**QLearner and SARSA Learner**

1. **Setting up the Environment**

Created a torch environment by running the following command

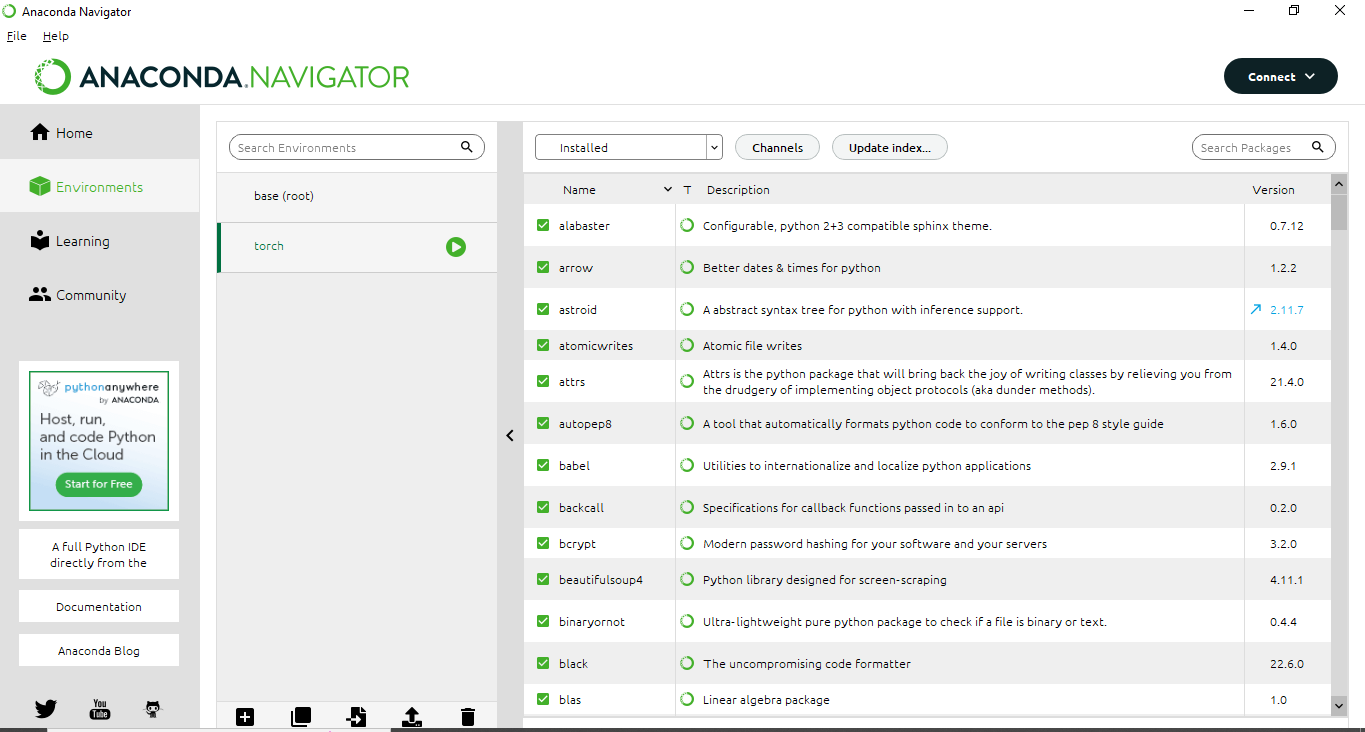
conda create -n torch python=3 pip



Environment torch is activated using the below command

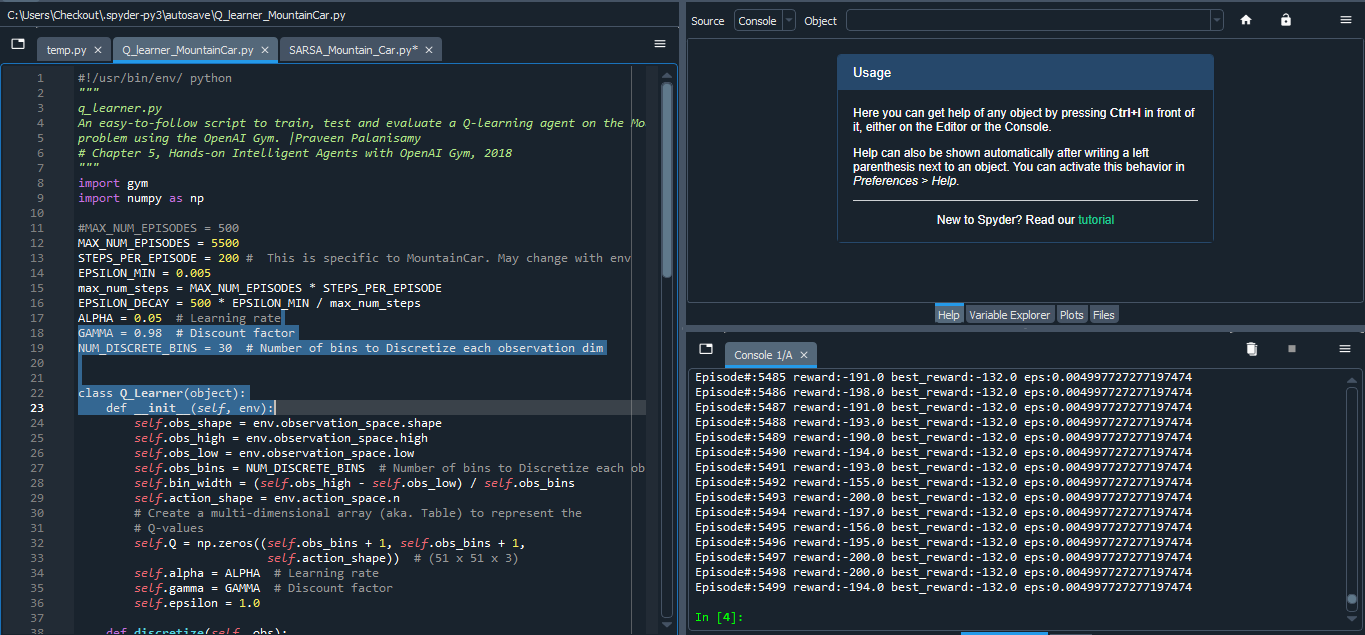
conda activate torch

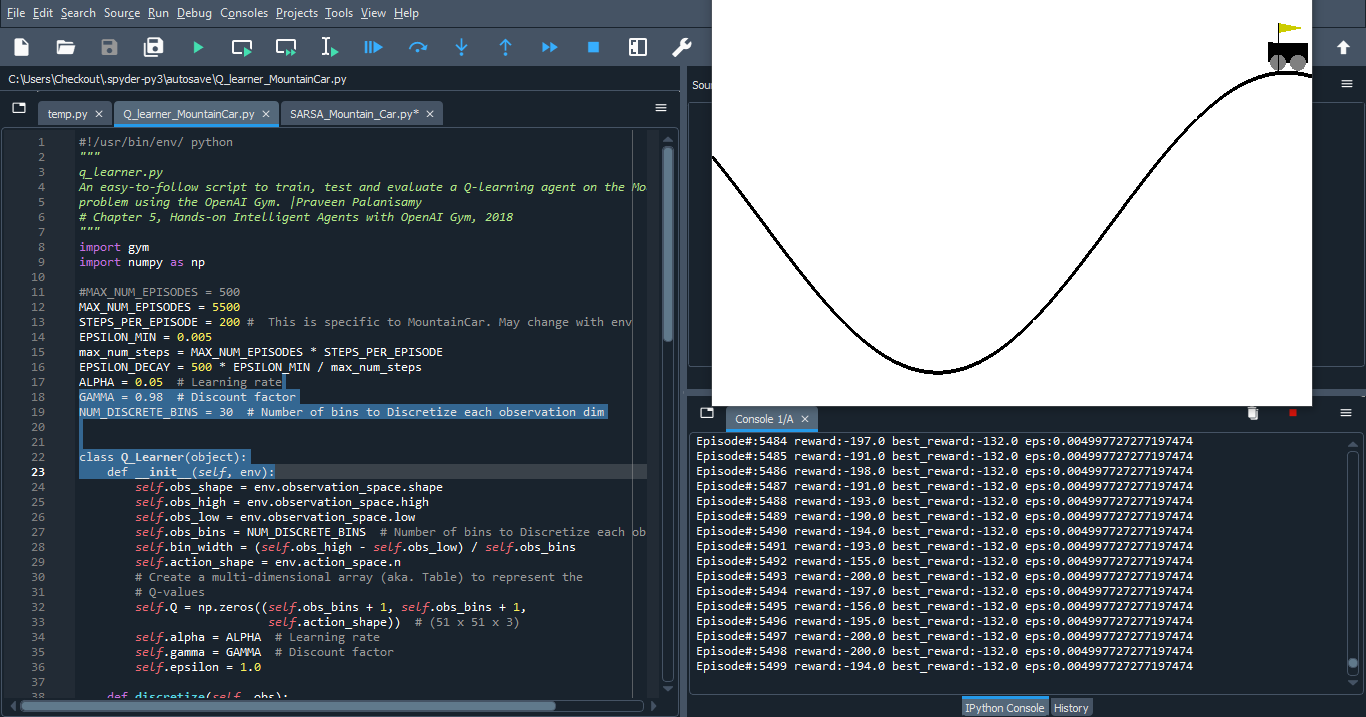
1. **Anaconda Navigator**



1. **QLearner**

**Output**





**Rewards**

Episode#:5492 reward:-155.0 best\_reward:-132.0 eps:0.004997727277197474

Episode#:5493 reward:-200.0 best\_reward:-132.0 eps:0.004997727277197474

Episode#:5494 reward:-197.0 best\_reward: -132.0 eps:0.004997727277197474

Episode#:5495 reward:-156.0 best\_reward:-132.0 eps:0.004997727277197474

Episode#:5496 reward:-195.0 best\_reward:-132.0 eps:0.004997727277197474

Episode#:5497 reward:-200.0 best\_reward: -132.0 eps:0.004997727277197474

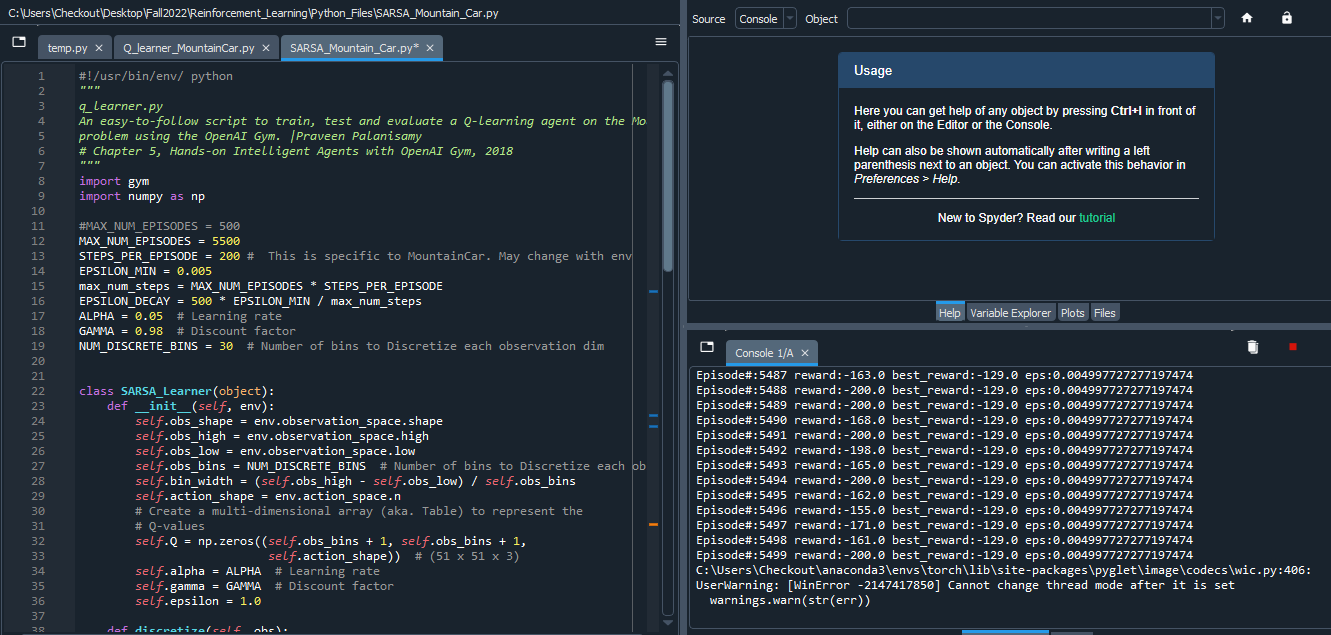
Episode#:5498 reward:-200.0 best\_reward:-132.0 eps:0.004997727277197474

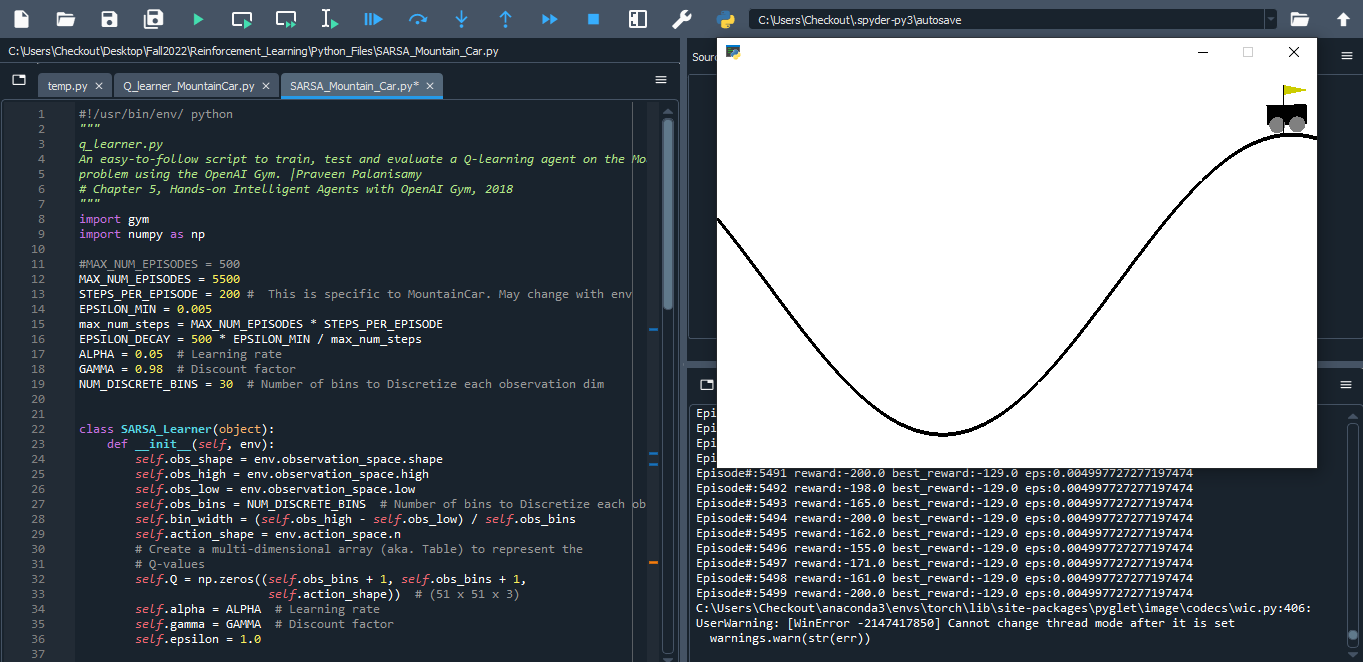
Episode#:5499 reward:-194.0 best\_reward:-132.0 eps:0.004997727277197474

**Best Reward is -132.0 with eps: 0.004997727277197474**

1. **SARSA Learner**

**Output**





**Reward**

Episode#:5493 reward:-165.0 best\_reward:-129.0 eps:0.004997727277197474

Episode#:5494 reward:-200.0 best\_reward:-129.0 eps:0.004997727277197474

Episode#:5495 reward:-162.0 best\_reward:-129.0 eps:0.004997727277197474

Episode#:5496 reward:-155.0 best\_reward:-129.0 eps:0.004997727277197474

Episode#:5497 reward:-171.0 best\_reward:-129.0 eps:0.004997727277197474

Episode#:5498 reward:-161.0 best\_reward:-129.0 eps:0.004997727277197474

Episode#:5499 reward:-200.0 best\_reward:-129.0 eps:0.004997727277197474

**Best Reward is -129.0 with eps: 0.004997727277197474**

**Conclusion**

With the same hyperparameters during training, SARSA and QLearner perform similarly. By comparing the two, it was possible to see that Qlearner offered a greater reward than SARSA.