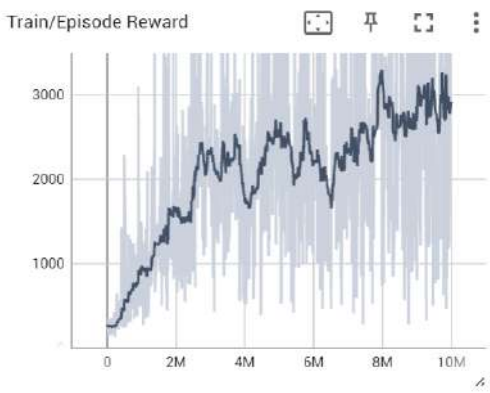
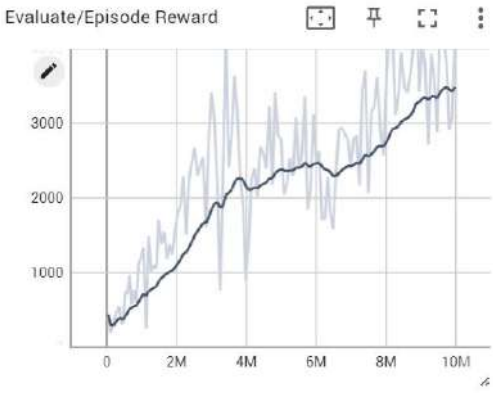
RL lab2

Experiment Result

DQN

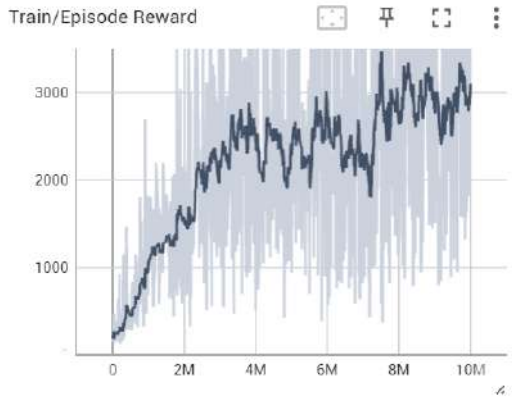


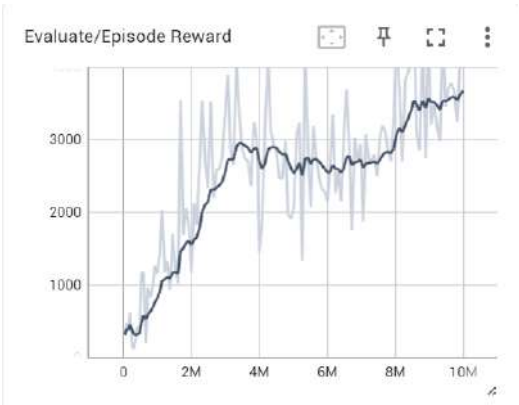



```
(RLenv) C:\Users\stanl\RL_s>python main.py
A.L.E: Arcade Learning Environment (version 0.8.1+53f58b7)
[Powered by Stella]
==========================================
Evaluating...
C:\Users\stanl\Envs\RLenv\lib\site-packages\gym\utils\passive_env_checker.py:233: DeprecationWarning: `np.bo
ol8` is a deprecated alias for `np.bool_`.  (Deprecated NumPy 1.24)
  if not isinstance(terminated, (bool, np.bool8)):
episode 1 reward: 4170.0
episode 2 reward: 3980.0
episode 3 reward: 4520.0
episode 4 reward: 3540.0
episode 5 reward: 4400.0
average score: 4122.0
==========================================
```

Experiment Result & Discussion (Bonus)

DDQN

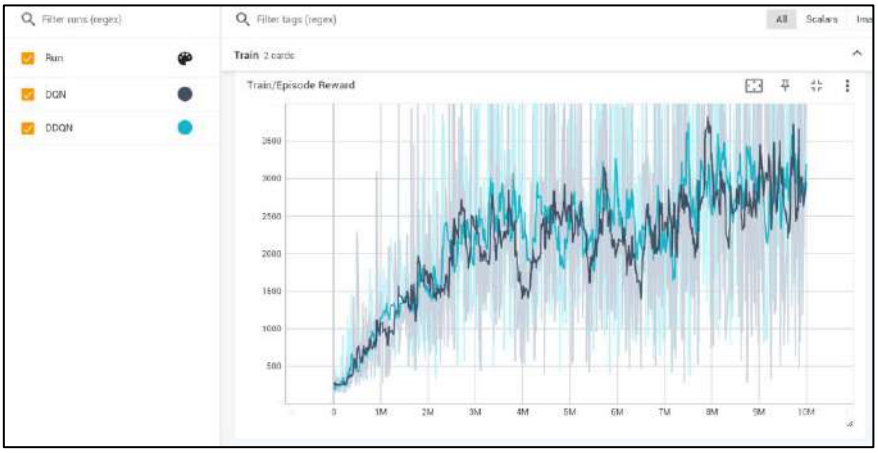


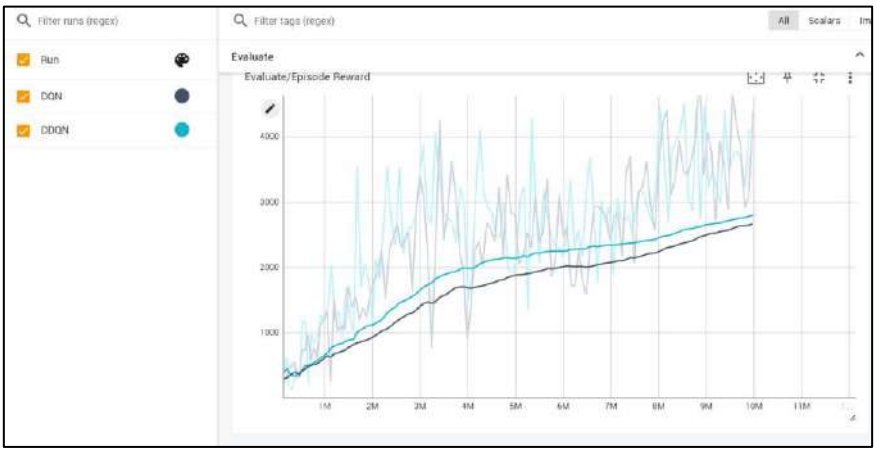


```
(RLenv) C:\Users\stanl\RL_s>python ddqn_main.py
A.L.E: Arcade Learning Environment (version 0.8.1+53f58b7)
[Powered by Stella]
==========================================
Evaluating...
C:\Users\stanl\Envs\RLenv\lib\site-packages\gym\utils\passive_env_checker.py:233: DeprecationWarning: `np.bo
ol8` is a deprecated alias for `np.bool_`.  (Deprecated NumPy 1.24)
  if not isinstance(terminated, (bool, np.bool8)):
episode 1 reward: 3510.0
episode 2 reward: 4160.0
episode 3 reward: 3130.0
episode 4 reward: 3430.0
episode 5 reward: 5020.0
average score: 3850.0
==========================================
```

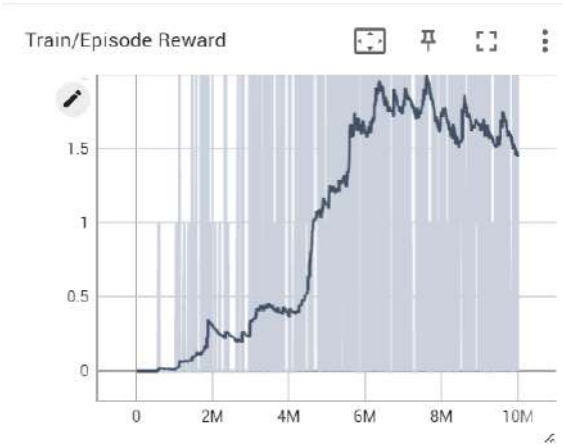
DDQN & DQN comparison

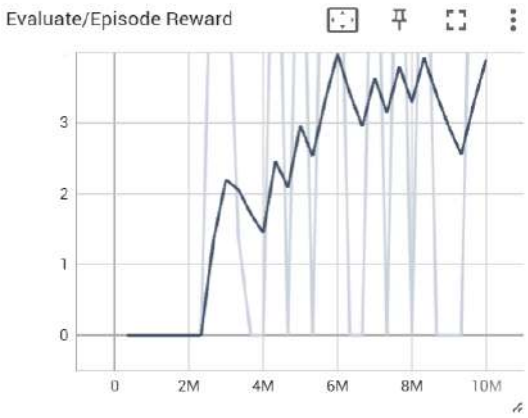
The training curves are similar, and the evaluate rewards are a little of DDQN is slightly better than DQN (I set the eval_epsilon to 0). My result show no obvious difference between DQN and DDQN because I only trained 10M for each algorithm, and it seems that the reward for DQN is still increasing. Maybe if I train to 20M for both algorithms, where the overestimating issue occurs in DQN, the reward of DQN model will start to increase at some point, and the DDQN model will increase steadily.





Enduro (DQN)






```
(RLenv) C:\Users\stanl\RL_s>python enduro_main.py
A.L.E: Arcade Learning Environment (version 0.8.1+53f58b7)
[Powered by Stella]
==========================================
Evaluating...
C:\Users\stanl\Envs\RLenv\lib\site-packages\gym\utils\passive_env_checker.py:233: DeprecationWarning: `np.bo
ol8` is a deprecated alias for `np.bool_`.  (Deprecated NumPy 1.24)
  if not isinstance(terminated, (bool, np.bool8)):
episode 1 reward: 27.0
episode 2 reward: 30.0
episode 3 reward: 25.0
episode 4 reward: 34.0
episode 5 reward: 38.0
average score: 30.8
==========================================
```