Name(s): Jennifer Lim, Yoonjae Hwang
Netid(s): jylim3, yoonjae2

Mean Reward Reached using DQN and DDQN:

DQN

```
episode: 1934   score: 7.0   memory length: 447845   epsilon: 0.3112649200096604    steps: 392    lr: 6.400000000000001e-06     evaluation reward: 6.73
episode: 1935   score: 14.0   memory length: 448377   epsilon: 0.3102115600096537    steps: 532    lr: 6.400000000000001e-06     evaluation reward: 6.82
episode: 1936   score: 5.0   memory length: 448701   epsilon: 0.30957004000964966    steps: 324    lr: 6.400000000000001e-06     evaluation reward: 6.82
episode: 1937   score: 3.0   memory length: 448913   epsilon: 0.309150280009647    steps: 212    lr: 6.400000000000001e-06     evaluation reward: 6.82
episode: 1938   score: 5.0   memory length: 449242   epsilon: 0.3084988600096429    steps: 329    lr: 6.400000000000001e-06     evaluation reward: 6.79
episode: 1939   score: 5.0   memory length: 449550   epsilon: 0.307889020009639    steps: 308    lr: 6.400000000000001e-06     evaluation reward: 6.81
episode: 1940   score: 9.0   memory length: 450026   epsilon: 0.30694654000963306    steps: 476    lr: 6.400000000000001e-06     evaluation reward: 6.83
episode: 1941   score: 9.0   memory length: 450464   epsilon: 0.3060793000096276    steps: 438    lr: 6.400000000000001e-06     evaluation reward: 6.88
episode: 1942   score: 5.0   memory length: 450758   epsilon: 0.3054971800096239    steps: 294    lr: 6.400000000000001e-06     evaluation reward: 6.9
episode: 1943   score: 8.0   memory length: 451203   epsilon: 0.3046160800096183    steps: 445    lr: 6.400000000000001e-06     evaluation reward: 6.93
episode: 1944   score: 8.0   memory length: 451639   epsilon: 0.30375280000961286    steps: 436    lr: 6.400000000000001e-06     evaluation reward: 6.92
episode: 1945   score: 6.0   memory length: 451995   epsilon: 0.3030479200096084    steps: 356    lr: 6.400000000000001e-06     evaluation reward: 6.92
episode: 1946   score: 5.0   memory length: 452287   epsilon: 0.30246976000960474    steps: 292    lr: 6.400000000000001e-06     evaluation reward: 6.89
episode: 1947   score: 10.0   memory length: 452795   epsilon: 0.3014639200095984    steps: 508    lr: 6.400000000000001e-06     evaluation reward: 6.92
episode: 1948   score: 5.0   memory length: 453125   epsilon: 0.30081052000959424    steps: 330    lr: 6.400000000000001e-06     evaluation reward: 6.92
episode: 1949   score: 7.0   memory length: 453487   epsilon: 0.3000937600095897    steps: 362    lr: 6.400000000000001e-06     evaluation reward: 6.94
episode: 1950   score: 5.0   memory length: 453812   epsilon: 0.29945026000958563    steps: 325    lr: 6.400000000000001e-06     evaluation reward: 6.92
episode: 1951   score: 9.0   memory length: 454283   epsilon: 0.2985176800957973    steps: 471    lr: 6.400000000000001e-06     evaluation reward: 6.96
episode: 1952   score: 5.0   memory length: 454573   epsilon: 0.2979434800095761    steps: 290    lr: 6.400000000000001e-06     evaluation reward: 6.92
episode: 1953   score: 5.0   memory length: 454877   epsilon: 0.297341560095723    steps: 304    lr: 6.400000000000001e-06     evaluation reward: 6.89
episode: 1954   score: 5.0   memory length: 455189   epsilon: 0.2967238000095684    steps: 312    lr: 6.400000000000001e-06     evaluation reward: 6.88
episode: 1955   score: 8.0   memory length: 455653   epsilon: 0.29580508000956257    steps: 464    lr: 6.400000000000001e-06     evaluation reward: 6.92
episode: 1956   score: 8.0   memory length: 456086   epsilon: 0.29494774000955715    steps: 433    lr: 6.400000000000001e-06     evaluation reward: 6.94
episode: 1957   score: 7.0   memory length: 456527   epsilon: 0.2940745600095516    steps: 441    lr: 6.400000000000001e-06     evaluation reward: 6.95
episode: 1958   score: 14.0   memory length: 457139   epsilon: 0.2928628000954396    steps: 612    lr: 6.400000000000001e-06     evaluation reward: 7.02
episode: 1959   score: 5.0   memory length: 457446   epsilon: 0.2922549400095401    steps: 307    lr: 6.400000000000001e-06     evaluation reward: 7.02
```

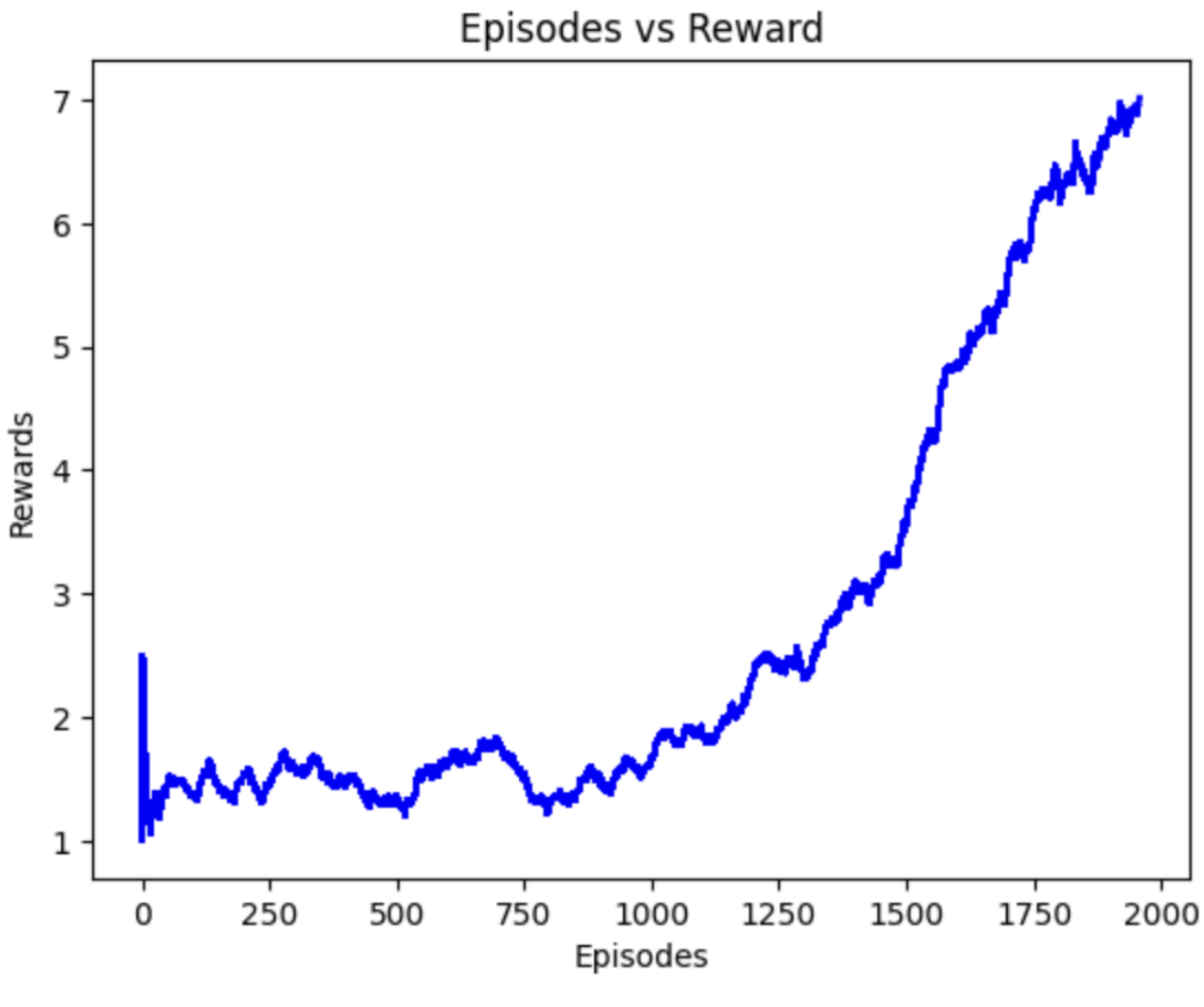
DDQN

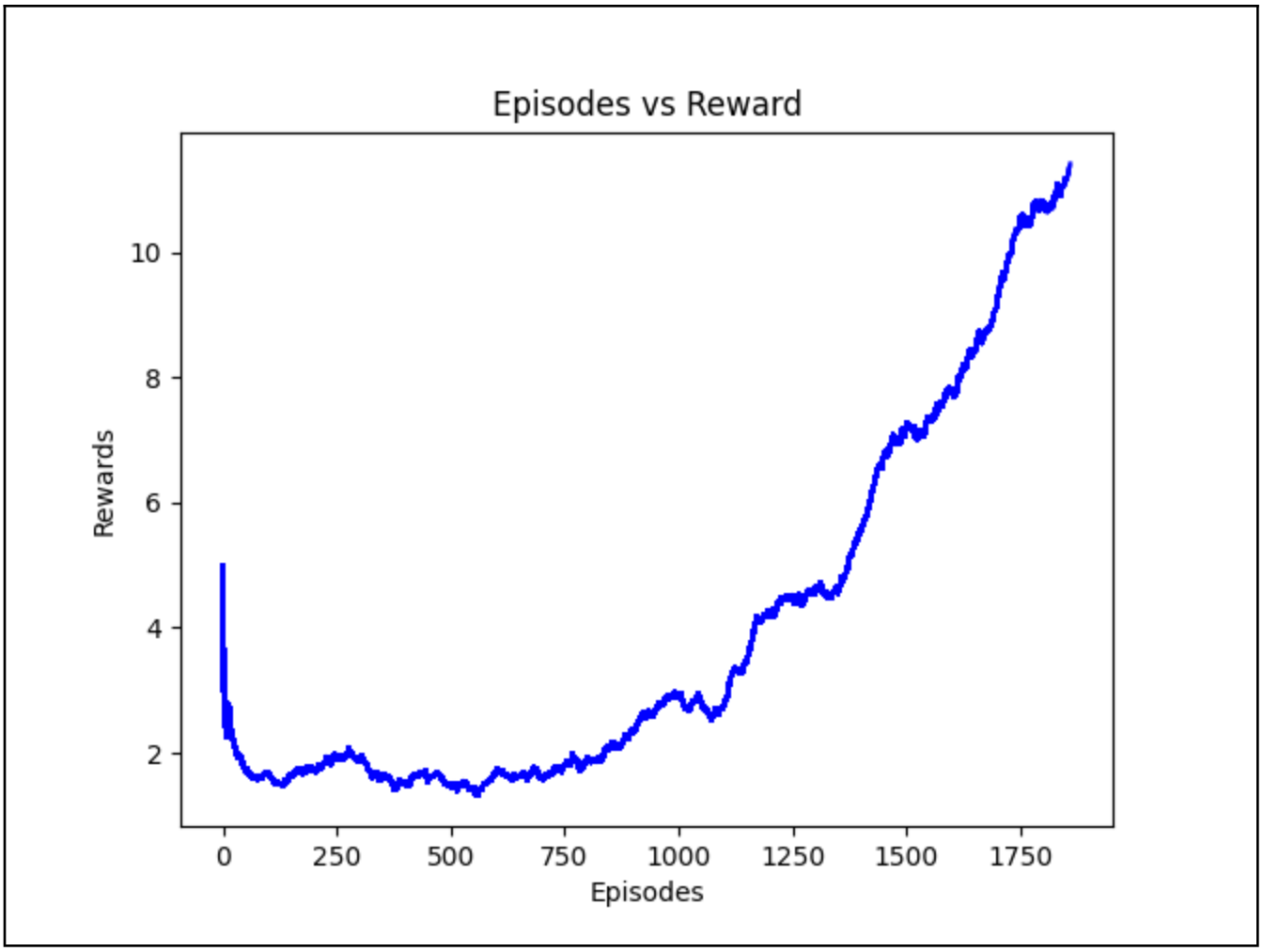
```
episode: 1822   score: 16.0   memory length: 500867   epsilon: 0.20628136000899616    steps: 601    lr: 2.5600000000000001e-06     evaluation reward: 10.78
episode: 1823   score: 18.0   memory length: 501485   epsilon: 0.20505772000898842    steps: 618    lr: 2.5600000000000001e-06     evaluation reward: 10.88
episode: 1824   score: 14.0   memory length: 502126   epsilon: 0.2037885400089804    steps: 641    lr: 2.5600000000000001e-06     evaluation reward: 10.87
episode: 1825   score: 12.0   memory length: 502686   epsilon: 0.20267974000897337    steps: 560    lr: 2.5600000000000001e-06     evaluation reward: 10.82
episode: 1826   score: 9.0   memory length: 503134   epsilon: 0.20179270000896776    steps: 448    lr: 2.5600000000000001e-06     evaluation reward: 10.87
episode: 1827   score: 13.0   memory length: 503729   epsilon: 0.2006146000089603    steps: 595    lr: 2.5600000000000001e-06     evaluation reward: 10.87
episode: 1828   score: 11.0   memory length: 504321   epsilon: 0.1994424400089529    steps: 592    lr: 2.5600000000000001e-06     evaluation reward: 10.9
episode: 1829   score: 15.0   memory length: 504866   epsilon: 0.19836334000894607    steps: 545    lr: 2.5600000000000001e-06     evaluation reward: 10.97
episode: 1830   score: 12.0   memory length: 505426   epsilon: 0.19725454000893905    steps: 560    lr: 2.5600000000000001e-06     evaluation reward: 11.02
episode: 1831   score: 18.0   memory length: 505973   epsilon: 0.1961714800089322    steps: 547    lr: 2.5600000000000001e-06     evaluation reward: 11.1
episode: 1832   score: 14.0   memory length: 506606   epsilon: 0.19491814000892427    steps: 633    lr: 2.5600000000000001e-06     evaluation reward: 11.08
episode: 1833   score: 6.0   memory length: 506964   epsilon: 0.19420930000891978    steps: 358    lr: 2.5600000000000001e-06     evaluation reward: 10.98
episode: 1834   score: 11.0   memory length: 507427   epsilon: 0.19329256000891398    steps: 463    lr: 2.5600000000000001e-06     evaluation reward: 10.94
episode: 1835   score: 13.0   memory length: 508093   epsilon: 0.19197388000890564    steps: 666    lr: 2.5600000000000001e-06     evaluation reward: 10.95
episode: 1836   score: 7.0   memory length: 508439   epsilon: 0.1912888000089013    steps: 346    lr: 2.5600000000000001e-06     evaluation reward: 10.94
episode: 1837   score: 13.0   memory length: 509045   epsilon: 0.1900889200088937    steps: 606    lr: 2.5600000000000001e-06     evaluation reward: 10.95
episode: 1838   score: 7.0   memory length: 509454   epsilon: 0.1892791000088886    steps: 409    lr: 2.5600000000000001e-06     evaluation reward: 10.93
episode: 1839   score: 14.0   memory length: 510130   epsilon: 0.18794062000888012    steps: 676    lr: 2.5600000000000001e-06     evaluation reward: 10.9
episode: 1840   score: 12.0   memory length: 510663   epsilon: 0.18688528000887344    steps: 533    lr: 2.5600000000000001e-06     evaluation reward: 10.93
episode: 1841   score: 18.0   memory length: 511214   epsilon: 0.18579430000886654    steps: 551    lr: 2.5600000000000001e-06     evaluation reward: 11.02
episode: 1842   score: 8.0   memory length: 511626   epsilon: 0.18497854000886138    steps: 412    lr: 2.5600000000000001e-06     evaluation reward: 11.02
episode: 1843   score: 16.0   memory length: 512327   epsilon: 0.1835905600088526    steps: 701    lr: 2.5600000000000001e-06     evaluation reward: 11.07
episode: 1844   score: 7.0   memory length: 512670   epsilon: 0.1829114200088483    steps: 343    lr: 2.5600000000000001e-06     evaluation reward: 11.07
episode: 1845   score: 8.0   memory length: 513088   epsilon: 0.18208378000884307    steps: 418    lr: 2.5600000000000001e-06     evaluation reward: 11.05
episode: 1846   score: 13.0   memory length: 513739   epsilon: 0.1807948000088349    steps: 651    lr: 2.5600000000000001e-06     evaluation reward: 11.08
episode: 1847   score: 16.0   memory length: 514383   epsilon: 0.17951968000882684    steps: 644    lr: 2.5600000000000001e-06     evaluation reward: 11.08
episode: 1848   score: 15.0   memory length: 514929   epsilon: 0.17843860000882    steps: 546    lr: 2.5600000000000001e-06     evaluation reward: 11.15
episode: 1849   score: 16.0   memory length: 515526   epsilon: 0.17725654000881252    steps: 597    lr: 2.5600000000000001e-06     evaluation reward: 11.19
episode: 1850   score: 10.0   memory length: 515998   epsilon: 0.1763219800088066    steps: 472    lr: 2.5600000000000001e-06     evaluation reward: 11.18
episode: 1851   score: 16.0   memory length: 516724   epsilon: 0.17488450000879752    steps: 726    lr: 2.5600000000000001e-06     evaluation reward: 11.15
episode: 1852   score: 10.0   memory length: 517080   epsilon: 0.17417962000879306    steps: 356    lr: 2.5600000000000001e-06     evaluation reward: 11.16
episode: 1853   score: 9.0   memory length: 517544   epsilon: 0.17326090000878724    steps: 464    lr: 2.5600000000000001e-06     evaluation reward: 11.18
episode: 1854   score: 15.0   memory length: 518176   epsilon: 0.17200954000877933    steps: 632    lr: 2.5600000000000001e-06     evaluation reward: 11.21
episode: 1855   score: 17.0   memory length: 518748   epsilon: 0.17087698000877216    steps: 572    lr: 2.5600000000000001e-06     evaluation reward: 11.24
episode: 1856   score: 8.0   memory length: 519210   epsilon: 0.16996222000876637    steps: 462    lr: 2.5600000000000001e-06     evaluation reward: 11.27
episode: 1857   score: 12.0   memory length: 519730   epsilon: 0.16893262000875986    steps: 520    lr: 2.5600000000000001e-06     evaluation reward: 11.34
episode: 1858   score: 12.0   memory length: 520290   epsilon: 0.16782382000875284    steps: 560    lr: 2.5600000000000001e-06     evaluation reward: 11.35
episode: 1859   score: 11.0   memory length: 520848   epsilon: 0.16671898000874585    steps: 558    lr: 2.5600000000000001e-06     evaluation reward: 11.38
episode: 1860   score: 12.0   memory length: 521423   epsilon: 0.16558048000873865    steps: 575    lr: 2.5600000000000001e-06     evaluation reward: 11.42
```

Uploaded Saved DQN/DDQN Model on Canvas (whichever performs better) : Yes

Uploaded your Agent.py and Agent_double.py file on Canvas : Yes

Plot of Mean Evaluation Reward for the model that reaches the target score (Either DQN or DDQN):

DQN

DDQN



Provide a few sentences to analyze the training process and talk about some implementation details:

The training process of DQN iteratively updates the network system to approximation Q-values. Agents collects data during training and the target values are computed using separated target network with fixed weights and is also periodically updated. We implemented both the agent and double agent to improve the performance of the DQN. Finally, we could conclude that DQN implementation uses target network to stabilize the training and prevent overfitting.

Extra Credit

Answer the questions accordingly if you did the corresponding part. The questions are just prompts. You should elaborate a bit more if you can.

1. What games did you apply the extra credit to? How does it work?
2. What other algorithm did you use? Explain and cite all your sources. Any issues you got in training your new algorithm.