Reinforcement Learning Methods for Stock Trading

Peisong Wang, Fengchi Yuan

Intelligent Transportation and Logistics Systems Lab, TBSI, Tsinghua University, China

ABSTRACT

• Apply reinforcement learning to stock trading

• To maximize the return of stock trading with a certain degree of risk

• DDPG, TD3, SAC, for stock trading.

• 30 stocks are used

• Results are compared with Dow Jones Industrial Average to show the potential of DRL for financial tasks.

MOTIVATION

·Stock trading is a muti-stage decision making problem, which may be suitable to use reinforcement learning to solve.

·Reinforcement learning has been applied to several financial research problems. [1]

METHOD

·Dataset: Dow Jones 30 stock information from Yahoo Finance

Training data period: 2009-01-01 to 2019-01-01

Trading data period: 2019-01-01 to 2021-10-31

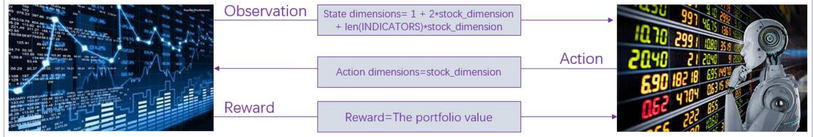
	date	open	high	low	close	volume	tic	day
0	2009-01-02	3.067143	3.251429	3.041429	2.762746	746015200	AAPL	4
1	2009-01-02	58.590000	59.080002	57.750000	44.219173	6547900	AMGN	4
2	2009-01-02	18.570000	19.520000	18.400000	15.418559	10955700	AXP	4
3	2009-01-02	42.799999	45.560001	42.779999	33.941101	7010200	BA	4
4	2009-01-02	44.910000	46.980000	44.709999	31.729940	7117200	CAT	4
...	...	...	...	...	...	...	...	...
94326	2021-10-29	454.410004	461.390015	453.059998	453.169403	2497800	UNH	4
94327	2021-10-29	209.210007	213.669998	208.539993	209.810745	14329800	V	4
94328	2021-10-29	52.500000	53.049999	52.410000	50.239872	17763200	VZ	4
94329	2021-10-29	46.860001	47.279999	46.770000	44.510624	4999000	WBA	4
94330	2021-10-29	147.910004	150.100006	147.559998	146.517624	7340900	WMT	4

94331 rows × 8 columns

·Preprocess the data(add some technical indicators)[2]

·A stock trading environment in OpenAI Gym-style

——Stock Dimension: 29, State Space: 291

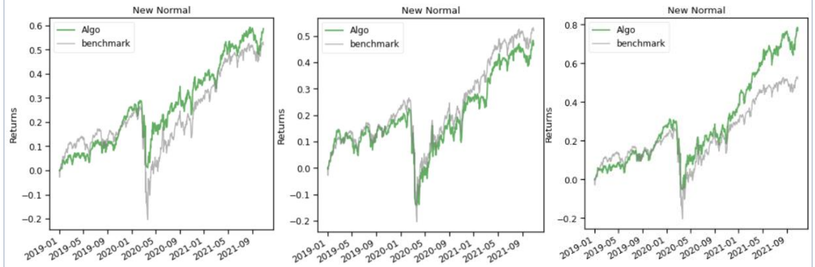


·Algorithms: DDPG, TD3, SAC from stablebaseline3[3]

RESULTS

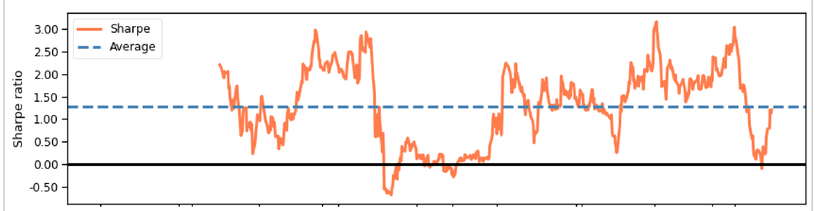
	Baseline(DJI)	DDPG	TD3	SAC
Annual return	15.979%	17.728%	14.628%	22.527%
Cumulative returns	52.01%	58.689%	47.149%	77.68%
Sharpe ratio	0.739	1.03	0.77	1.16
Max drawdown	-37.086%	-21.587%	-29.962%	-27.807%

Accumulative returns:

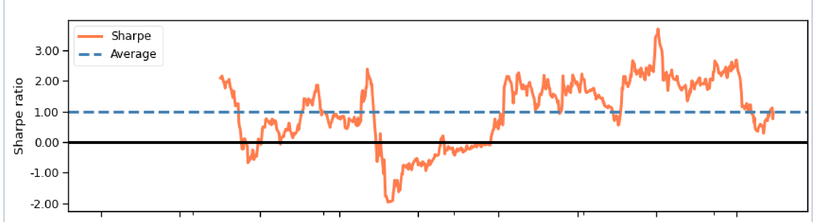


DDPG TD3 SAC

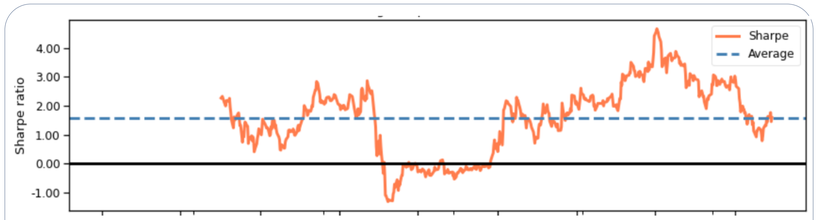
Rolling Sharpe ratio (6 months):



DDPG

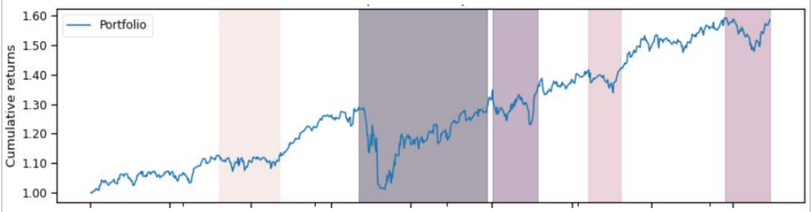


TD3

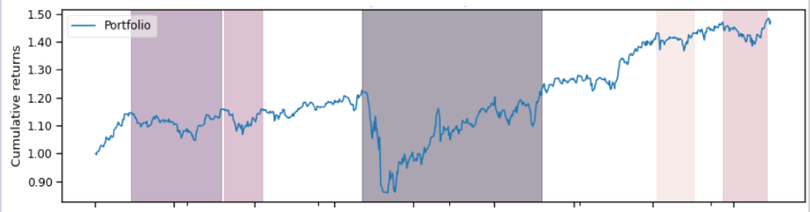


SAC

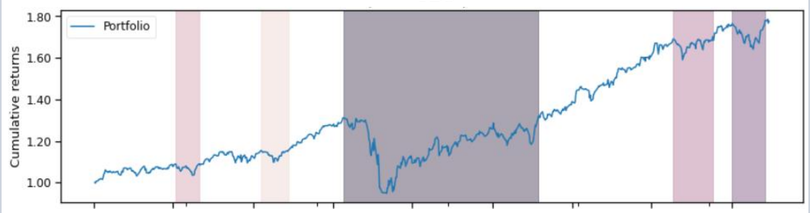
Top 5 drawdown periods:



DDPG



TD3



SAC

CONCLUSIONS

·We explored the potential of reinforcement learning agents (DDPG, TD3 and SAC) make decisions on stock trading.

·Annual return of DDPG and SAC agents outperforms the Dow Jones Industrial Average, while the TD3 fails to outperform.

·The Sharpe ratios' difference indicates that reinforcement learning agents are capable in balancing risk and return.

REFERENCES

[1] Zhuoran Xiong, Xiao-Yang Liu, Shan Zhong, Hongyang Yang and Anwar Walid. "Practical Deep Reinforcement Learning Approach for Stock Trading" Research Papers in Economics (2018): n. pag.

[2] http://finrl.org/

[3] https://stable-baselines3.readthedocs.io/en/master/