AI Reinforced Learning Lab
Group 43

Group Members
210620M
210621R
210626L
210627P

Comment on the influence the above parameters have on how fast q-learning can converge. Plot the necessary graphs to justify your answer.

Learning Rate

Learning Rate is the hyperparameter in the training model for the rate that the model learns from the previous states to update the q table.

High Learning Rates lead to faster initial learning because previous stages are being considered highly in updating the q table. This also causes high oscillations or divergence in the q table values.

Low Learning Rates lead to slower convergence in the q table. But this convergence is much more stable.

Discount Rate

Discount Rate is the parameter in the training model scenarios that determines how much the previously owned reward values from states are affected to the final cumulative reward. It affects the balance between immediate rewards and future rewards.

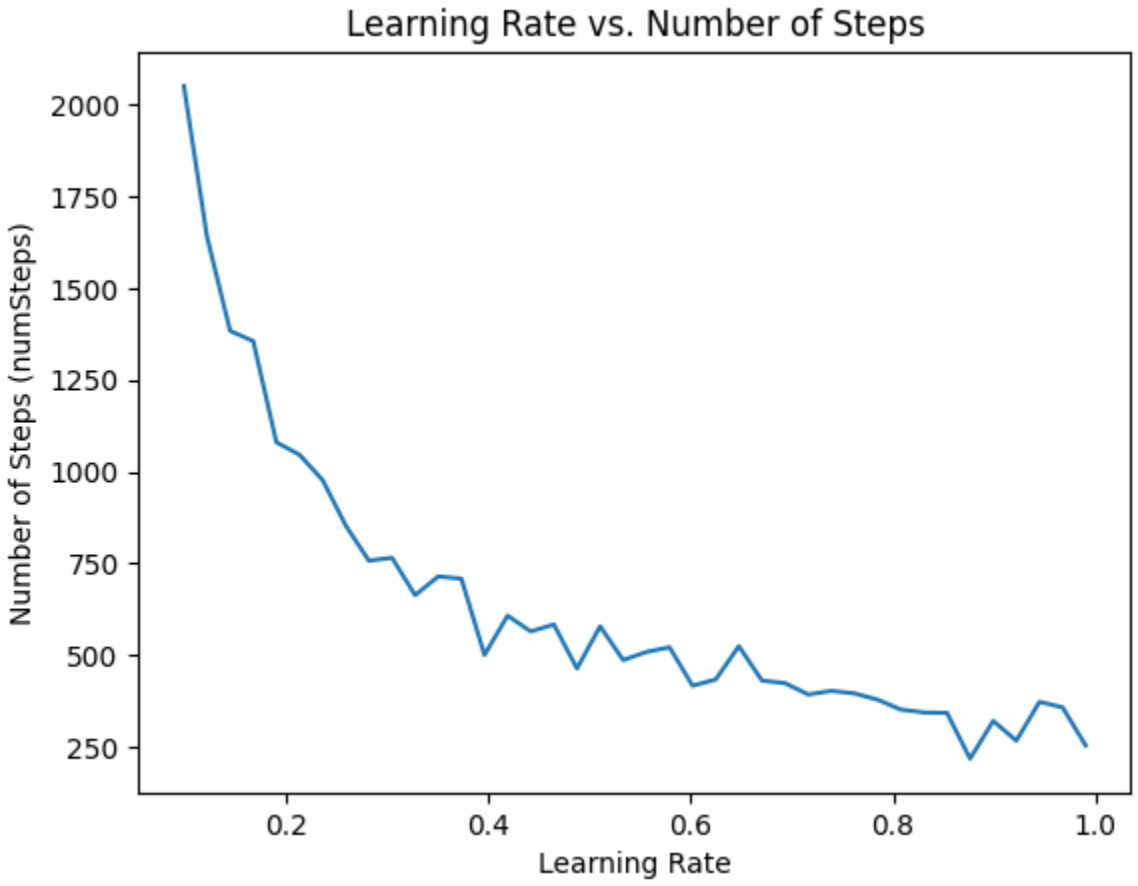
A high discount rate emphasizes long-term planning, potentially leading to slower convergence.

low discount rate focuses on immediate rewards, which might result in faster convergence.

Learning Rate vs Number of Episodes to Converge the Q table.

discount_rate = 0.8
epsilon = 1.0
decay_rate = 0.005
converging_epsilon = 0.001 (Q tables maximum tolerated change when converged)

Learning Rate	Number of Episoded (numSteps)
0.1	2052
0.122821	1641
0.145641	1384
0.168462	1356
0.191282	1080
0.214103	1046
0.236923	977
0.259744	853
0.282564	757
0.305385	765
0.328205	663
0.351026	714
0.373846	708
0.396667	500
0.419487	607
0.442308	564
0.465128	583
0.487949	463
0.510769	578
0.53359	486
0.55641	508
0.579231	521
0.602051	416
0.624872	433
0.647692	524
0.670513	430
0.693333	423
0.716154	392
0.738974	402
0.761795	395
0.784615	378
0.807436	351
0.830256	343
0.853077	342
0.875897	217
0.898718	320
0.921538	266
0.944359	372
0.967179	357
0.99	253



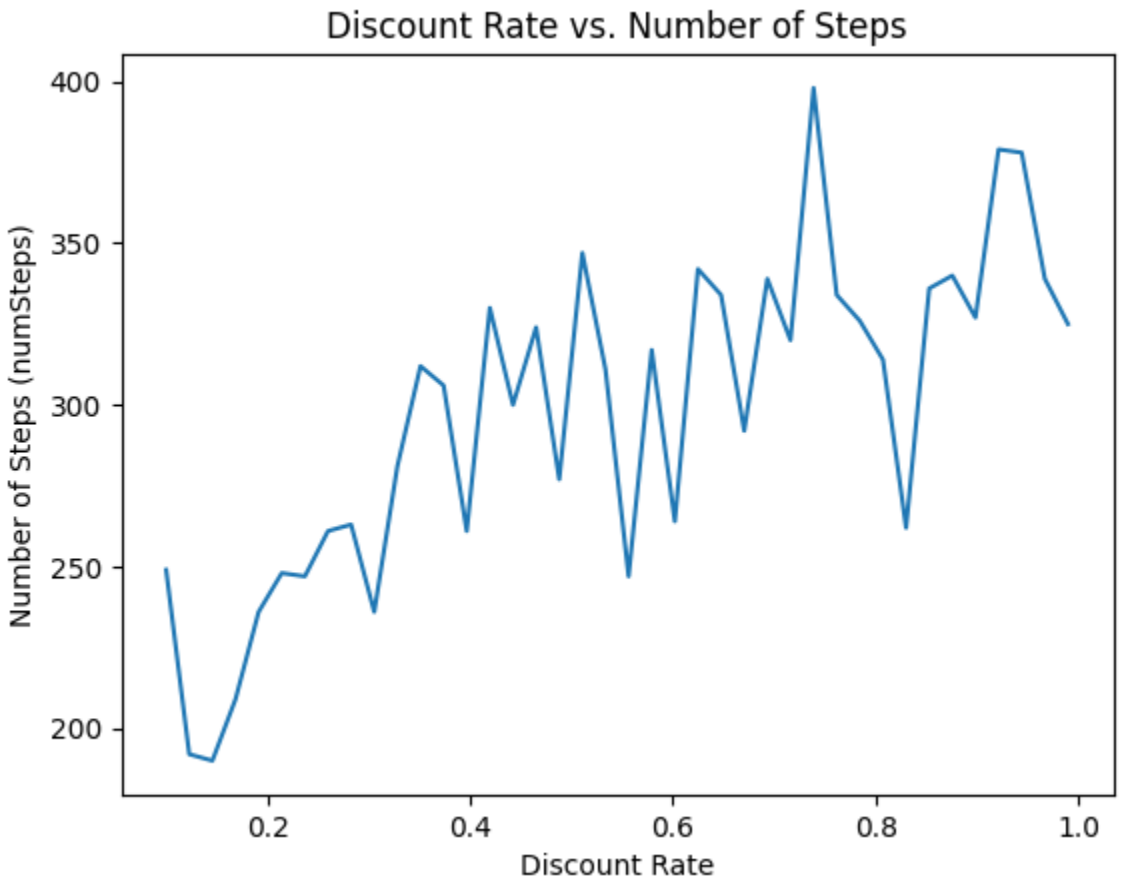
Here,

- High Learning Rate causes a faster convergens (Less number of steps)
- Low Learning Rate causes a much slower convergence (High number of steps)

Discount Rate vs Number of Episodes to Converge the Q table

learning_rate = 0.9
epsilon = 1.0
decay_rate = 0.005
converging_epsilon = 0.001 (Q tables maximum tolerated change when converged)

Discount Rate	Number of Steps (numSteps)
0.1	249
0.122821	192
0.145641	190
0.168462	209
0.191282	236
0.214103	248
0.236923	247
0.259744	261
0.282564	263
0.305385	236
0.328205	281
0.351026	312
0.373846	306
0.396667	261
0.419487	330
0.442308	300
0.465128	324
0.487949	277
0.510769	347
0.53359	311
0.55641	247
0.579231	317
0.602051	264
0.624872	342
0.647692	334
0.670513	292
0.693333	339
0.716154	320
0.738974	398
0.761795	334
0.784615	326
0.807436	314
0.830256	262
0.853077	336
0.875897	340
0.898718	327
0.921538	379
0.944359	378
0.967179	339
0.99	325



Here,

- The step count fluctuates drastically.
- Overall number of steps are increasing with the Discount rate.
- Low Learning Rate causes a much slower convergence (High number of steps)
- Fluctuations can occur as the agent explores the environment. During exploration, the agent may take suboptimal actions, leading to increased steps to reach the goal. As it learns more about the environment, it may switch to exploitation, reducing the number of steps.