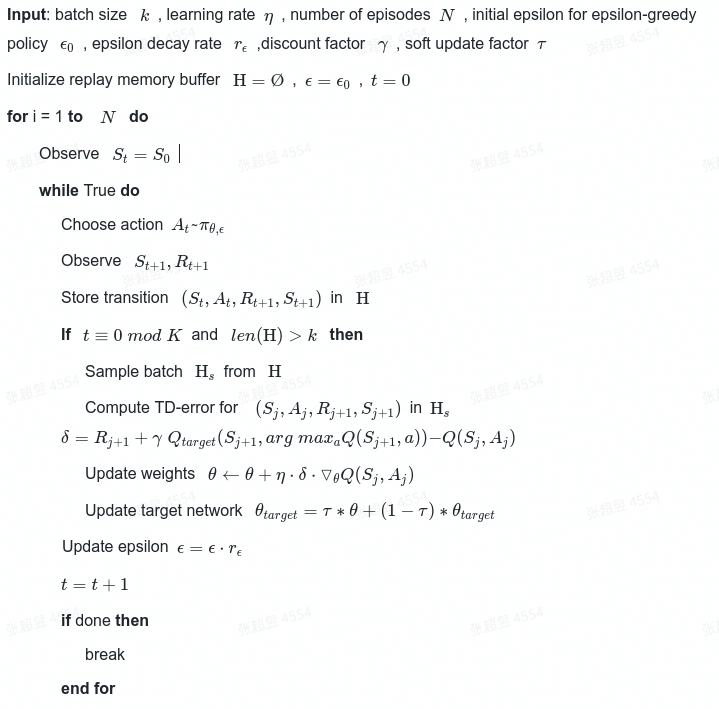
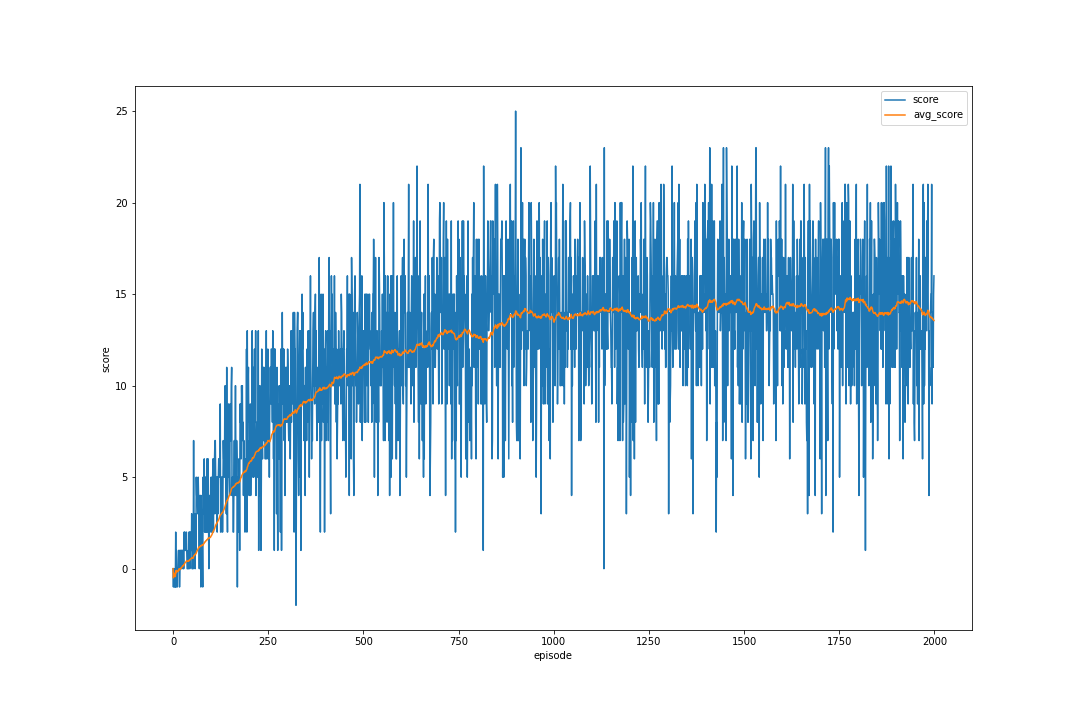
1. Learning Archetecture
   1. Algorithm

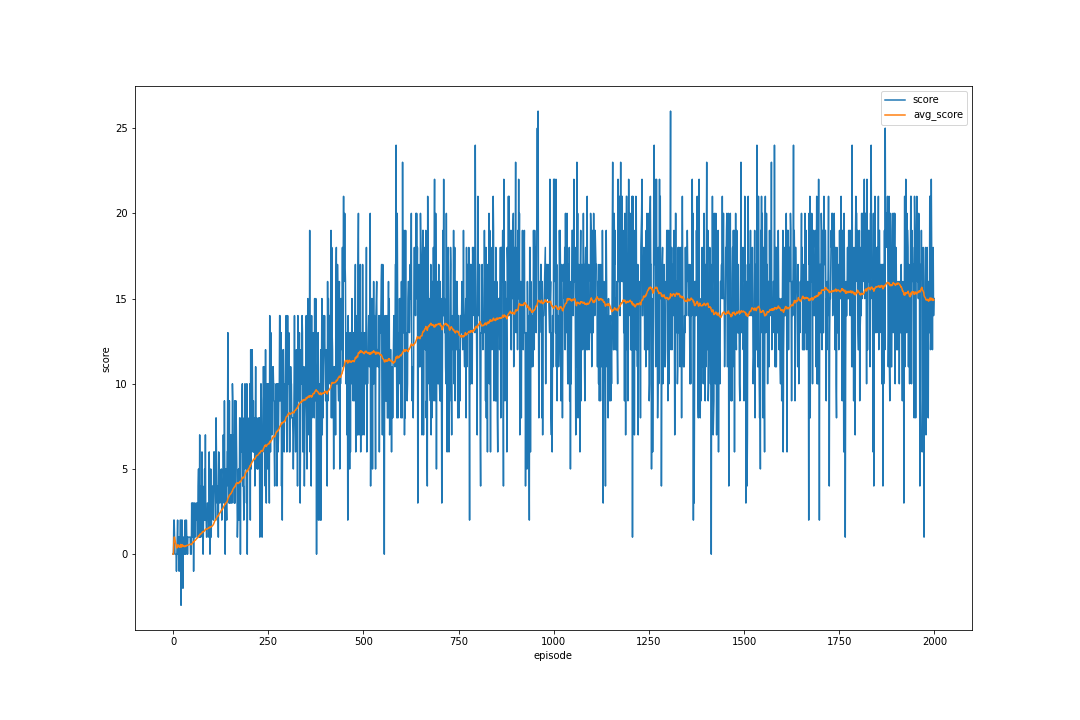


* 1. Model Structure
     1. Input size = state size = 37
     2. Hidden layers(2)
        1. Fully connected with 8\*37 rectifiers
        2. Fully connected with 8\*37 rectifiers
     3. Output layer of size 4(=action size)
  2. Hyperparameters
     1. Batch size 64
     2. Memory buffer size 1e5
     3. Number of episodes 2000
     4. Epsilon decay rate 0.995
     5. Target score 13.0
     6. Discount factor gamma 1e-3
     7. Learning rate 5e-4
     8. Update Period 4
     9. Sampling priority buffer 0.1
     10. SamplingWeightOrderIncreaseSpeed 1.0/1500.0
     11. SamplingPriorityOrderIncreaseSpeed 1.0/1500.0

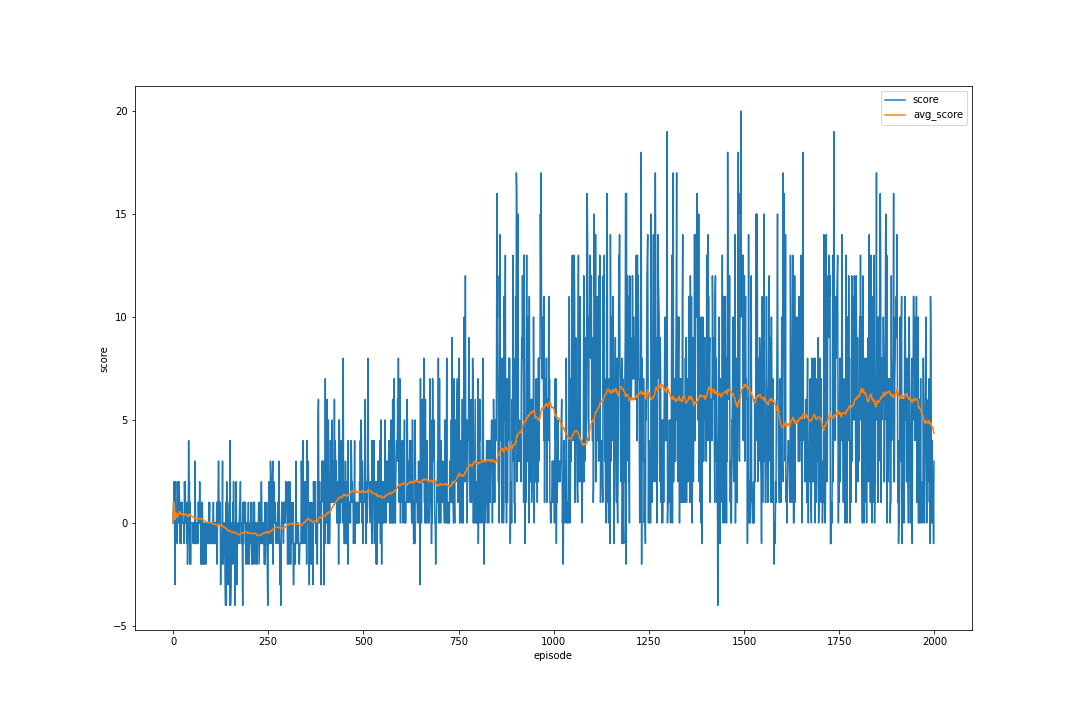
1. Results
   1. Original DQN
      1. Folder: dqn\_result\_2022\_04\_02\_14\_24\_08
      2. Result:



* 1. Double DQN
     1. Folder: dqn\_result\_2022\_04\_02\_16\_29\_25
     2. Result:



* 1. Double DQN + Prioritized Experience Replay
     1. Folder: dqn\_result\_2022\_04\_11\_12\_18\_33
     2. Result:



1. Conclusion
   1. Both original DQN and Double DQN can converge fast enough(within 1000 episodes)
   2. Comparing to original DQN, Double DQN reached a highier final score(14.92 vs 13.56)
   3. Prioritized Experience Learning can converge, but the final score is much lower than others(4.35), the reason maybe implementation error or inappropriate hyperparameters
2. Future Improvements
   1. Will try differrent hyperparameters by using google’s ML hypertun subsystem Vizier
   2. Will try to correct implementation of Prioritized Experience Replay
   3. Will try to implement Dueling DQN
   4. Will try learning from pixels
   5. Will try Adaptively Parametric ReLU