environments.

3 Optimization parameters

The values for the parameters ε_{start} , ε_{decay} and ε_{end} are respectively 1.0, 0.995 and 0.01. They have been selected empirically among several different combinations, as they seem to offer a good compromise between speed and robustness of convergence.

4 Results

With the configuration detailed above, the agent is able to solve the environment in less than 500 episodes, which is quite efficient, as the recommended upper limit is 1800 episodes. In the figure 1, we can see the evolution of the average score until the threshold score of 13.0 is reached.

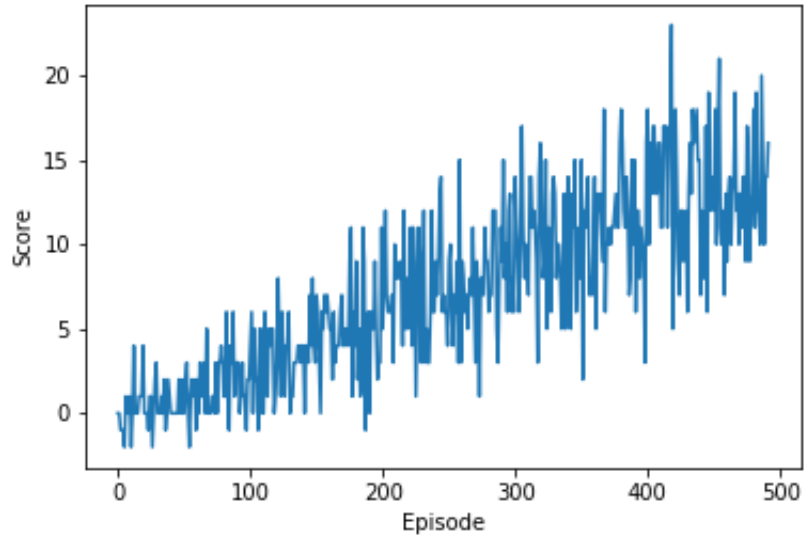


Figure 1: Plot of the score throughout learning

5 Possible improvements

In the next future, I will try the following extensions to the Deep Q-Network algorithm that have been presented in the course:

- Double DQN (DDQN)
- Prioritized experience replay
- Dueling DQN

as they potentially could provide huge performance improvements.