

ML2021Spring HW12 Report

<醫學工程所一> <卓嘉虹>

<r09528018>

Reward
253

The methods I used to pass the strong baselines include:

1. Implement discount accumulative reward
2. Number of batch=700
3. Break when average total reward exceed certain threshold