AI Assignment-4

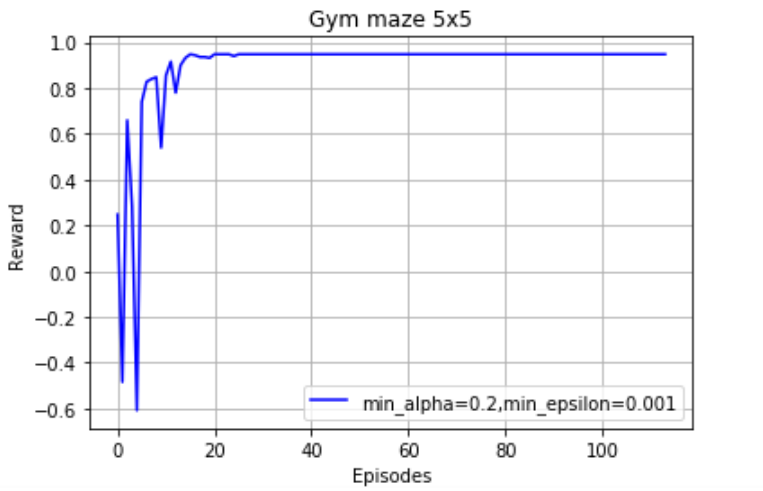
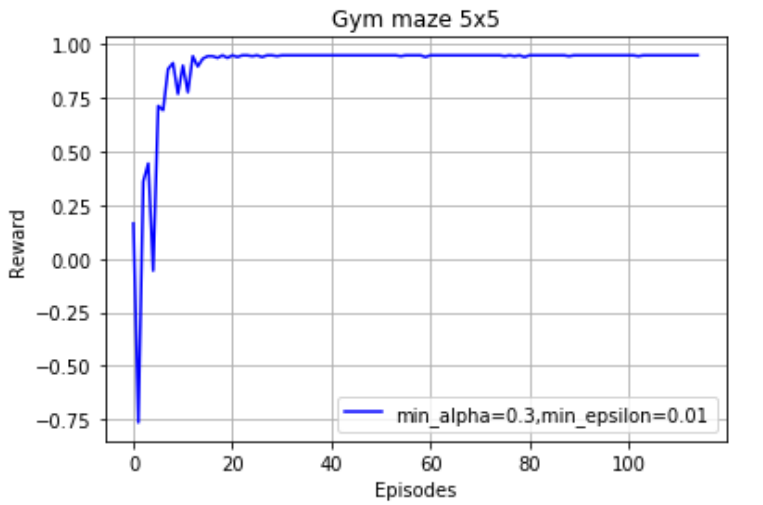
MT18117-Subhani shaik

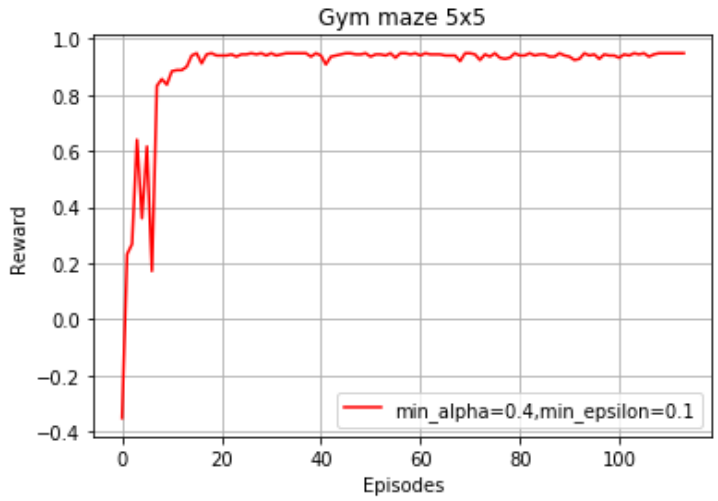
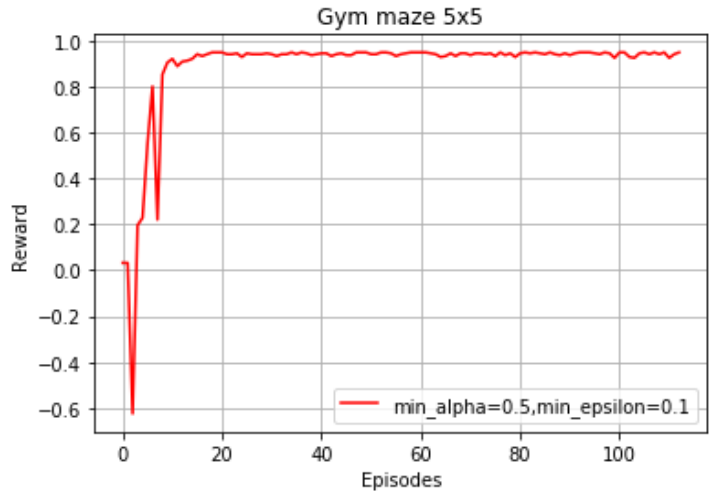
1.
a,b).

Initially I have taken the sample 5x5 maze with minimum learning rate (alpha)=0.2 and minimum epsilon(exploration rate)=0.001 and in every episode I am decaying both the parameters.

So, that exploration rate will effect the model like initially it acts as exploitation and then later as exploration.

I tried this with different values for hyperparameters and used discount rate(gamma)=0.99.

Sno	Hyperparameters	Plot
1	min_alpha=0.2 min_epsilon=0.001 gamma=0.99	
2	min_alpha=0.2 min_epsilon=0.001 gamma=0.99	

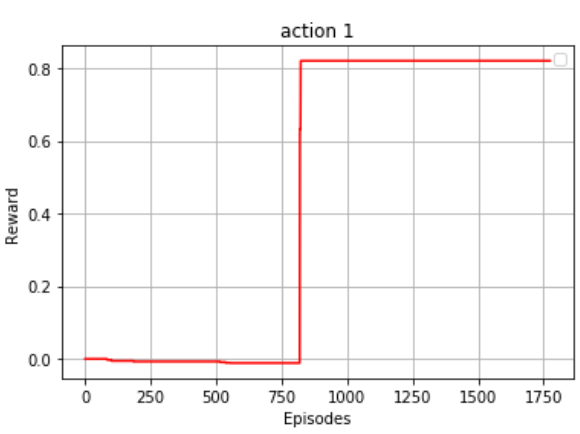
3	min_alpha=0.2 min_epsilon=0.001 gamma=0.99	
4	min_alpha=0.2 min_epsilon=0.001 gamma=0.99	

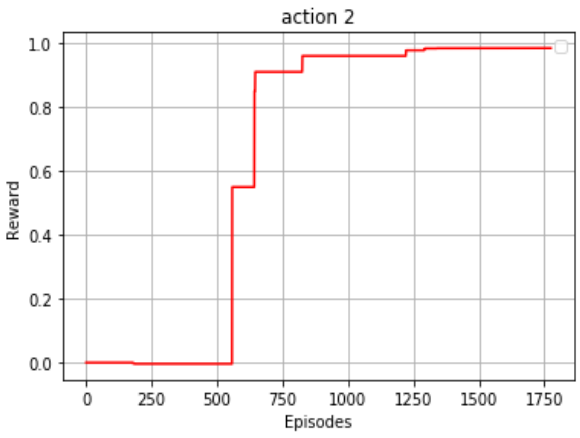
By increasing learning rate(alpha) it is oscillating and over-shooting convergence.

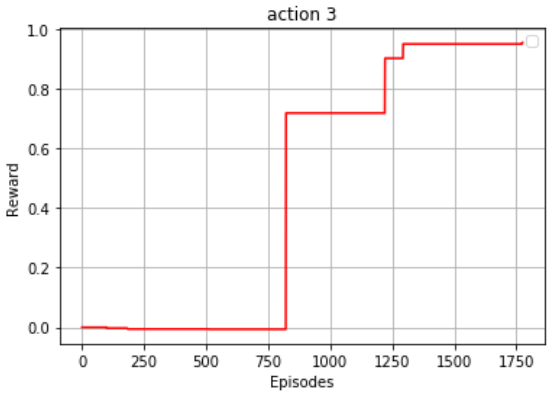
c,e.

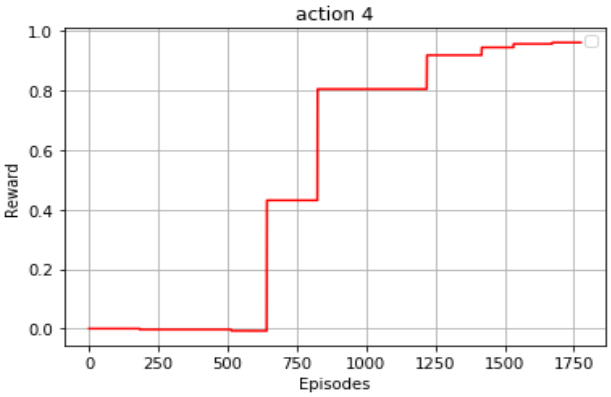
I have chosen particular state (3,2) and taken different actions on it in Qtable.Initially the probability for this state is same on every action. Later it changes by updating the Qtable values, so that the probability of actions changed.

Plots for Qtable value for chosen state:







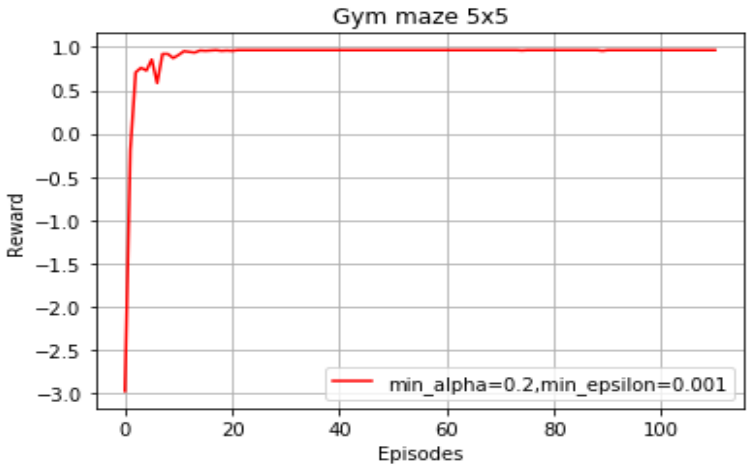


Here I plotted graphs for every action on the same chosen state.On every action initially it is same i.e, start from 0 and then later changed due to change in updating Qtable with best action values.

d.

After trained on different hyperparameters, I picked best one i.e., model 4 hyperparameters Qtable and tested on new random 5x5 maze and it is converging.

Plot for new 5x5 random maze:



Assumptions:

- In every episode, if it is giving path less the size of maze then I am counting it and if this counter equals to 100 times then I am stopping it by considering convergence of optimal path.