

Bonus task

Vidyadhari chintala
50290085

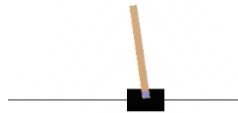
Introduction:

This report provides details on how Reinforcement learning mechanism is used in order to play games. It address the process of solving cart pole & Atari - breakout games. The environments for these games are provided by GYM which is a toolkit for developing and comparing reinforcement learning algorithms.

Problem statement:

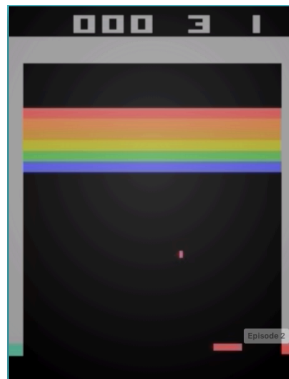
Cart Pole:

The environment of cart pole problem is as shown in the figure. The task is to balance the pole on the cart against gravity such that it would not fall down.



Atari breakout:

The task is to play the game of Atari and score as many points as possible.



CartPole problem:

The environment is important from Gym. Using the following hyper parameter values the model was able to train to some extent as shown in the fig:

$\text{GAMMA} = 0.99$

$\text{LEARNING_RATE} = 0.001$

$\text{MEMORY_SIZE} = 10000$

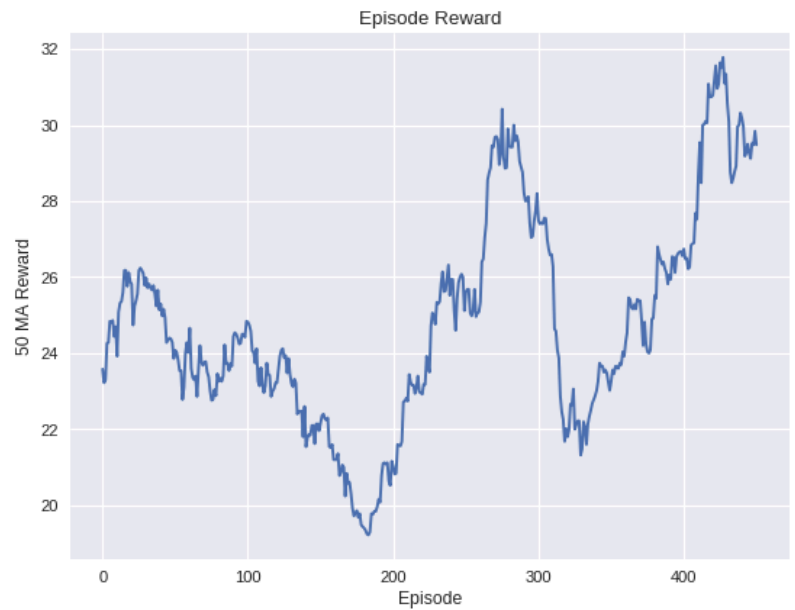
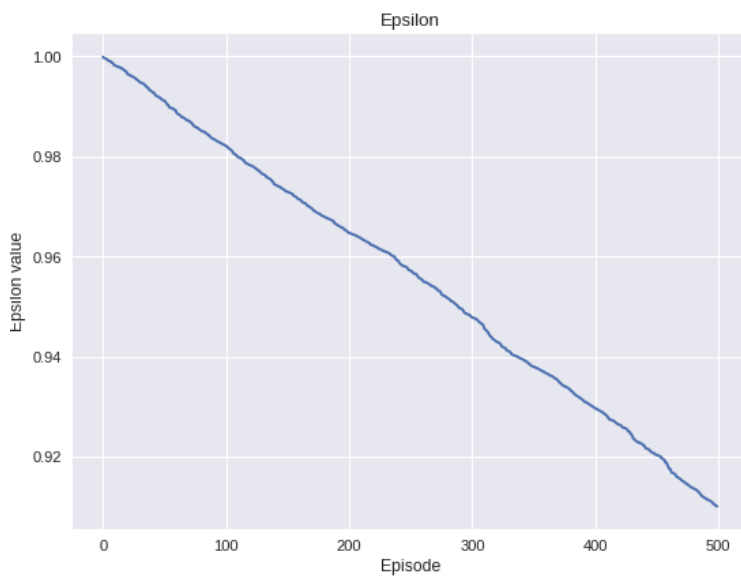
BATCH_SIZE = 20

max_epsilon = 1.0

min_epsilon = 0.01

Lambda = 0.000008

Number of episode: 500



It could be observed that as the exploration rate is gradually reduced with the number of iterations the agent gets trained where the reward was increasing. As the exploration decay given was very small the agent would take much time to get trained. It could also be observed that as the exploration rate is very slowly decreasing the agent ends into many random actions that shouldn't actually be the case.

Tuning the exploration decay factor - lambda:

GAMMA = 0.99

LEARNING_RATE = 0.001

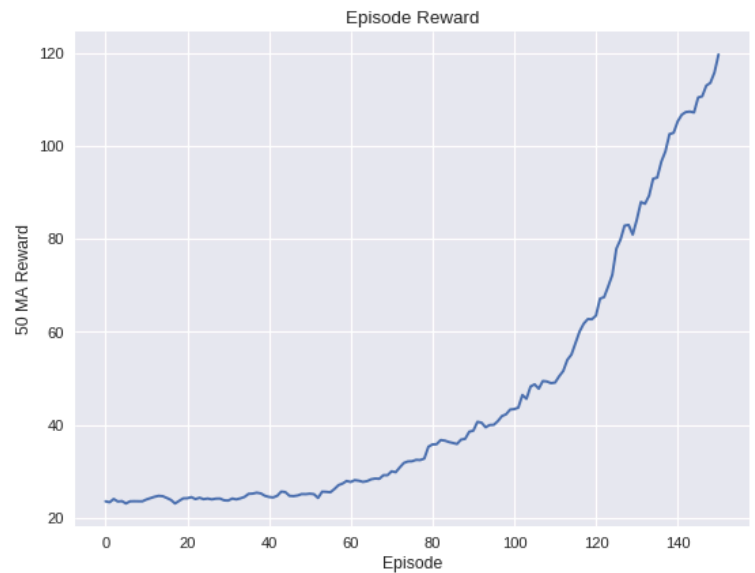
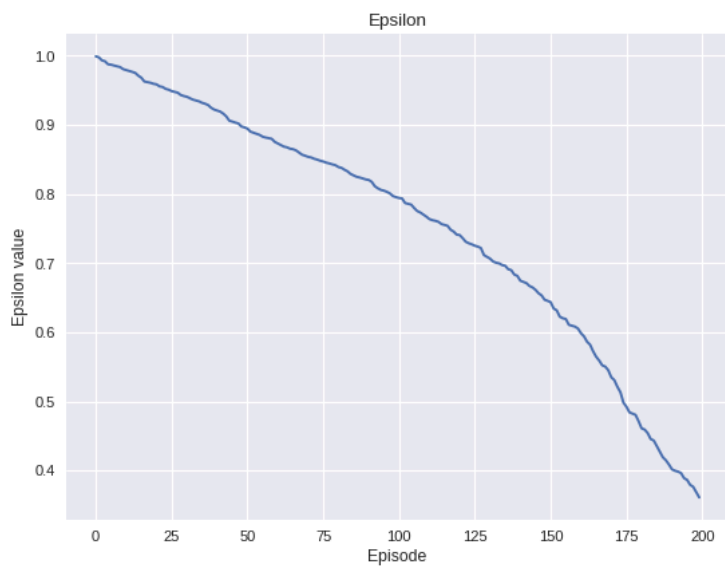
BATCH_SIZE = 20

max_epsilon = 1.0

min_epsilon = 0.01

Lambda = 0.0001

Number of episode: 200



Now it could be clearly observed that as the exploration rate is reduced to 0.4 for 200th iteration the agents receiving the reward at each state kept on increasing indicating the fact the agent has good knowledge about its environment that would help to achieve its final goal.

Atari- Breakout:

Using reinforcement learning for training Atari-breakout:

ENV_NAME = "Breakout-ram-v0"

GAMMA = 0.99

LEARNING_RATE = 0.001

BATCH_SIZE = 20

max_epsilon = 1.0

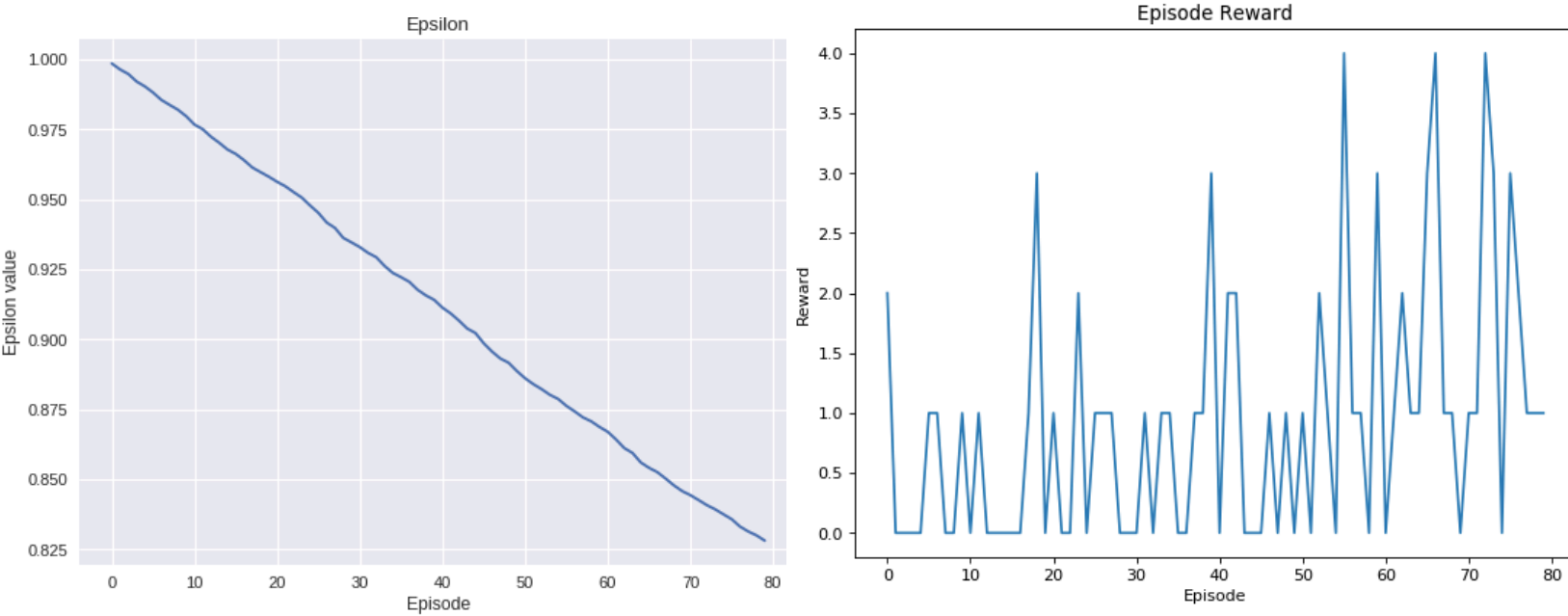
min_epsilon = 0.01

lamb = 0.00001

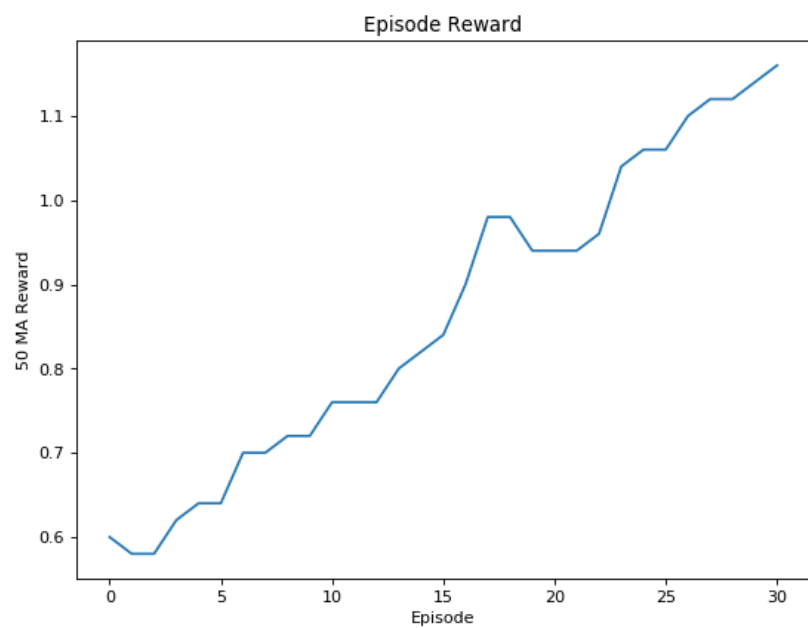
Number of episode: 80

Timesteps: 20000

It could be observed that the agent is getting trained i.e its rewards increases as the number of episodes are increased and also as the exploration rate decreases. If we increase the number of episodes further, then the agent would be completely trained and would have more knowledge about the environment in which it is working.



Smoothened output:



Increasing the number of episodes resulted in better results:

ENV_NAME = "Breakout-ram-v0"

GAMMA = 0.99

LEARNING_RATE = 0.001

BATCH_SIZE = 20

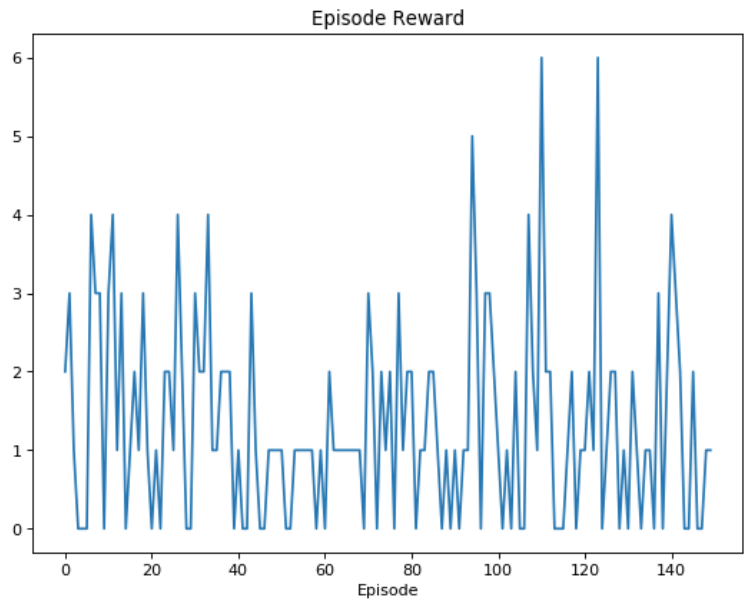
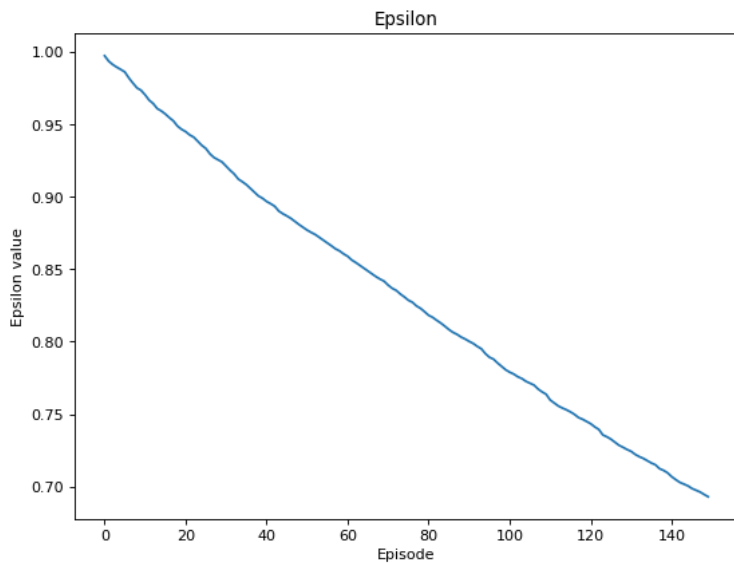
$\text{max_epsilon} = 1.0$

$\text{min_epsilon} = 0.01$

$\text{lamb} = 0.00001$

Number of episode: 150

Timesteps: 20000



The rewards has increase from 4 to 6 as we increase the number of episodes from 80 to 150.

Conclusions:

- As the number of episodes increase the rewards are increasing
- Decrease in exploration rate as the number of episodes increases is also very important factor for the agent to get trained