

# **Introduction to Machine Learning**

## **Project 4 Bonus Report**

**UBID: sppandey@buffalo.edu**

**PERSON NO: 50288608**

**Implement DQN using at least TWO environments from OpenAI's Gym library (one of them has to be from Atari). You can use Stable Baselines implementation of DQN, which provides a detailed documentation. In the report for each implemented environment reason how quickly your agent were able to learn using pre-implemented DQN function, plot graphs of the rewards, etc .**

**Answer:**

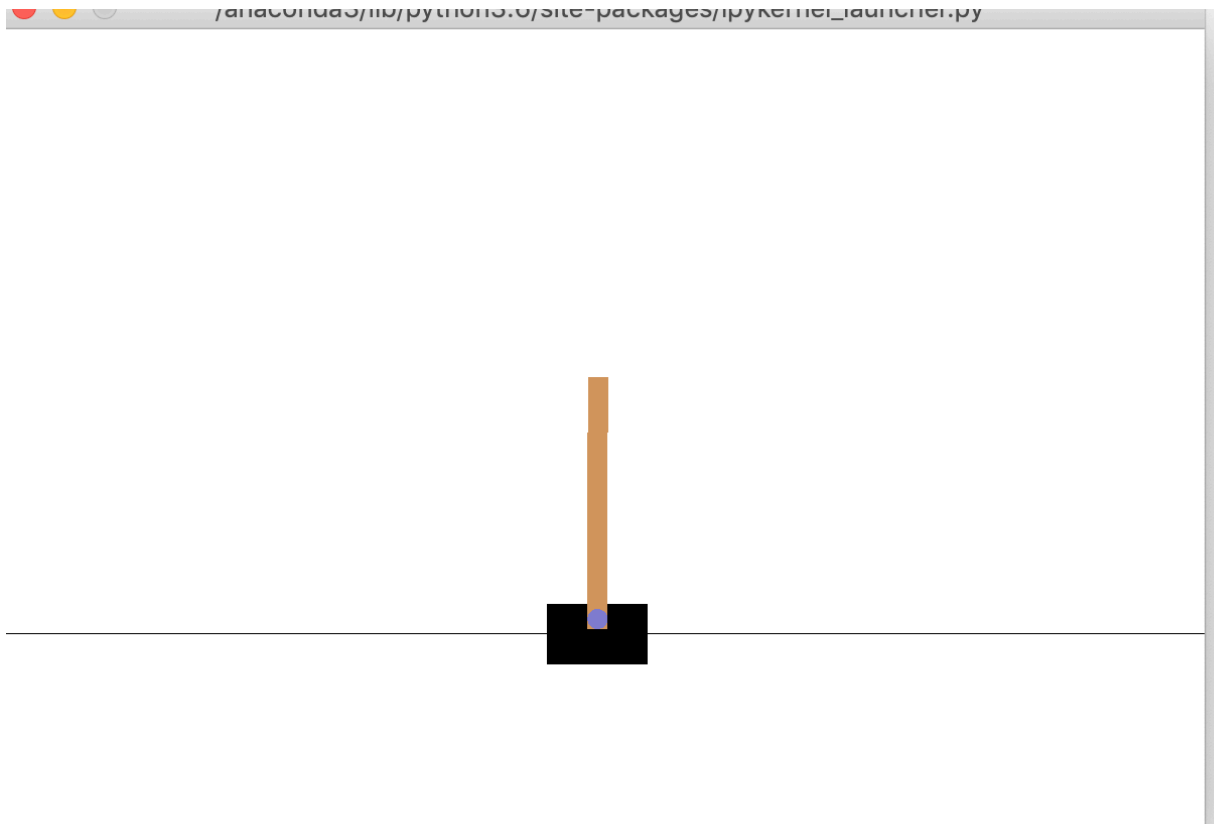
### **Part a) Implementation of DQN using Gym Library :**

We have tried to implement basic code using the Stable baselines implementation of DQN and managed to achieve good results for the cartpole-v gym make.

We tried different steps for training and testing the model, we achieved accurate results when steps were set to 100000.

Result obtained when steps were set to 100000:

% time spent exploring	80
episodes	100
mean 100 episode reward	20.5
steps	2033
% time spent exploring	20
episodes	200
mean 100 episode reward	60.7
steps	8103
% time spent exploring	2
episodes	300
mean 100 episode reward	112
steps	19296
% time spent exploring	2
episodes	400
mean 100 episode reward	119
steps	31234
% time spent exploring	2
episodes	500
mean 100 episode reward	108
steps	42014
% time spent exploring	2
episodes	600
mean 100 episode reward	125
steps	54534
% time spent exploring	2
episodes	700
mean 100 episode reward	162
steps	70680



Since we got pretty good results with number of steps = 100000 we tried another run with steps increased to 1100000 just to check what results could return, thus we got the following results and the above finding still remains close to the best answer we desire.

% time spent exploring	76
episodes	100
mean 100 episode reward	26.9
steps	2661

% time spent exploring	31
episodes	200
mean 100 episode reward	50.7
steps	7732

% time spent exploring	2
episodes	300
mean 100 episode reward	114
steps	19162

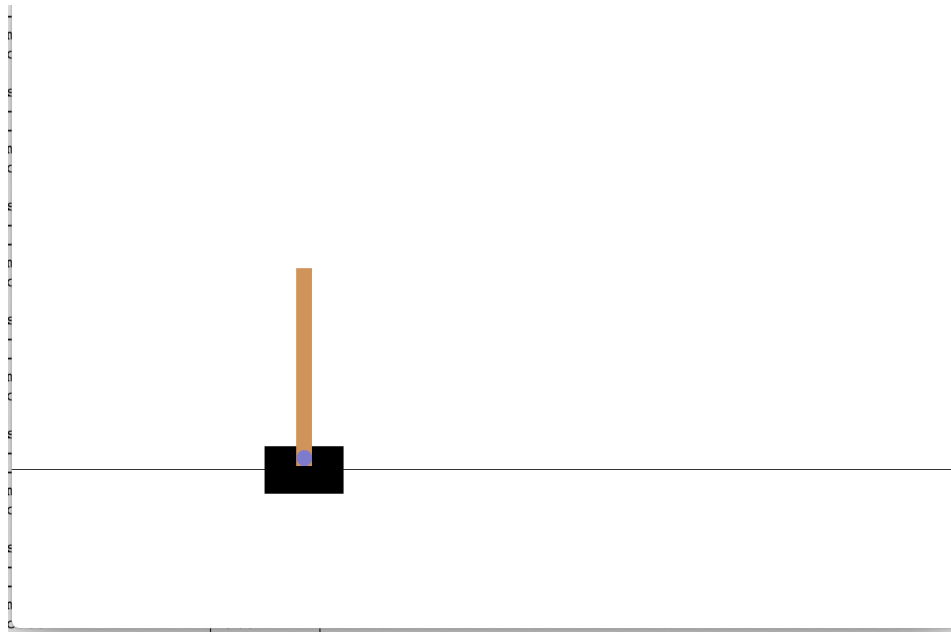
% time spent exploring	2
episodes	400
mean 100 episode reward	103
steps	29469

% time spent exploring	2
episodes	500
mean 100 episode reward	114
steps	40858

% time spent exploring	2
episodes	600
mean 100 episode reward	121
steps	52934

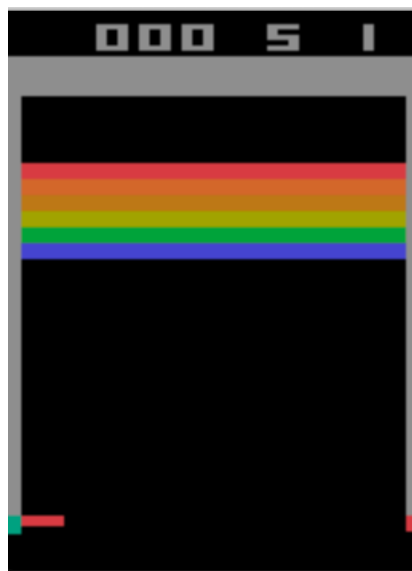
% time spent exploring	2
episodes	700
mean 100 episode reward	172
steps	70139

% time spent exploring	2
episodes	800
mean 100 episode reward	227
steps	92799



### Part a) Implementation of DQN using Atari:

Atari code took significant time to run. We tried to run the code for smaller steps such as 500,700, 900, etc. Few of the python graphic output we could capture is as follows:





Observation: using smaller steps lead to extremely poor performance as the ball didn't even touch the block at the bar. The results however started improving once moved from 500 to 700 and further. Since a significant amount of time was taken by Atari (it needs steps in million to be accurate) we tried on smaller batches and could achieve some success.