

ML2021Spring HW12 Report

<電機工程研究所> <黃國郡>

<r08921005>

Public Score	Private Score
286	286

The methods I used to pass the strong baselines include:

1. Change the optimizer to **Adam**:
“`self.optimizer = optim.Adam(self.network.parameters(), lr=0.001)`”
2. Change **EPISODE_PER_BATCH = 20**
3. Change **NUM_BATCH = 2000** (total time is 16726 sec)
4. Collect **seq_rewards** in each **episode**
 - a. When each episode is done, then apply **discount_rate** as **0.99** to **seq_rewards**
 - b. `Rewards[]` extend the **seq_rewards** that after applying **discount_rate**.

(Your report should be written in English. Do not exceed 100 words describing your methods, but you may add comments to your code to make other students easier to understand.)