DQL vs DQLT

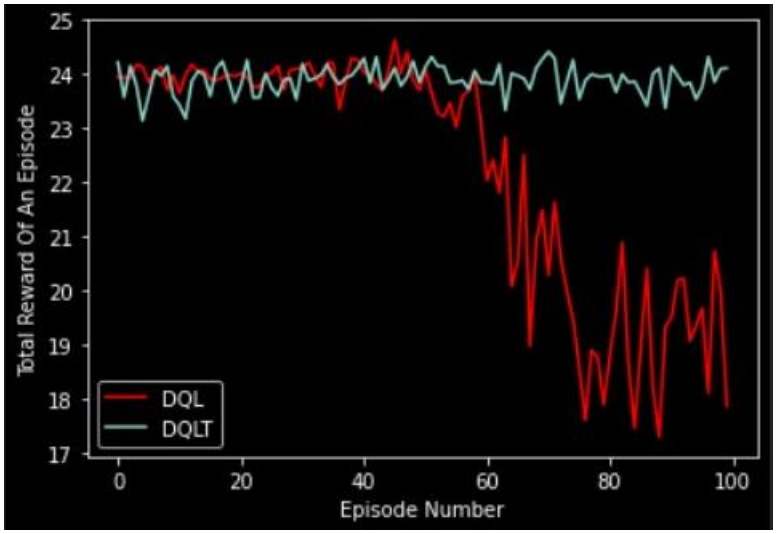
EVERYTHING IS AVERAGED OVER 50 RUNS

1.

Total Episodes=100

alpha=0.7

e=0.1



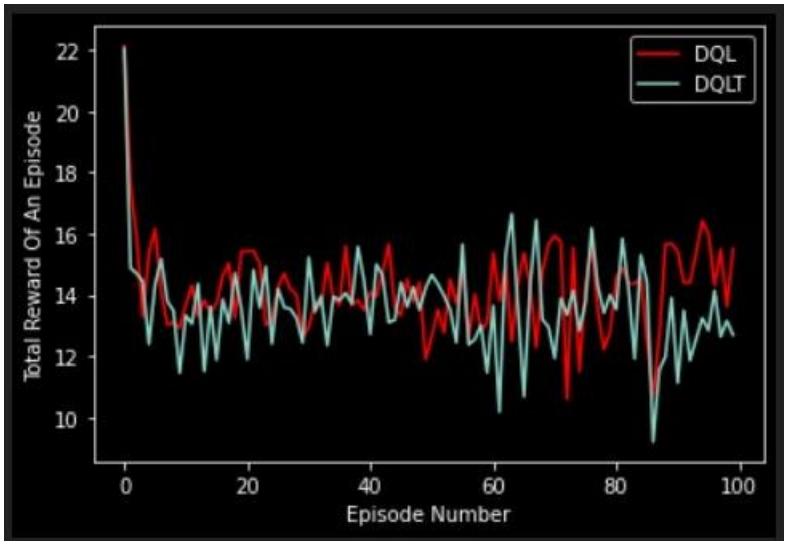
DQLT performs better. Alpha is high.

2.

Total Episodes =100

alpha=0.01

e=0.2



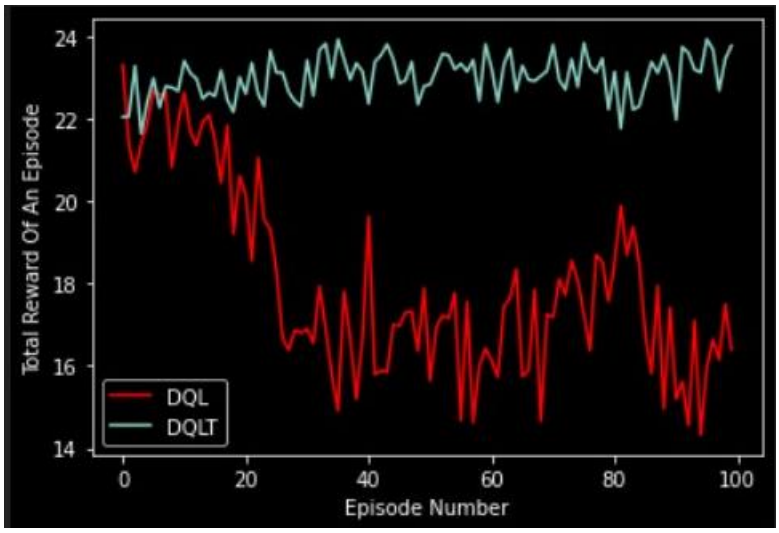
DQL performs better. Alpha is low.

3.

Total Episodes=100

alpha=1

e=0.2



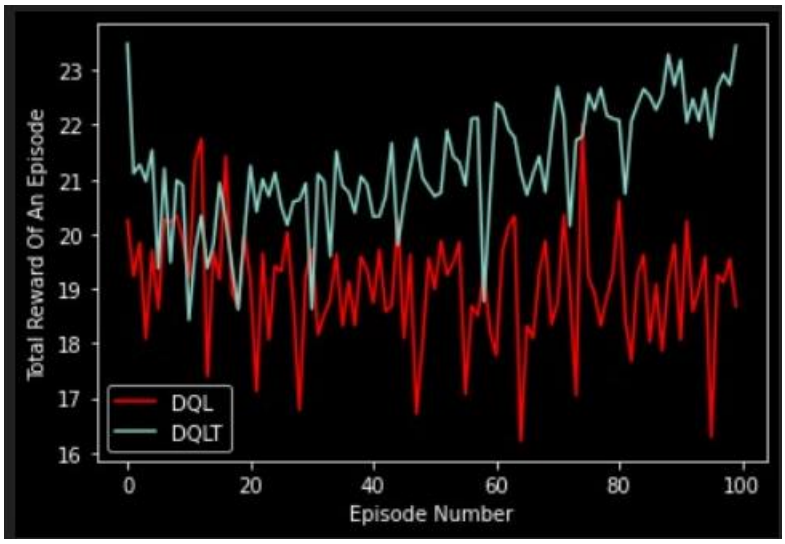
DQLT performs better. Alpha is high

4.

Total Episodes=100

alpha=0.1

e=0.2



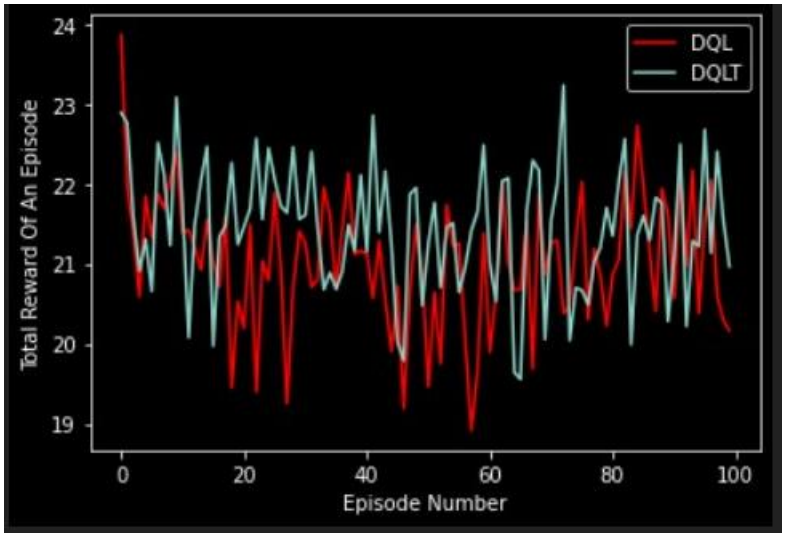
DQLT performs better. Alpha is modrate.

5.

Total Episodes=100

alpha=0.05

e=0.15



EXPLANATION

We see that for larger alphas DQL performs poorly, its performance degrades over time. However, DQLT is sort of constant with its rewards. For smaller alpha (0.01), DQL starts to out perform DQLT in the later episodes.

DQL picks the maximizing action as a target deterministically, however DQLT picks the maximizing action as a target with e-greedy policy. So once in a while our target could be the non greedy Q value. Hence, for larger steps, if we are too biased, we could degrade our performance while picking the maximizing target. But if out step size is lower, then even for times we are biased, we don't take a large step towards wrong estimates and hence DQL is a bit better for lower alpha.

DQL is getting biased way too easily by maximizing target. Its performance is hindering and as a result of high alpha it's performance is reducing even more. For DQLT, it is picking its target with e-greedy for atleast 1 table and as a result for higher alphas it is performing better. For lower alpha, it isn't learning al that well as well since the update step we are taking is too small.

Constraints:

Averaging over 50 runs

Episode Length 50 Steps

Episodes =100

Increasing this was taking way too long on my computer, so I had to do with this.