Project report

1. Learning algorithm

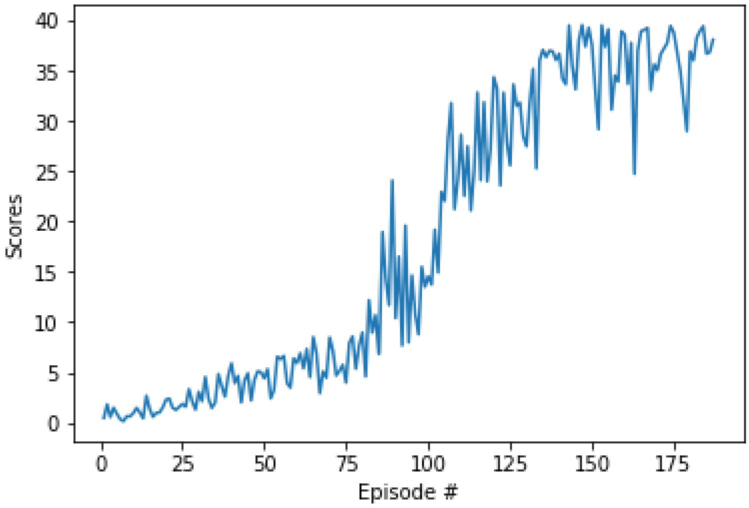
In this project, ddpg algorithm was implemented and tested for the single agent with version 1 environment.

2. DDPG Agent Hyper-parameters and models

Hyper-parameters	● BUFFER_SIZE = int(1e6) ● BATCH_SIZE = 128 ● GAMMA = 0.99 ● TAU = 1e-3 ● LR_ACTOR = 2e-4 ● LR_CRITIC = 2e-4 ● WEIGHT_DECAY = 0.0	
Model	Actor Network	Critic Network
	● 33 x 128 fc1 ● relu ● 128 x 128 fc2 ● relu ● 128 x 4 fc3 ● Tanh	● 33 x 128 fc1 ● relu ● (128+4) x 128 fc2 ● relu ● 128 x 1

3. Rewards

The average test score reached the target (above 30 over 100 episodes) at the 187th episode.



```
In [ ]:
```

4. Future Work

The ddqn algorithm with 20 Agents was tried multiple times but never exceeded 30 score over 100 episodes. The 20 Actor Critic ddpg agents shared the replay buffer with each agent stacking different experiences in the buffer in a sequential manner. And 20 Actor Critic agents shared neural network parameters such that each agent updates parameters in addition to the previously updated parameters by another agent.

It reached close to target score but kept becoming unstable at a certain point to drop significantly. Out of wits to improve the performance at this point, I leave more tests as a future work. The training with multiple agents was very slow even when gpu was used. In the future work, ddpg and regular actor critic n-step agents for multiple agents will be tested more out of the course.