Parmeter Name:

S: current State, A: current Action, R: Reward, S'=next State

W⁻=weights of Separate target Network

Q_Target_Next Equation And Q_Target:

Q_target_next= maxq̂(S',A,W⁻)

Q_target=reward +(gamma*Q_target_next*(1-done))

My Loss function and Optimizer is

Loss function=(1/2)*(Q_Expected-Q_target)^2

Optimizer: Adam, Learning rate: 0.0005

Batch_size=64:

My Qnetwork consists of this.

Input Layer vector:37 , hidden1 Layer size: 64 ,hidden2 Layer size:64, output Layer size: 4

My Plot of Rewards

```
Episode 100     Average Score: 0.97
Episode 200     Average Score: 3.52
Episode 300     Average Score: 7.54
Episode 400     Average Score: 9.50
Episode 495     Average Score: 13.03
Environment solved in 395 episodes!     Average Score: 13.03
Score: 16.0
```