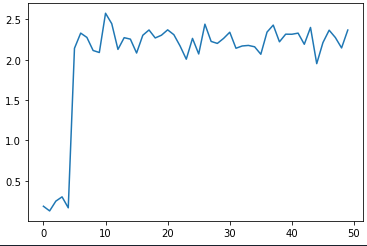
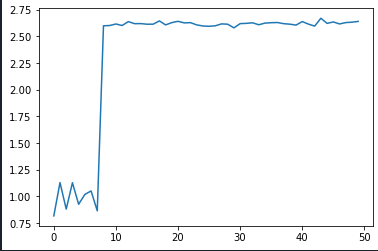
Soft update; MSE loss trade size 1, action space 9, AAPL dataset

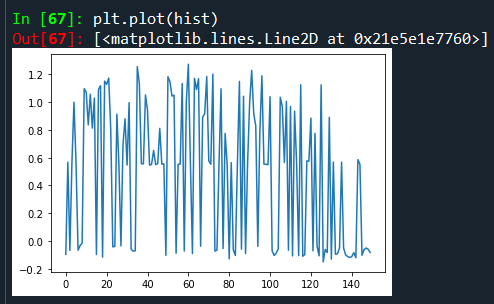
Evaluate per 10 episode, total 500 episodes

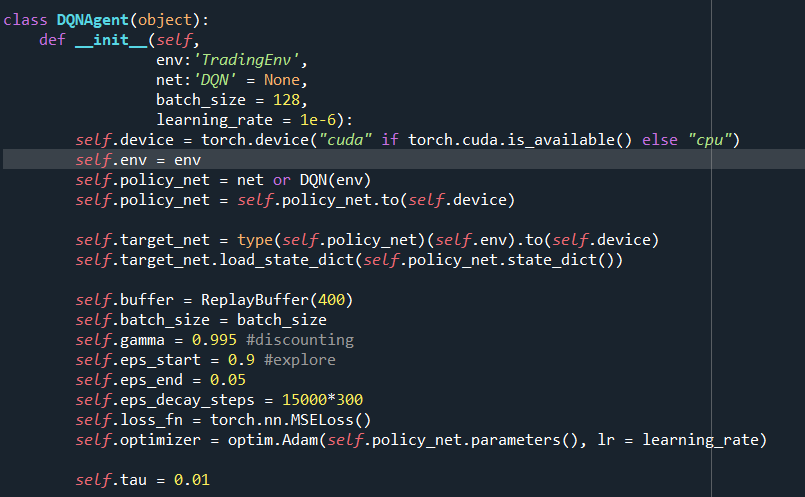






Trade size 0.1, 1500 episode, MSE loss, TA data set





2 看下loss是怎么回事 为什么这么小 也没收敛 Huber loss？

前期学的太快了导致进入了local minimum？还是用huber？

3. 为什么会产生负收益？检查一下

4. 算法太长视：调整gamma？

5. 把update那里的函数 for param in self.policy\_net.parameters():

param.grad.data.clamp\_(-1,1)

搞清楚什么意思