

180100095

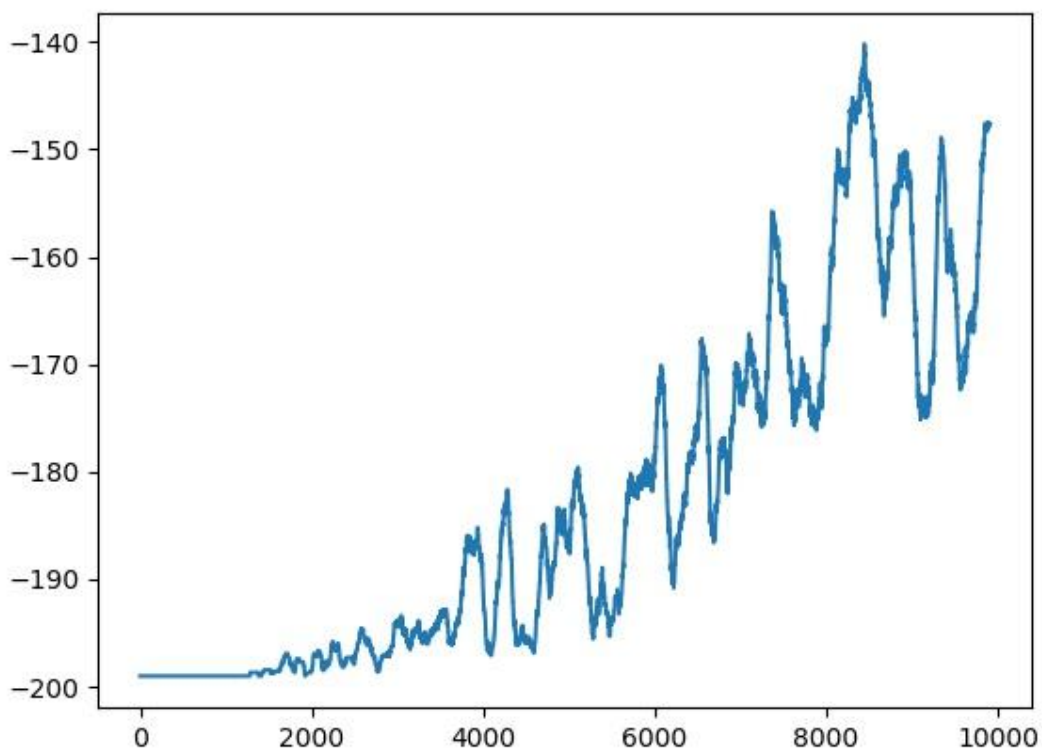
Rishab Khantwal

TASK 1

$\epsilon_{T1} = 0.05$

$\text{learning_rate}_{T1} = 0.1$

Discrete_space of 40 values was used making weights of shape [40,40,3]



- 1) The expected reward of over 100 episodes is increasing as episodes are increasing
- 2) Final reward is almost close to the reward on the test

TASK 2

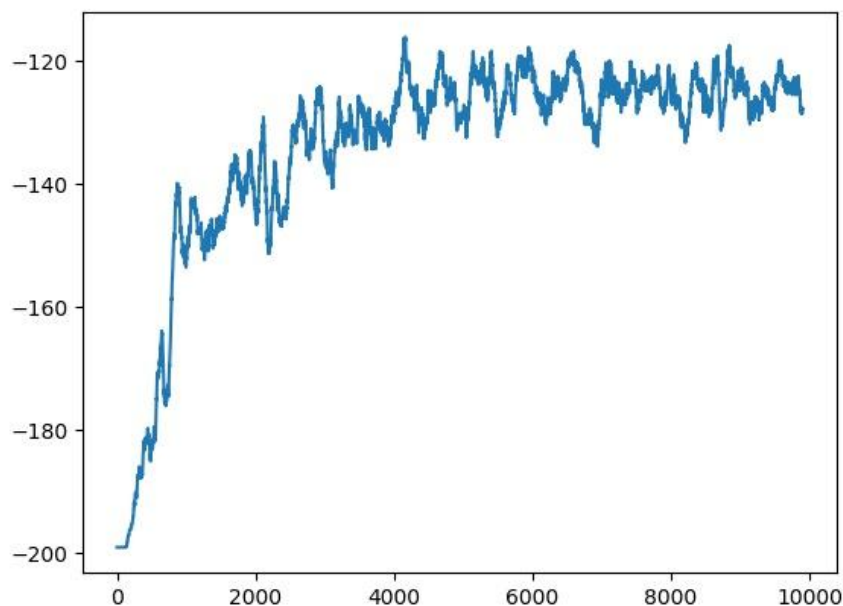
RBF =

1/5	1/4	1/3	1/4	1/5
1/4	1/3	1/2	1/3	1/4
1/3	1/2	1	1/2	1/3
1/4	1/3	1/2	1/3	1/4
1/5	1/4	1/3	1/4	1/5

$$W = W + \lambda * F * \text{RBF}$$

Element wise multiplication

All weights around the main state are updated based on the observation. The update is made proportional to the RBF value of that index.



- 1) Value increased faster than tabular one
- 2) The final value is not the same as in test value