RL P3

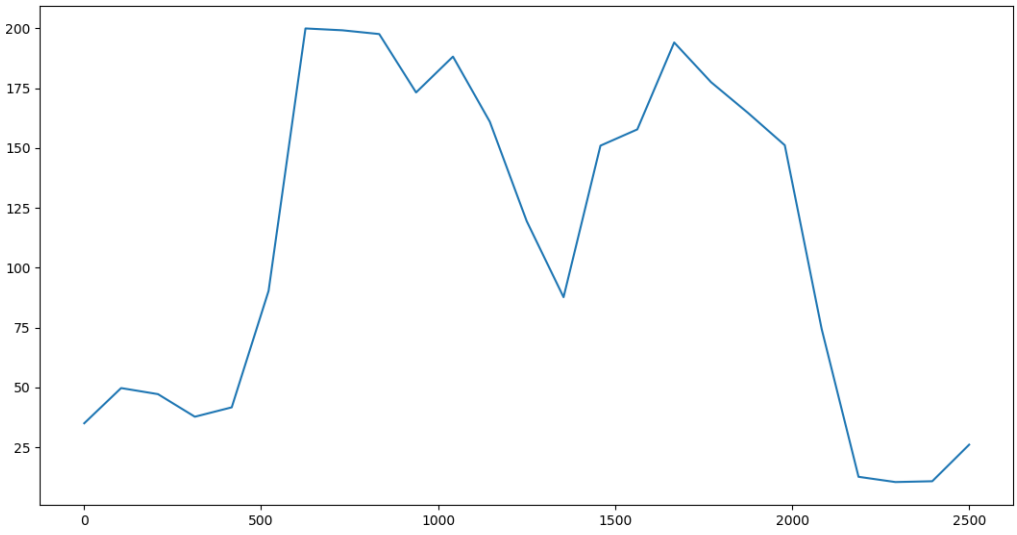
Rishabh

May 2019

1 1.DQN

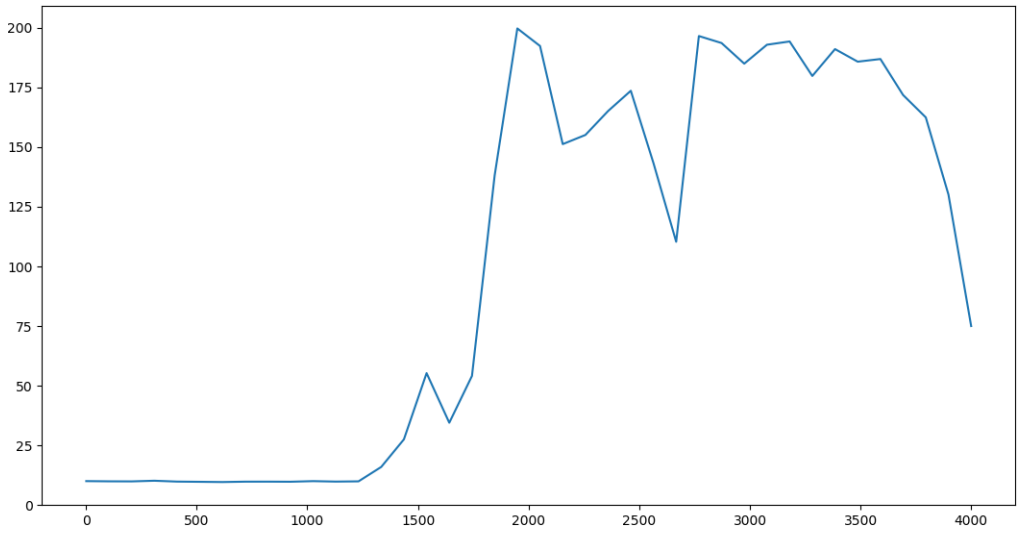
Standard Hyperparameters used:
Hidden Layer Size:100 , Learning rate = 0.0001 , Batch size =20 , Epsilon = 0.1 , Memory = 10000

Figure 1:



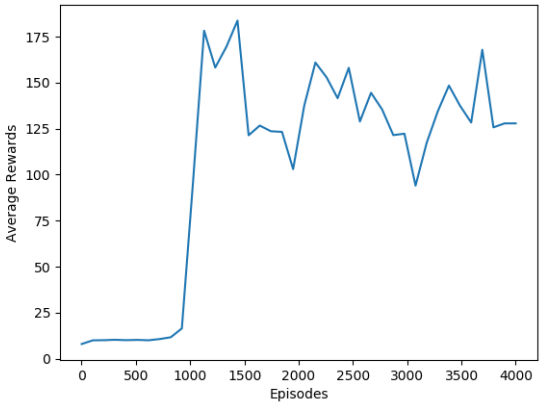
BATCH$_SIZE = 10$

Figure 2:



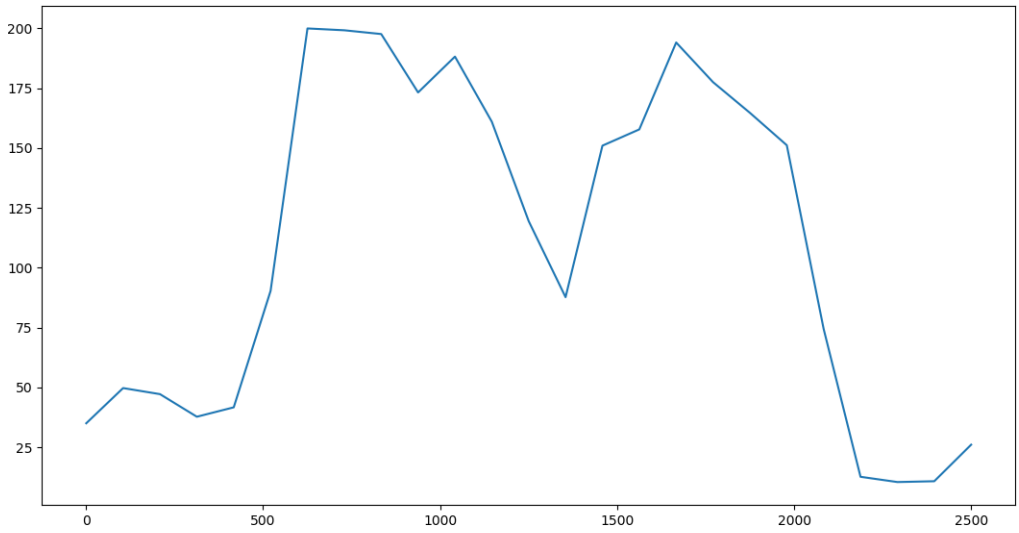
BATCH SIZE = 30

Figure 3:



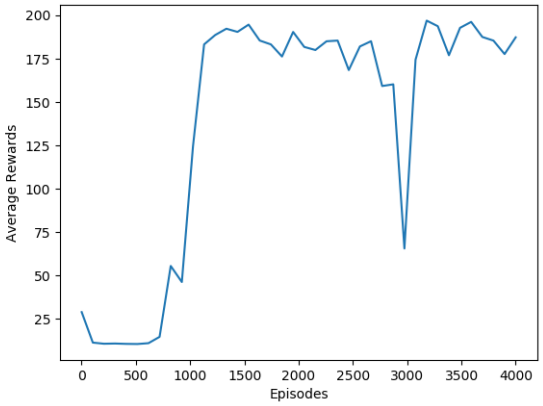
HIDDEN LAYER SIZE = 20

Figure 4:



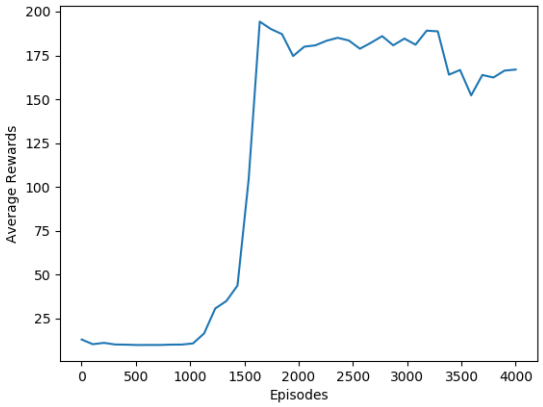
EPSILON = 0.2

Figure 5:



MEMORY =40000

Figure 6:



2 Conclusion

We saw that with increase in memory with few episodes DQN is solved.This is because we get more variety when we sample from larger memory.Apart from that with increase in batch size average reward value decreases.Also on increasing epsilon we see that cart-pole is solved much faster.

If we remove target network or experience replay then there is no superevision hence no learning.