

# Deep Reinforcement Nano Degree Project 1: Navigation

Peng Zhao

December 18, 2019

## Abstract

This report briefly summarizes the deep reinforcement nano degree project 1, which is training the agent to collect yellow banana using deep reinforcement learning method. The problem is solved after 191 episodes in this experiment.

## 1 DQN Architecture

The Deep Q-Network consists of 3 layers with Relu activation. The input data size of the first layer is the state size, which is 37. The output size of the first layer is 64. The output size of the second layer is also set to 64 and the out put of the third layer is the action size 4.

## 2 Parameter

The choice of  $\varepsilon$ -greedy parameter  $\varepsilon$  can significant effect the speed of learning. In this experiment, I use the following equation to determine this parameter:

$$\varepsilon = \max(\varepsilon_0 \times d, \varepsilon_{min}) \quad (1)$$

where  $\varepsilon_0$  is the initial value of  $\varepsilon$ ,  $d$  is decay rate for each step in episode,  $\varepsilon_{min}$  is the minimum value that  $\varepsilon$  can use. In this experiment, we use  $\varepsilon_0 = 0.5$ ,  $d = 0.98$ ,  $\varepsilon_{min} = 0.01$ .

## 3 Result

The environment has been solved in this experiment, which is shown in the following figure. As shown, after an initial training, the agent began to obtain higher scores after about 25 episodes. The required score (13.0) is finally reached after 191 episodes.

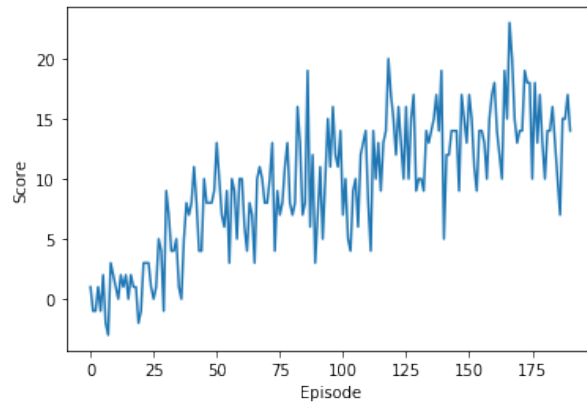


Figure 1: Result