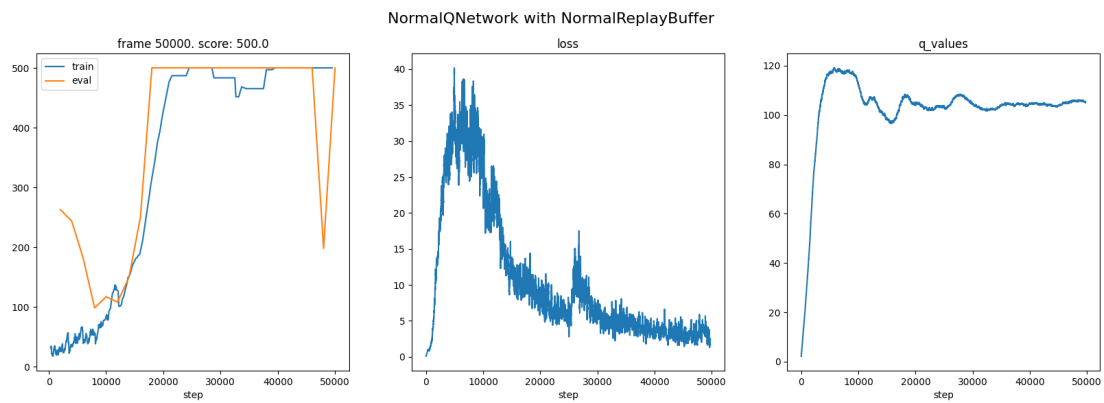


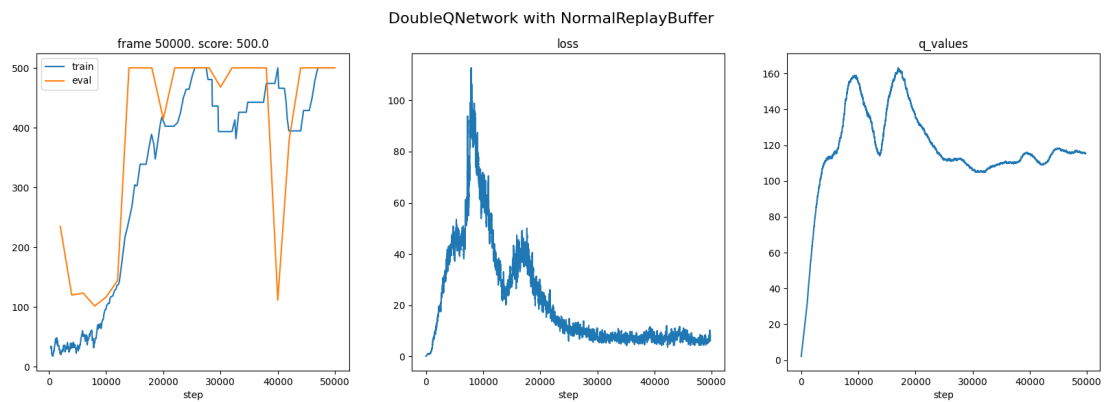
DRL Homework2

DQN



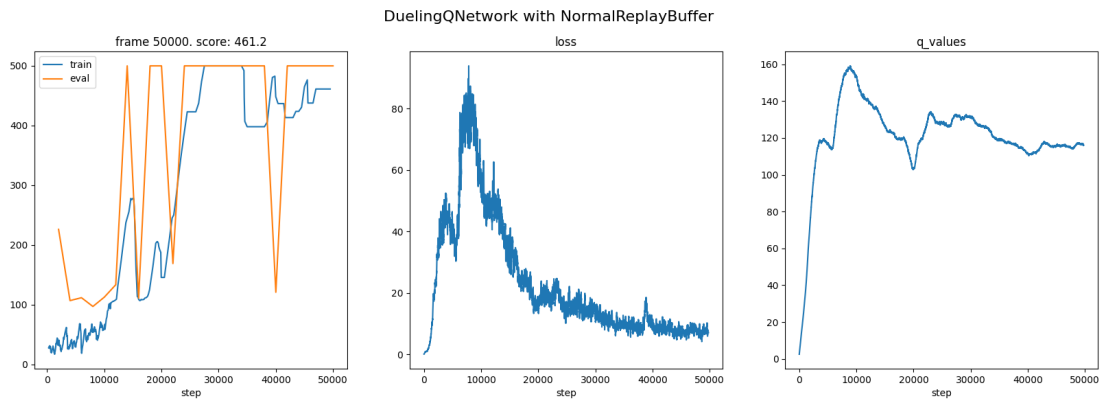
Seems steady but the Q value is not very big.

Double DQN



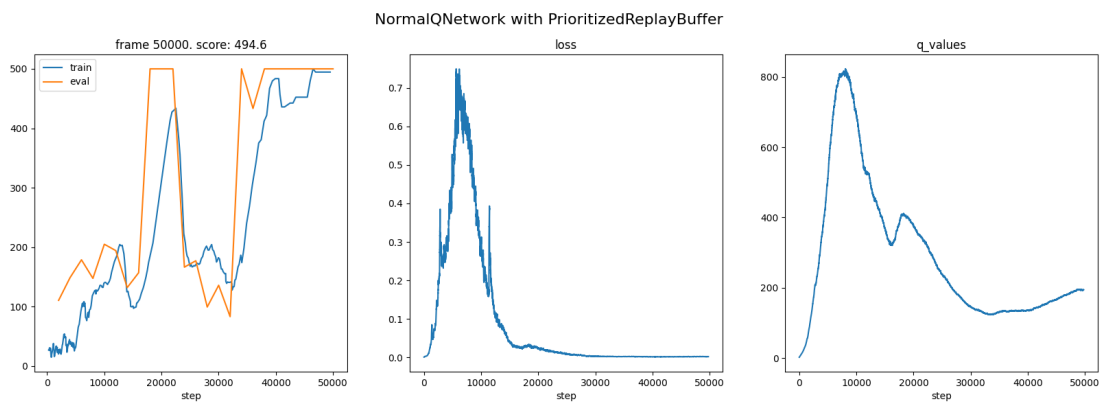
Sometimes not steady, but some times can find the greater Q value.

Dueling DQN



More unstable than the DQN and Double DQN. But it can find the biggest Q value among the DQN and Double DQN.

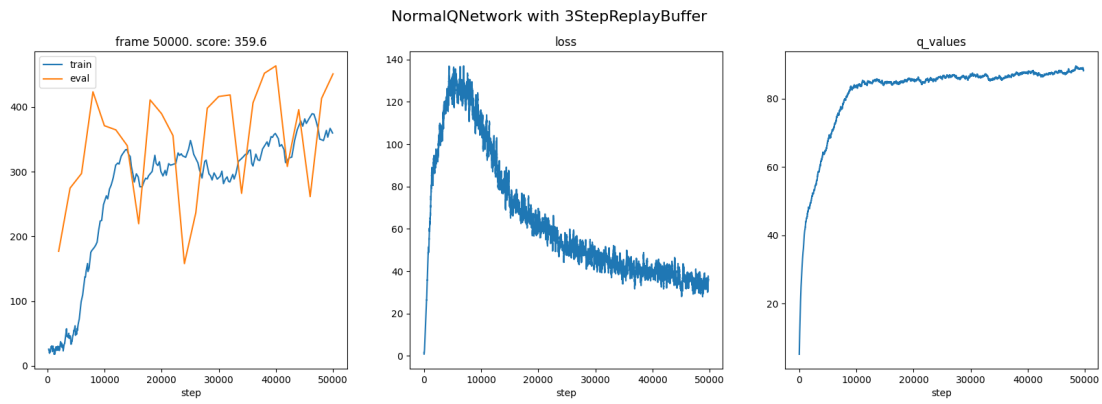
Prioritized Experience Replay



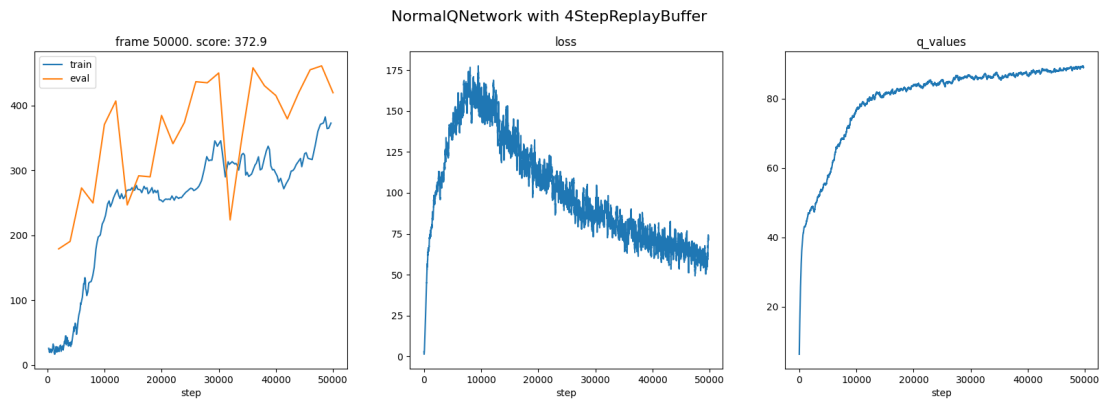
PER's training takes longer and fluctuates more. Since it will exceed the original capacity, I will adjust the capacity to 60000 here.

N-Step Return

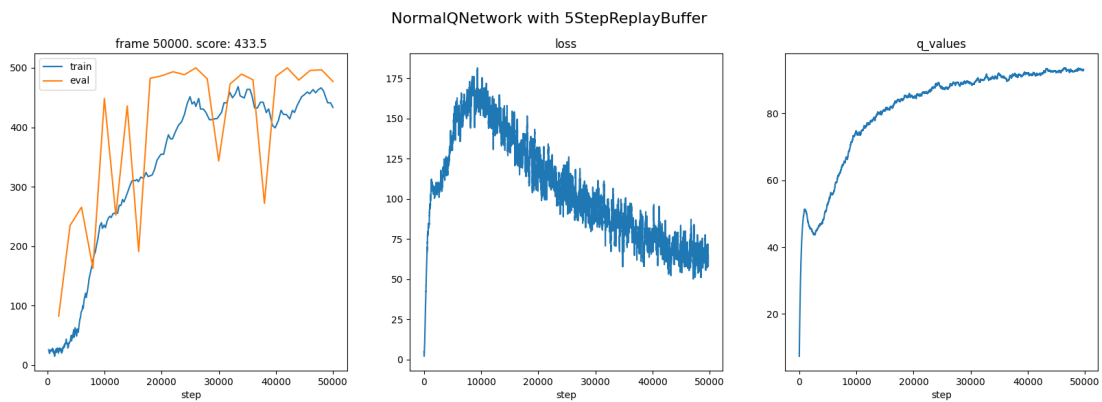
n=3:



n=4



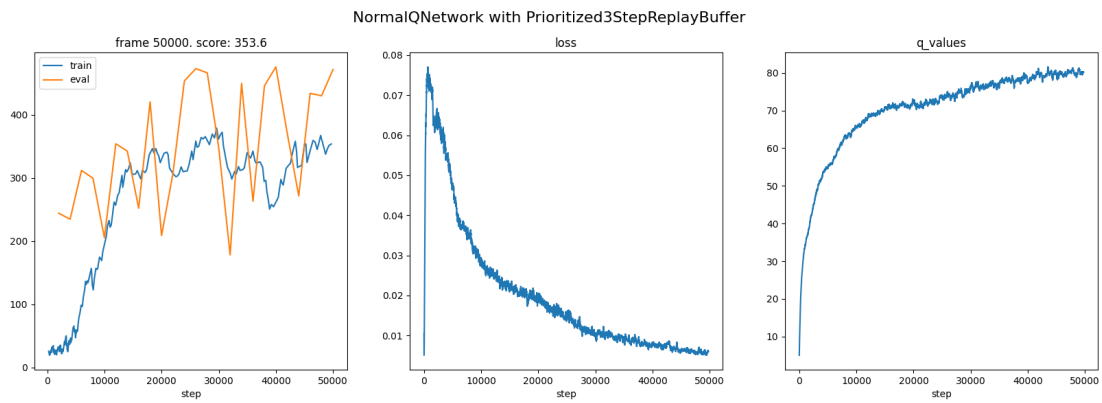
n=5



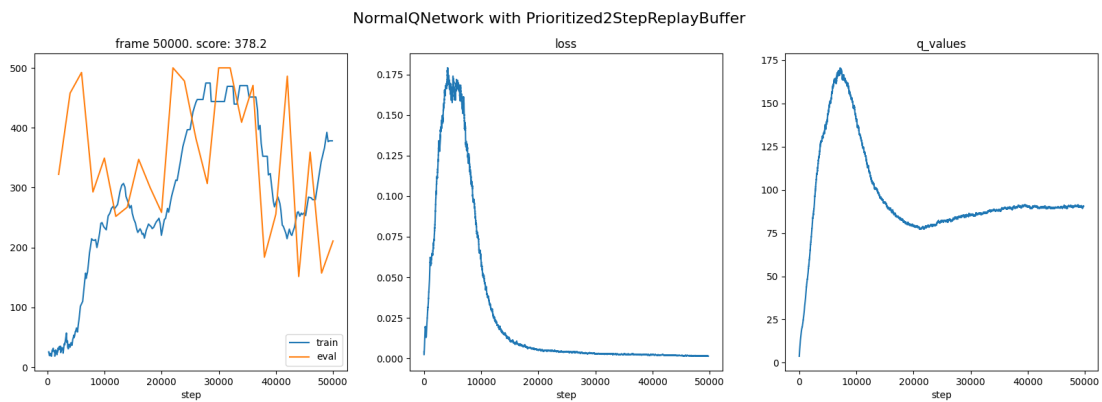
The bigger n, the wroth performance.

NStep+PER

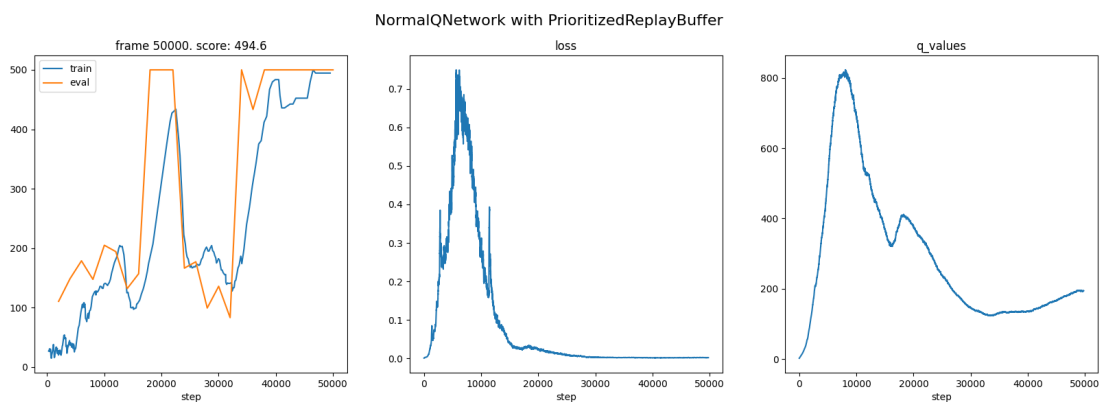
n=3



n=2

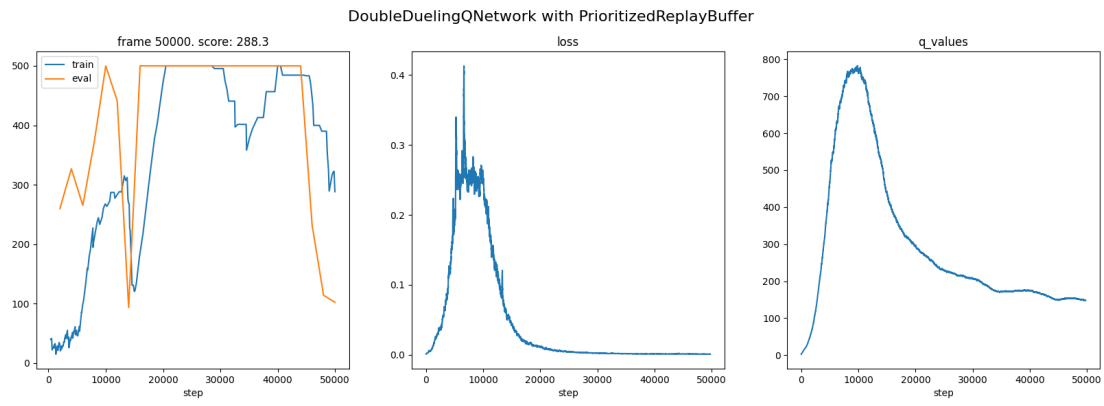


n=1



The bigger n, the wroth performance.

All in One



It seems the State Of The Art in these models.