



Practical Deep Reinforcement Learning at Lower Risk for Stock Trading

3.1. MDP definition

Considering the stochastic and interactive nature of the trading market, model the stock trading process as a Markov Decision Process (MDP) (S, A, P, R, γ) .

- State $s = [p, h, b]$: a set that includes the information of the prices of stocks $p \in P_s^D$, the amount of holdings of stocks $h \in Z_s^D$, and the remaining balance $b \in R_s$, where D is the number of stocks that we consider in the market and Z_s denotes non-negative integer numbers.
- Action a : a set of actions on all D stocks. The available actions of each stock include selling, buying, and holding, which result in decreasing, increasing, and no change of the holdings h , respectively.
- Reward $r(s, a, s')$: the change of the portfolio value when action a is taken at state s and arriving at the new state s' . The portfolio value is the sum of the equities in all held stocks $p^T h$ and balance b .
- Policy $\pi(s)$: the trading strategy of stocks at state s . It is essentially the probability distribution of a state s .
- Action-value function $Q_\pi(s, a)$: the expected reward achieved by action a at state s following policy π .

Abstract:

Stock trading plays a crucial role in investments. However, it is challenging to obtain optimal strategy in the dynamic and complex stock market, and stock trading is also considered a high-risk investment finance activity. Reinforcement Learning can provide lot of technologies to do optimize at lower risks. In this paper, we explore different deep reinforcement learning based approaches to get optimal and lower risk stock trading strategy, and get approaches beat Dow Jones market index, which use 30 of the stocks historical data. The result shown our approach beat the market baseline and reduced the risk.

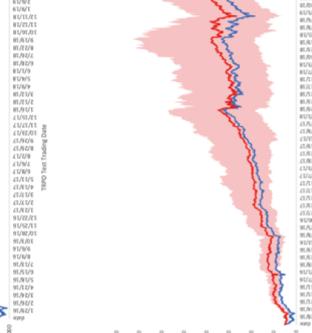
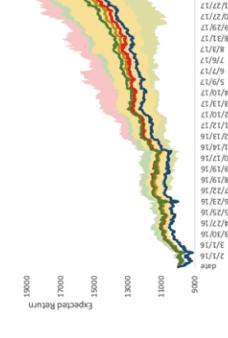
How much money we can earn?

Our start trading money is \$10000. Our baseline expected return, which hold \$10000 value of stock at the first trading day, will at market value \$15176.64, at the last trading day. While our favor approach DDPG Price Volume's average expected return is \$1756.57, which yield yearly return 23.2%, and is 6.9% better than DJI.

Table 1. Standard Deviation Analysis on Different Approaches

DRL ALG	STOCK FEATURE(S)	STD	MEAN
DOW JONES	INDEX	1516.64	1516.64
DDPG	PRICE ONLY	1844.63	1716.57
DDPG	PRICE W/ VOLUME	1796.35	17352.44
DDPG	PRICE W/ ALL TI	1565.71	17067.60
A2C	PRICE ONLY	2652.92	17277.21
A2C	PRICE W/ VOLUME	1952.76	16678.47
A2C	PRICE W/ ALL TI	3227.02	17584.00
TRPO	PRICE ONLY	1029.91	16819.74
TRPO	PRICE W/ VOLUME	1049.88	16823.93
TRPO	PRICE W/ ALL TI	1076.40	16899.82

Table 2. Different Variance Penalty on DDPG



Peter888@stanford.edu