Project report

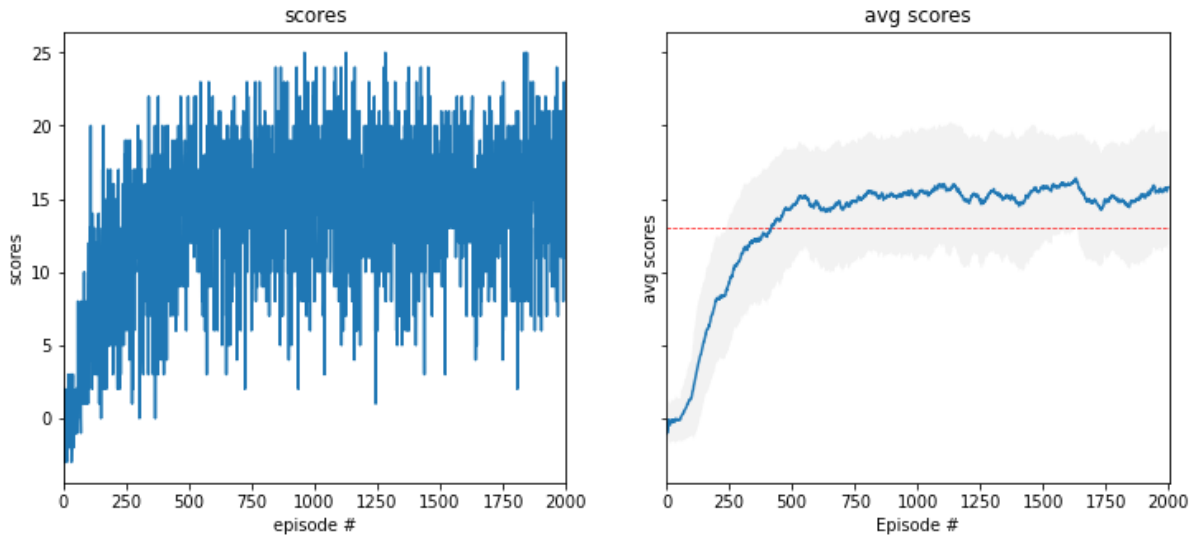
1. Results

In this project, Three TD learning algorithms(DQN, Double-DQN, Prioritized Experience Replay) were implemented and tested. The average test score of 5 episodes after training the model 2000 episodes resulted in **17.4, 18.6,** and **17.0** ! Below are the hyperparameter settings used.

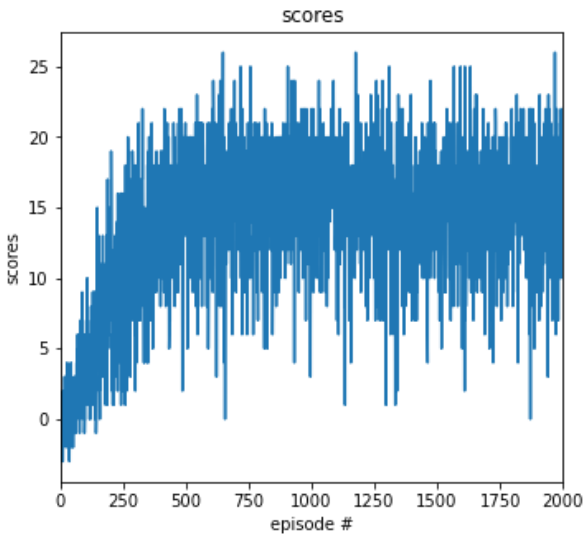
Hyper-parameters	• epsilon • epsilon decay • iterations • replay buffer size • minibatch size • discount factor • target parameters soft update TAU • learning rate • how often to update the network	: initial:1.0, min.:0.01, :**0.955** : 2000 : int(1e5) : 64 : 0.99 : **1e-3** : 5e-4 : 4
Model	• input • relu(fc1(input,128)) • relu(fc2(128, 64)) • fc3(64,4)	:37 states :128 units :64 units :4 action size

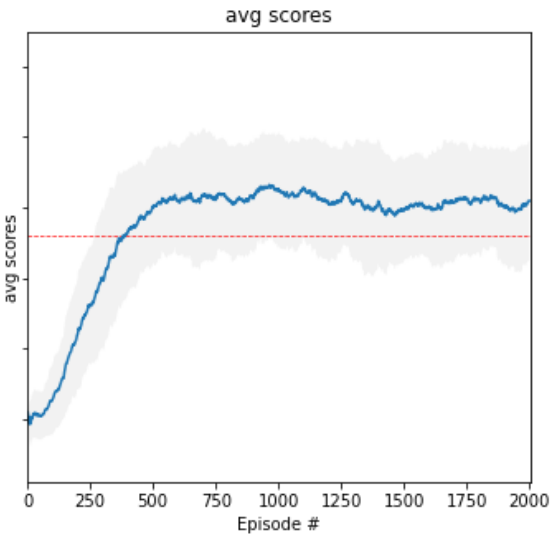
2. Rewards

DQN

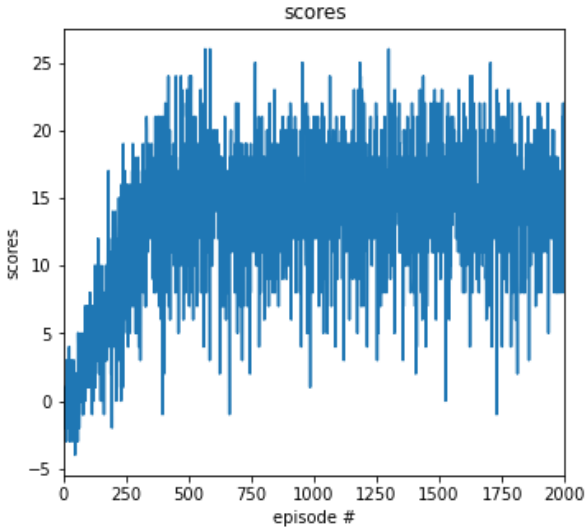


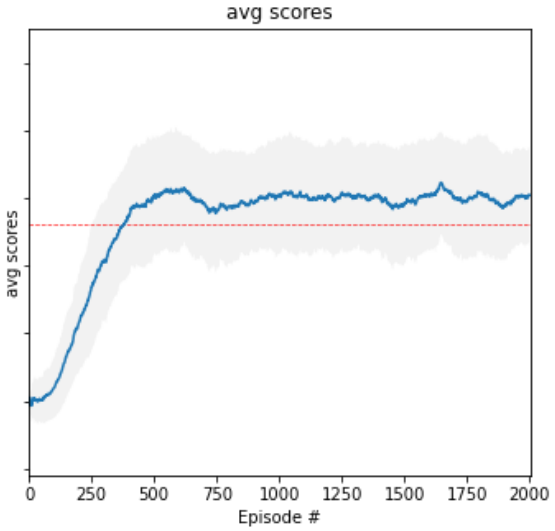
Double-DQN





Prioritized Experience Replay





Future Work

Even if all test results are over the threshold score of 11, Prioritized Experience Replay algorithm showed a similar test result without significant performance enhancement, which is not expected. The reason is, at this point, presumed to be the environment, which is too simple for the algorithm to make a difference. In the future, more tests will be followed for the analysis of behavior depending on different hyperparameter settings and environments.