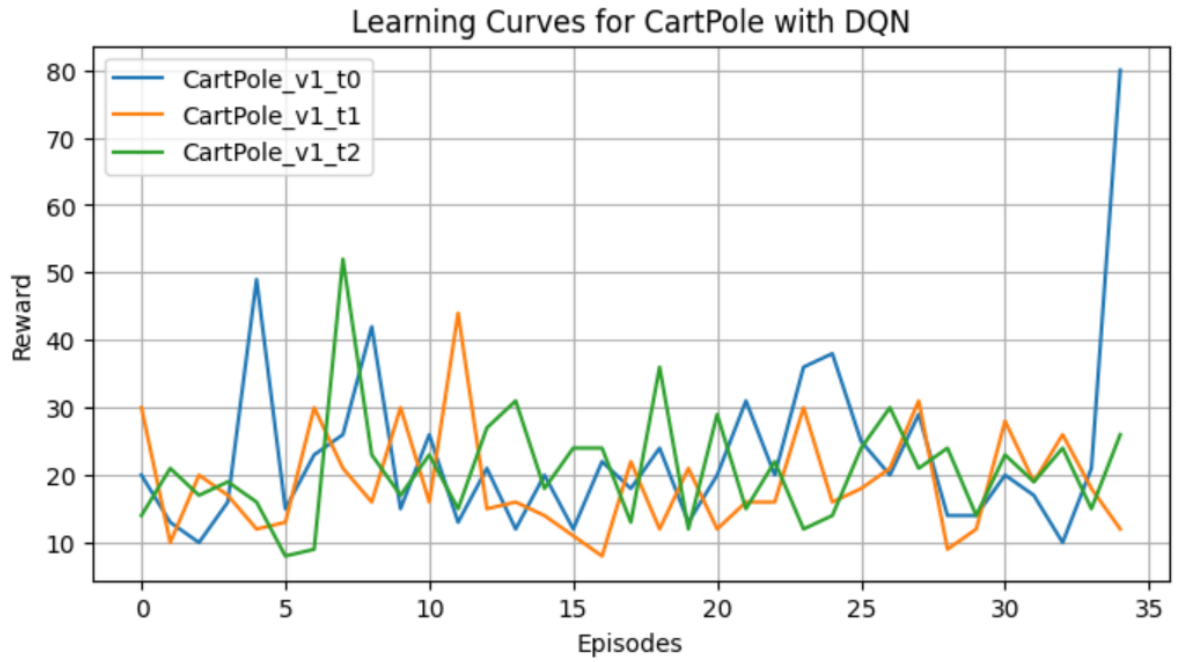
2232-CSE-6369-001
SPEC TOPS ADV INTELLIGENT SYS
Homework 2: Deep-Q-Network and Actor-Critic

STUDENT NAME: Naveyah Injam
STUDENT ID: 1002029985

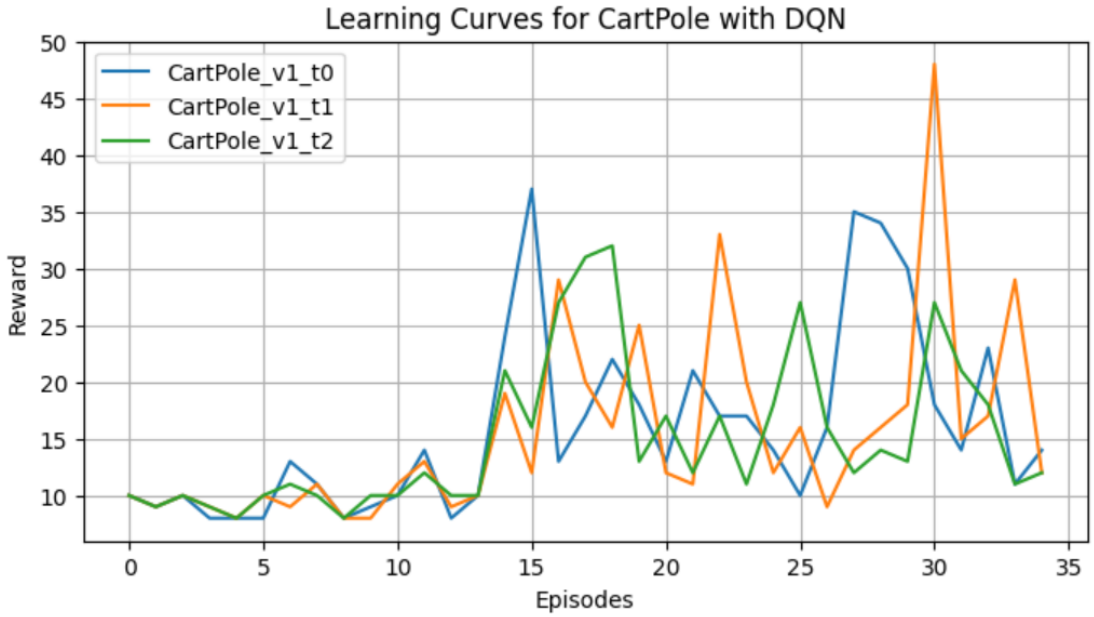
Experiment I (CartPole with DQN)

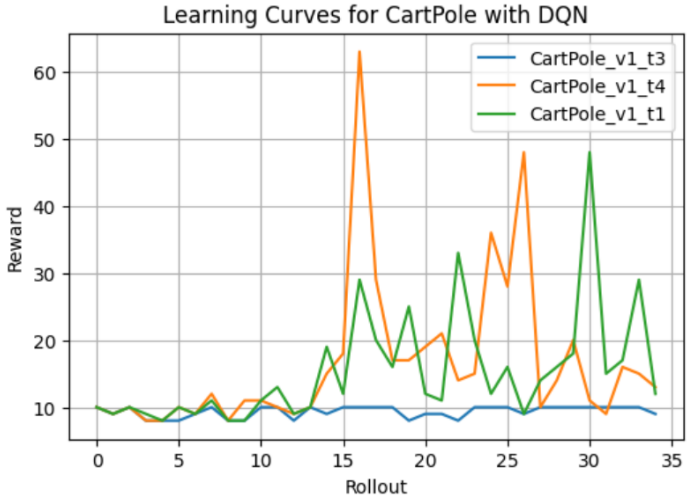
Deliverables for report:

1. Create a graph that compares the learning curve from the four trials above. Label the curves as t0, t1, t2. Report optimum value as τopt



2. Create a graph that compares the learning curve from the four trials above. Label the curves as t0, t1, t2. Report observed optimum values as ϵ init opt , ϵ min opt .





3. Answer the following questions

(a) How does changing target network update rate affect the learning curve? Can you justify your observation?

If the update rate is high target network updates more frequently else if it's low then the target network updates less frequently. However changing target network update rate affecting learning curve depends on the environment i.e. Cartpole. As we increase target network update rate, the learning of tasks is bit faster which means faster reaction to the changes happening in the environment. Higher updates may lead to poor performance while lower updates results in slow learning but gives better performance overall.

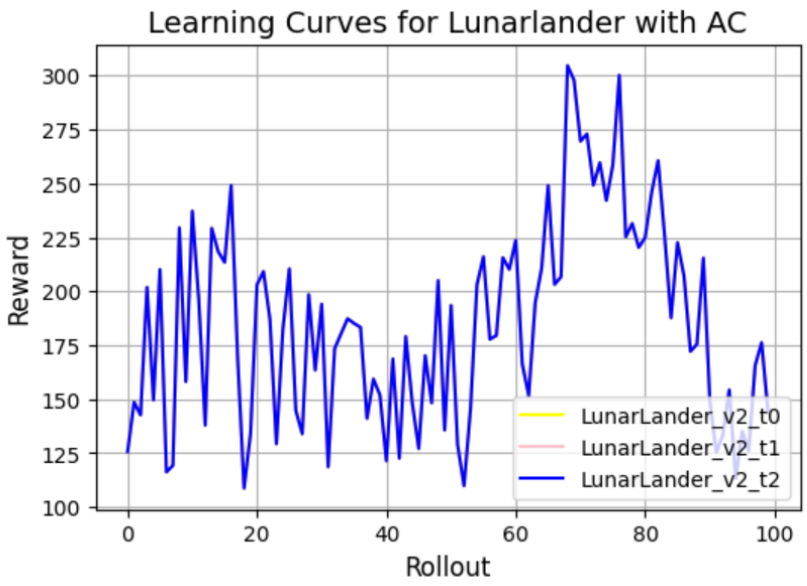
(b) How does changing range for ϵ affect the learning curve? Can you justify your observation?

Changing the range for ϵ affects learning curve that is having higher range shows that there is more exploration which helps agent learn new tasks faster and improves performance

Experiment II (LunarLander with AC)

Deliverables for report:

1. Create a graph that compares the learning curve from the three trials above. Label the curves as t0, t1, t2.



2. Answer the following questions
(a) How does changing the critic network update parameters (number of iterations and number of epochs) affect learning performance? How can you justify this relationship?

Changing the critic network by increasing the number of iterations and epochs will gradually improve the learning performance. As we try to approximate the optimal Q-function it will help give us better policy decisions and give us higher rewards.
We can justify this by looking at the results of graph

GITHUB REPO LINK:
HW2