COMP9444 Neural Networks and Deep Learning

Project 3 - Deep Reinforcement Learning

Team - 🤖🧐

Hayer Parameters

- **GAMMA**: The reward discount factor, the value should be range of (0,1), we choose 0.95 to do not decay the reward.
- **Initial epsilon**: The initial greedy rate, we use 0.9 to speed up our train.
- **Final epsilon**: The final greedy rate, we use 0.1 to stop the action when reach a high score.
- **Epsilon decay steps**: Greedy decay frequency, we just use 100 the same as TEST_FREQUENCY to test the result.

Neural Network

As for the neural network, We tried both [4,3,2] layers CNN network. Then we found the deeper network could lead to unstable result. Finally, we just use one layer CNN to train our behaviour and it is faster as well. The hidden nodes in the network is 100 nodes.

Batch

Obviously, the batch instance in our train can improve model performances. However, we have finish the first 100 episode in 1 min.Thus, we have find a balance in better performance and short time. After, we tried different values, we just found the 200 is suitable in our case.

Replay

The number of replay is the one of the major factor in our project. We tried [100, 1000, 5000, 10000] as our replay number, it shows that the more case included the better score we can get. 5000 can achieve 200 points in the first 100 episode in shortest time.