5/30/2023

Deep Learning

Project 3 – Deep Reinforcement Learning

Αχιλλέας Παλάσκος (113)
ΠΜΣ – ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ

Contents

1. Atari (PongNoFrameskip-v4)

1.1 Environment

As input to the Neural Network I used 2 consecutive frames of the game so as to capture the movement of the ball. The images were resized to lower dimensions and scaled from 0 to 1.

1.2 Replay Buffer

The Replay Buffer was used to store the current states, actions, rewards, next states and whether or not a terminal state was reached. The total memory size was set to 20000, while when this value was initially set to 50000 I ran out of memory. The replay buffer memory was sampled with a batch size of 32.

1.3 Model Architecture

The model consists of 3 convolutional layers with 32, 64 and 64 filters accordingly and ReLU activations. The kernel sizes and strides were set to 8 and 4, 4 and 2, 3 and 1 accordingly. The convolutional layers were followed by 2 fully-connected layers consisted of 512 units each with ReLU and Linear activations accordingly. In fact, 2 networks with the aforementioned architecture were used. One that was trained and one that was updated every 1000 iterations.

1.4 Methods

4 RL methods were implemented in order to solve this RL problem:

- Deep Q-Learning
- Double Deep Q-Learning
- Dueling Deep Q-Learning
- Dueling Double Deep Q-Learning

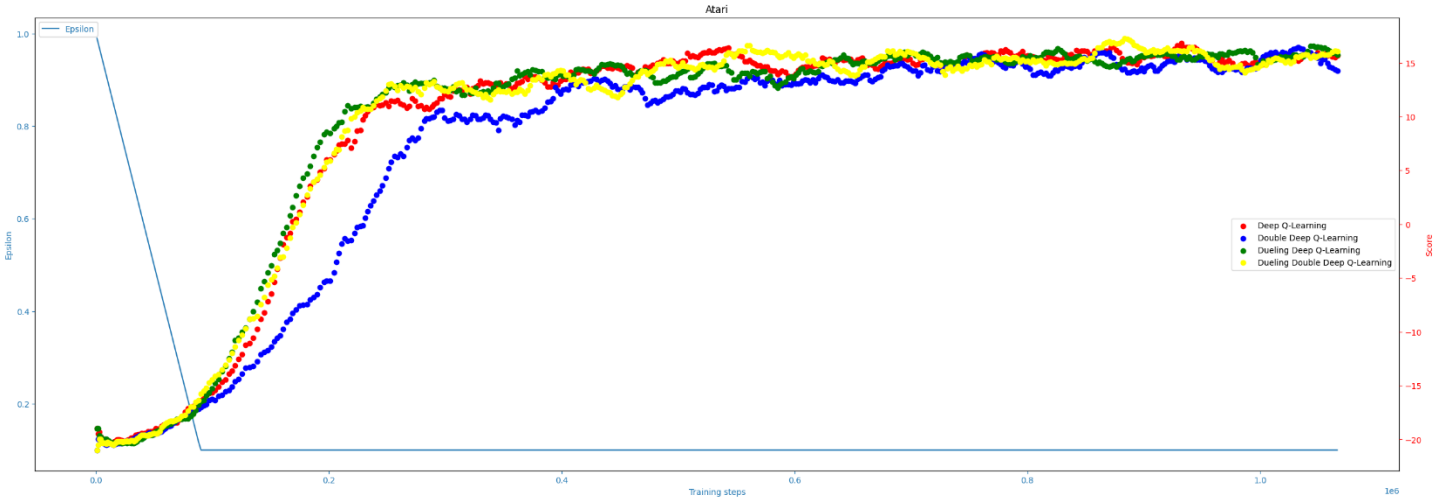
1.4 Hyper-parameters

Number of games, learning rate and γ were set to 500, 10^{-4} and 0.99 accordingly.

1.5 Epsilon

It was initiated to 1 and decreased by 10^{-5} in each iteration until 0.1. After this value was reached it was kept constant.

1.6 Results



1.7 Observations

As we can see from the figure above, learning actually starts when the value of ε has dropped to its minimum value of 0.1. Dueling Deep Q-Learning presents the fastest learning, while Deep Q-Learning is clearly the slowest as far as learning is concerned. However, overall it seems that all 4 methods achieve almost the same score ($\cong 15$) after 500 game played which indicates the game is solved by all of them.