ML and DL in Reinforcement Learning

April 25, 2022

1 Machine Learning and Deep Learning in Reinforcement Learning

You will find in this jupyter notebook a collection of graphs drawned during the training of our different approchaes to solving the Taxi Driver Game from openAI's gym. Each approach comes with the parameters used during training as well as the evolution of the exploration rate (*epsilon*), the reward and the number of steps for each episode done.

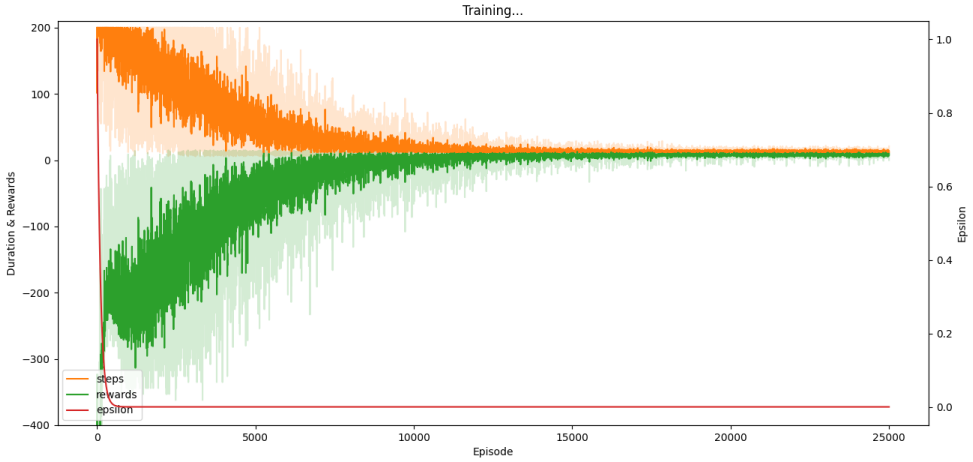
For more informations on the *Deep Q Learning* approach you can read the *DQN.ipynb* notebook located in the *DQN/* folder.

1.1 Machine Learning

1.1.1 Q-Learning

Parameters: - *Number of episodes*: 25 000, - *Learning Rate*: 0.01, - *Gamma*: 0.99, - *Starting Epsilon*: 1, - *Ending Epsilon*: 0.001, - *Epsilon Decay Per Episode*: 0.01

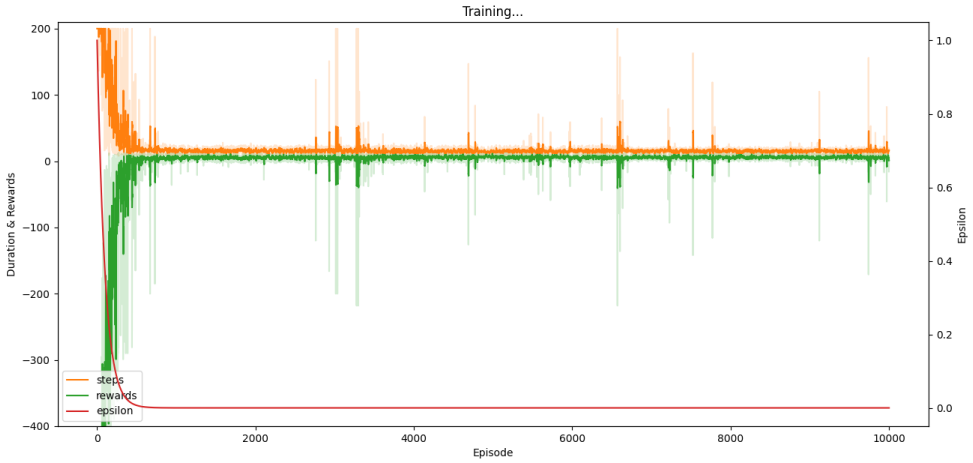
Training Graph:



1.1.2 SARSA

Parameters: - *Episodes*: 10 000, - *Alpha*: 0.85, - *Gamma*: 0.99, - *Starting Epsilon*: 1, - *Ending Epsilon*: 0.001, - *Epsilon Decay per Episode*: 0.01

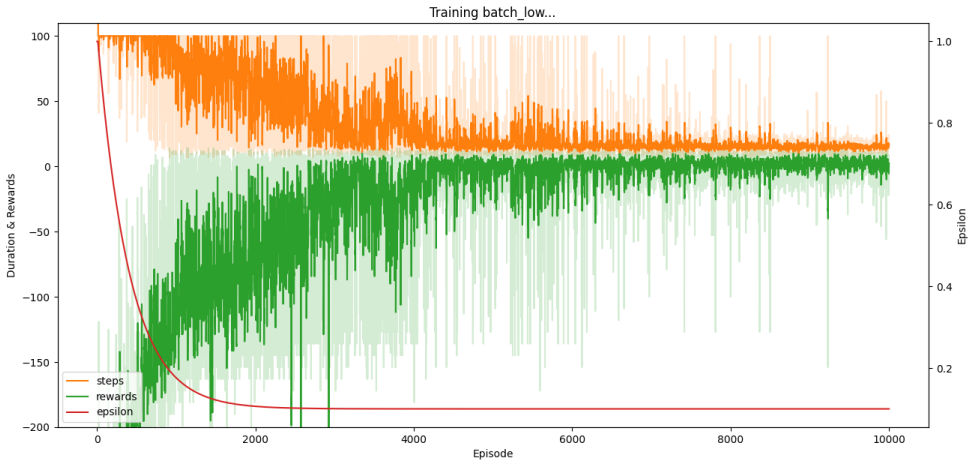
Training graph:



1.1.3 Deep Q Learning

First Architecture Best Parameters (Low Batch Size) Parameters: - *Episodes*:10 000, - *Batch Size*:32, - *Gamma*:0.99, - *Starting Epsilon*:1, - *Ending Epsilon*:0.1, - *Decay Factor Epsilon*:400, - *Number of episodes between model update*: 20, - *Max number of steps per episodes*: 100, - *Warmup Episodes*: 10, - *Starting Learning Rate*: 0.001, - *Ending Learning Rate*: 0.0001, - *Decay Factor Learning Rate*: 5 000, - *Memory Size*: 50 000,

Training graph:



Second Architecture Parameters: - *Episodes*:10 000, - *Batch Size*:128, - *Gamma*:0.99, - *Starting Epsilon*:1, - *Ending Epsilon*:0.1, - *Decay Factor Epsilon*:400, - *Number of episodes between model update*: 20, - *Max number of steps per episodes*: 100, - *Warmup Episodes*: 10, - *Starting Learning Rate*: 0.001, - *Ending Learning Rate*: 0.0001, - *Decay Factor Learning Rate*: 5 000, - *Memory Size*: 50 000,

Training graph: