HW5_Q-learning

Basic part

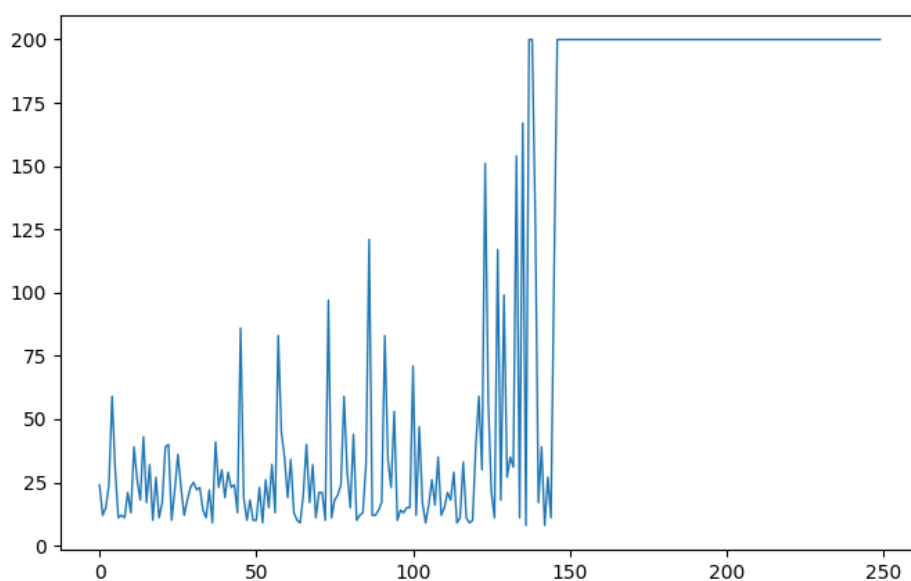**Describe how you implement the Q-learning algorithm.**

In basic part, we don't have to deal with too complex action. There were only few simple actions the car could do. Therefore, we just need to record the action and its reward, and we could find the best move!

Describe difficulties you encountered.

The main drawback of Q-learning is that the learning process is expensive for the agent, specially, in the beginning steps. Because every state-action pair should be visited frequently in order to converge to the optimal policy

Summarize your implementation.

For the purpose of convergence, we must apply greedy algorithm. By doing so, we could update learning rate and epsilon efficiently.

Advanced part:

Describe how you implement the Q-learning algorithm.

In the advanced part, we are dealing with a difficulty in Q-learning algorithm. We implemented it with neural network to solve the problem which requires a lot of space to record the action and rewards. Based on the neural networks, we don't have to use space efficiently.

Summarize your implementation.

There have two main different model. One is for forwarding, the other is for recording use. While we have a lot of experience, we will let the net learn. During the processing step, we want to speed up the convergence rate of training. Therefore, we updated the reward function to make the most balanced one get the highest reward scores. It leads the training process to converge rapidly.