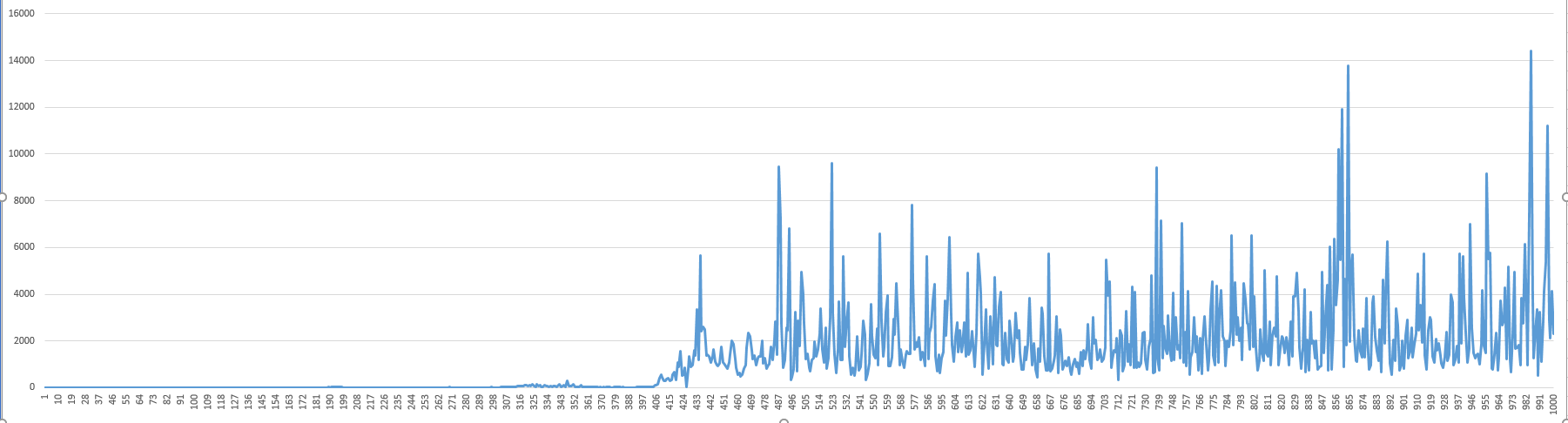
Lab 6 report

0412237 陳逸豐

⚫ Report (70%)

1.A plot shows episode rewards of at least 1000 training episodes (10%)



2.Describe your implement of network structure & Loss function (20%)

The DQN network is composed of two network which are Q target network and Q evaluation network. Q target network and Q evaluation network have two layer, both of them are fully connected layer. Loss function is mean square error between the output of Q target network and Q evaluation network.

3.Describe how you implement the training process of deep Q-learning (20%)

First, it will use epsilon-greedy action select method to choose the action, and apply the action on the Q evaluation network, get the reward, and store the state and the after state to do gradient descent learning to update parameter.

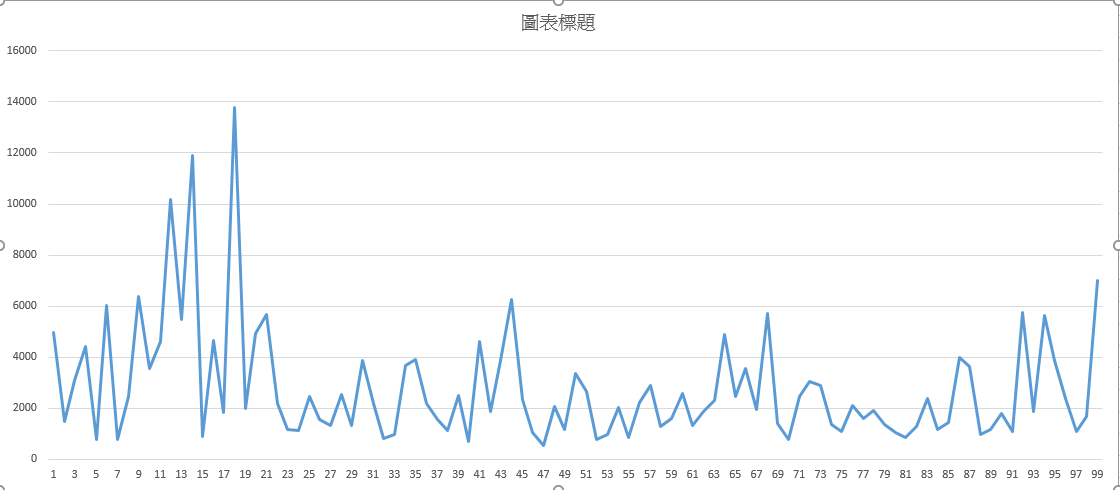
4.Describe the way you implement of epsilon-greedy action select method (10%)

There will be 90% to select the action which will get max reward and 10% to select the action randomly.

5.Describe how the code work (the whole code) (10%)

The code will play the game for 1000 episodes to learn how to play. During the episode, it will train the DQN network to get the best performance by using the training algorithm in the spec.

⚫ Performance (30%)

◼ Average reward during testing: Average ÷ 2.0

Avg:2769.673