DRL for Stock Trading in Bangladesh

Supervised by: Ajmain Yasar Ahmed
Presented by: 1805091 & 1805093

Motivating Factors

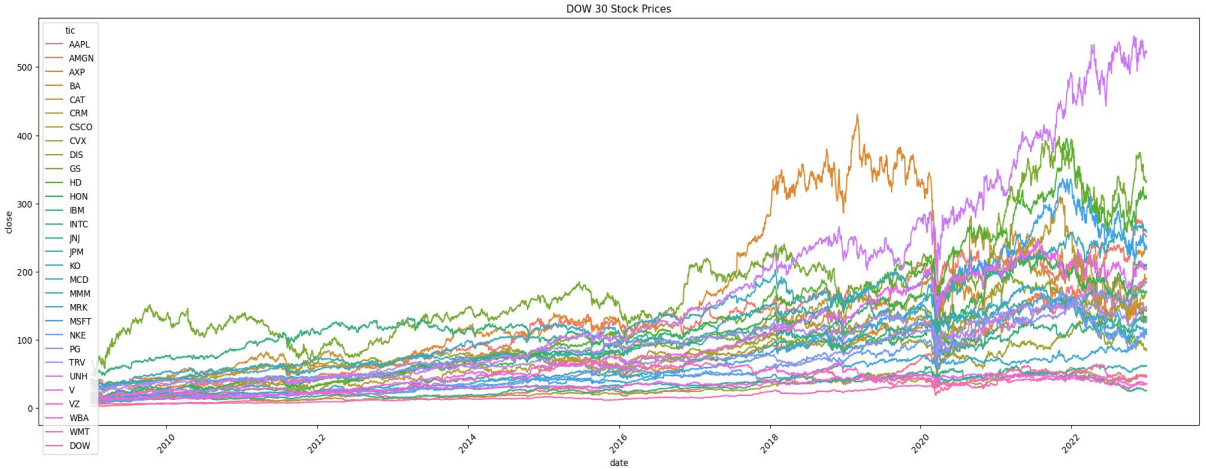
- Stock Exchange is a very unpredictable & turbulent sector. Predicting it and trading automatically is an interesting problem itself.
- Other Neural Network based models try to predict the future values of portfolios, not trade automatically.
- Reinforcement Learning based models would learn the ways of the stocks in a more intense way as it continuously trades and reinforces itself.
- Stock Market in Bangladesh is so much unstable comparing to other countries', it raises a curiosity to see how does Deep RL models work here.

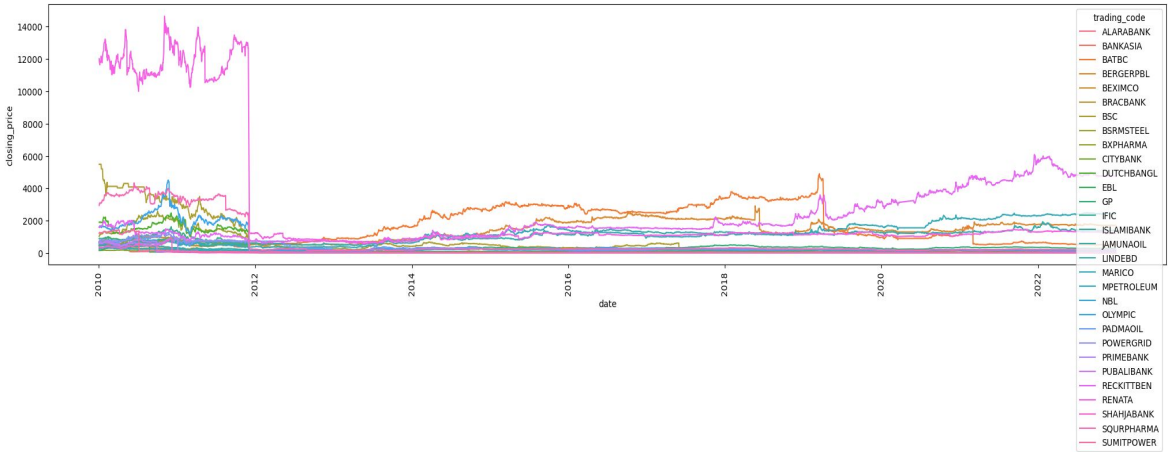
Overview

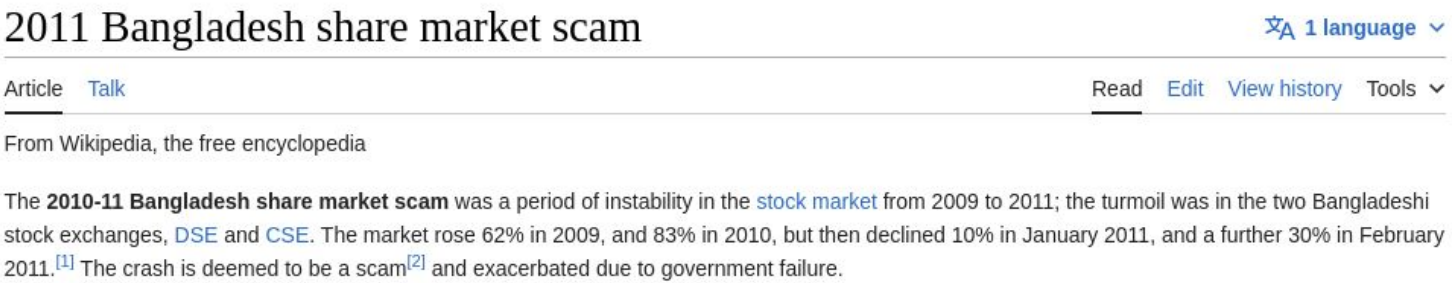
- Model adopted from FinRL
- Dataset taken from Kaggle

Trends in DJI & DSE

2009 to 2022







2011 Bangladesh share market scam

1 language

Article Talk Read Edit View history Tools

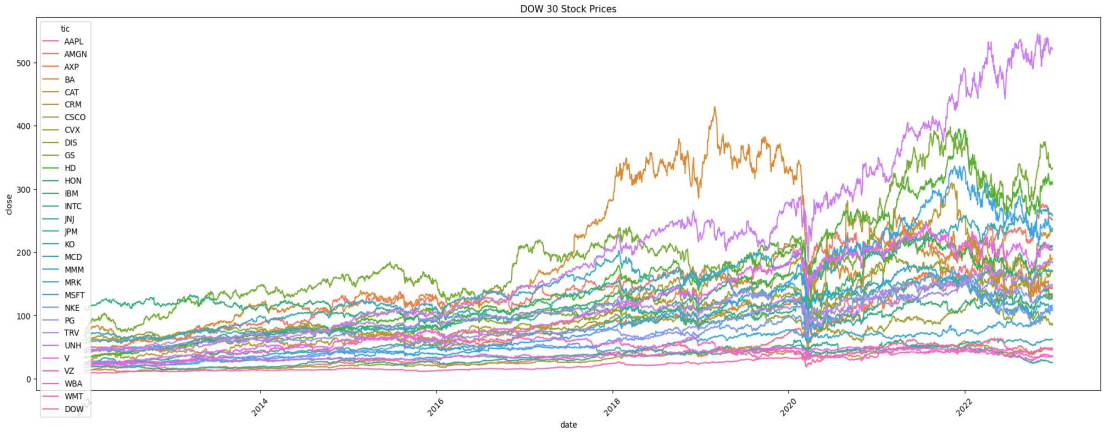
From Wikipedia, the free encyclopedia

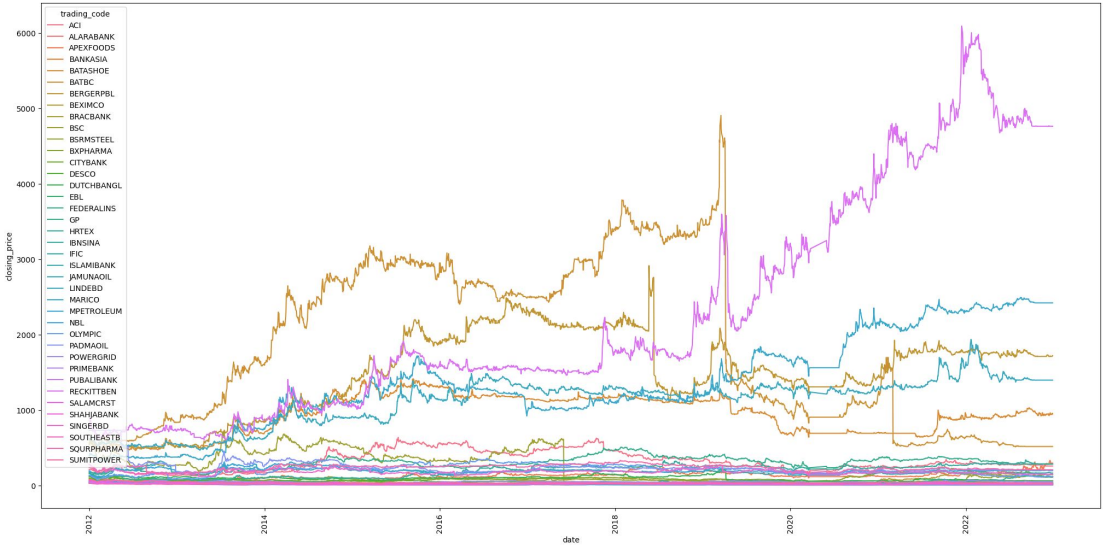
The **2010-11 Bangladesh share market scam** was a period of instability in the stock market from 2009 to 2011; the turmoil was in the two Bangladeshi stock exchanges, DSE and CSE. The market rose 62% in 2009, and 83% in 2010, but then declined 10% in January 2011, and a further 30% in February 2011.[1] The crash is deemed to be a scam[2] and exacerbated due to government failure.

Why DSE fluctuate too much around 2011?

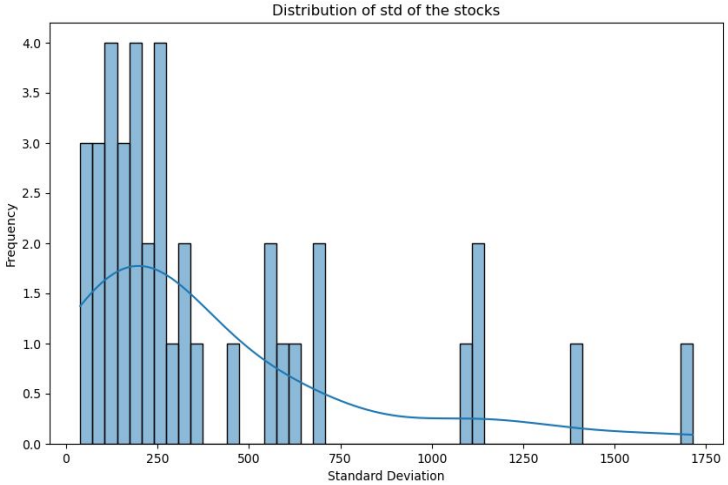
Trends in DJI & DSE

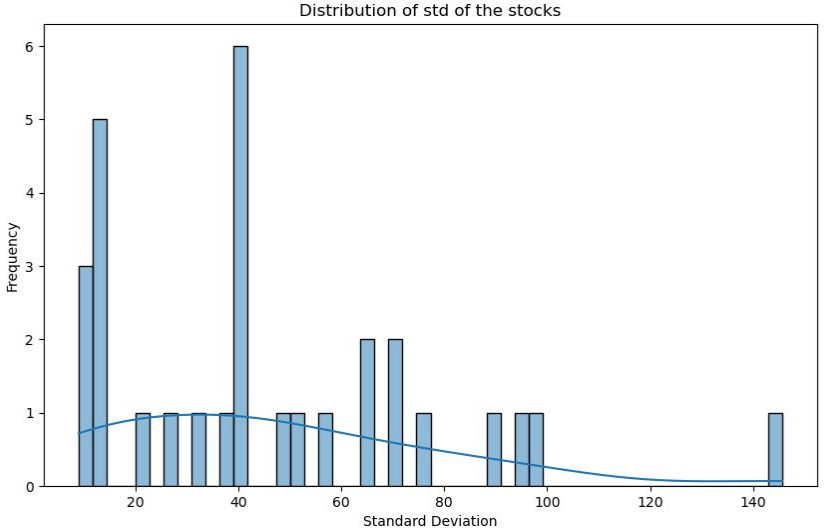
2012 to 2022





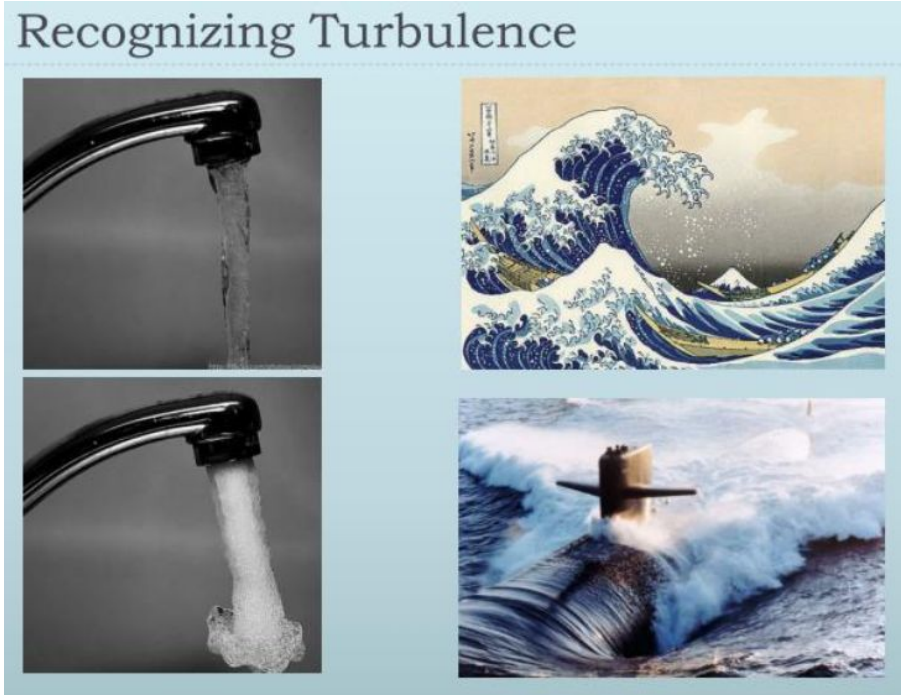
Distribution of Standard Deviation of DSE and DJI top companies



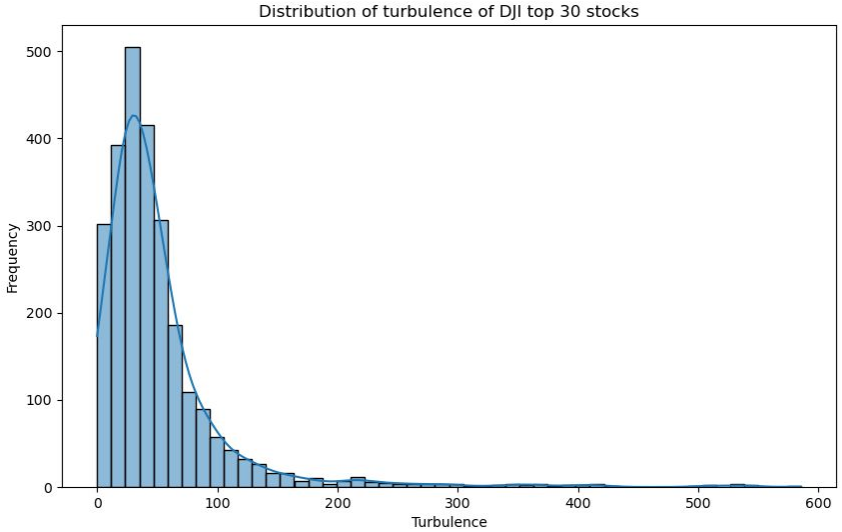


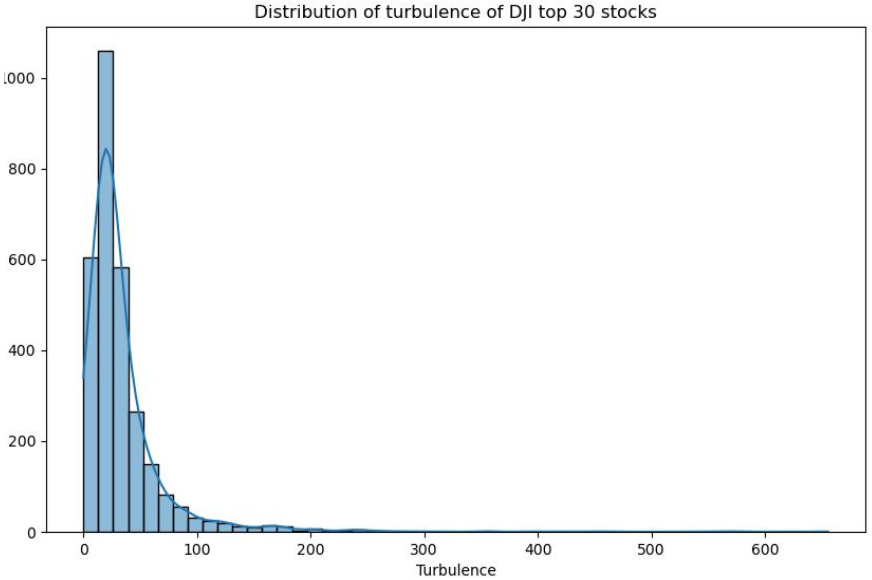
Turbulence

Market turbulence is the unexpected rising and falling of the stock market



Turbulence distribution in DSE and DJI





Volatility Index - VIX

- Measures insecurity in the stock market.
- When the VIX is high, people are not buying which is also called market fear.
- "If the VIX is high, it's time to buy"
- "When the VIX is low, look out below!"
- Buying when the VIX is high and selling when it is low is a strategy, but that needs to be considered against other factors and indicators too.

Stock Market Indicators:

MACD (Moving Average Convergence Divergence): Trend-following momentum indicator based on moving averages.

Bollinger Bands (Boll_ub and Boll_lb): Volatility bands indicating potential overbought or oversold conditions.

RSI (Relative Strength Index) 30: Momentum oscillator for identifying overbought or oversold conditions.

CCI (Commodity Channel Index) 30: Momentum oscillator for assessing trend strength and overbought/oversold conditions.

DX (Directional Movement Index) 30: Measures trend strength in a security.

Close_30_SMA and **Close_60_SMA**: Simple moving averages based on closing prices over 30 and 60 days respectively.

Models Used

- A2C [Advantage Actor Critic]
- DDPG [Deep Deterministic Policy Gradient]
- PPO [Proximal Policy Optimization]
- TD3 [Twin-Delayed Deep Deterministic Policy Gradient]
- SAC [Soft Actor-Critic]

Mean Variance Optimization - MVO

- Measuring an asset's risk against its likely return and investing based on that risk/return ratio.
- The analysis helps investors determine the biggest reward at a given level of risk or the least risk at a given level of return.
- The expected return is a probability expressing the estimated return of the investment in the security.
- If two different securities have the same expected return, but one has lower variance, the one with lower variance is preferred.
- Similarly, if two different securities have approximately the same variance, the one with the higher return is preferred.

Dow Jones Industrial Average - DJI

- Stock market index that tracks 30 large, publicly-owned companies
- Primarily focuses on the US Market
- Could be somewhat relatable to other markets too

What Models Don’t See

- Natural Calamities
- War
- Government Decisions
- Company Policies
- New Company in Market
- Acquisitions etc.

That’s why an expert stock trader is most of the times more valuable than an automated software.

Progress

- Regenerated original results.
- Successfully incorporated existing RL based models from FinRL in Dhaka Stock Exchange dataset.
- Designed reward function for better outcome.

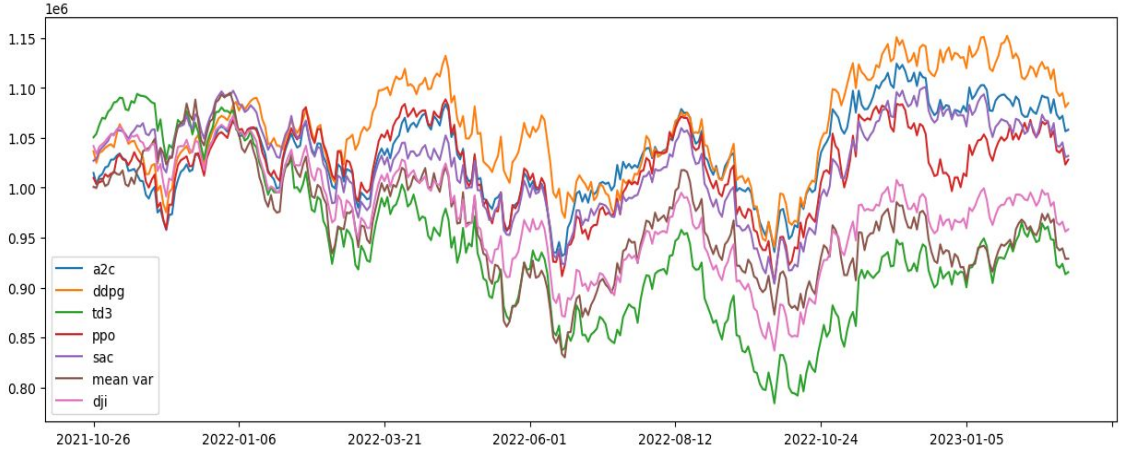
Results

Default Reward Function

- Only depends on the return at the end of the day.
- Reward Scaling controls the rate.

Reward = daily_return * reward_scalling

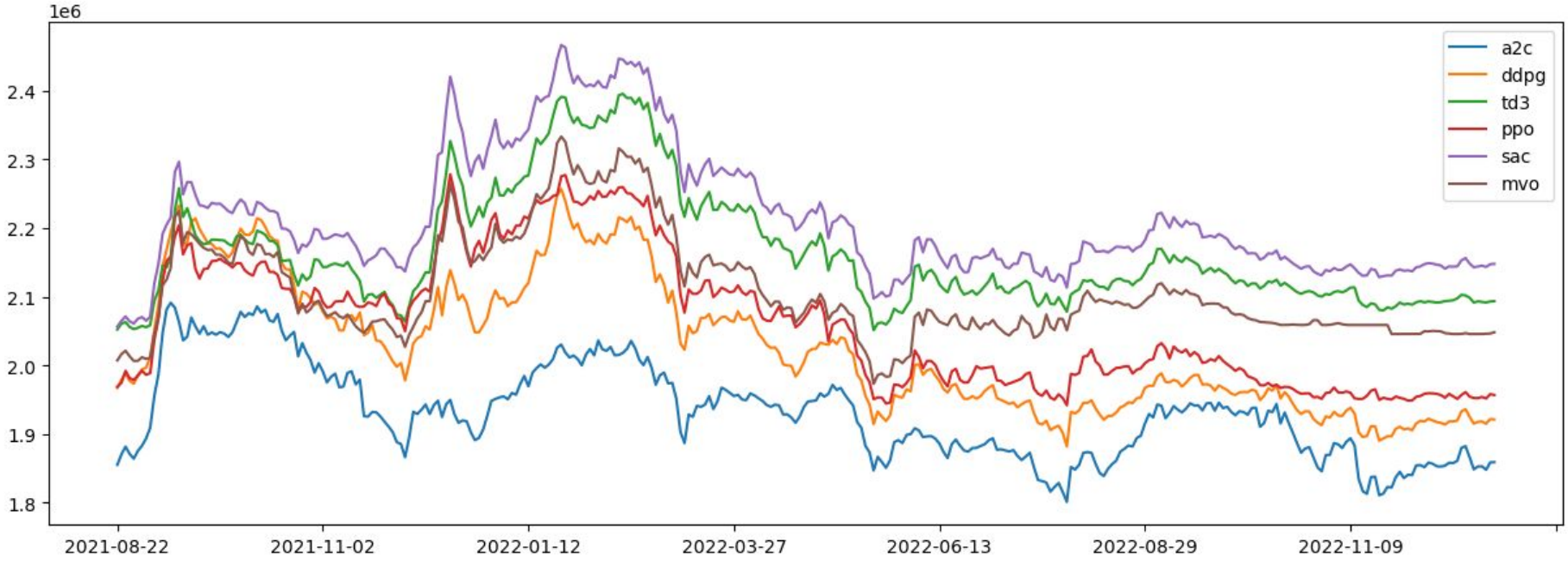
Dataset: DJIA top 29



	initial_value	final_value	annual_return	sharpe_ratio	max_drawdown
a2c	1014670.51891483	1058016.19458915	1.40415669825364	3.28099854532285E-05	-151957.461378052
ddpg	1036320.01684446	1084571.66802854	1.52853575693324	3.06601491965349E-05	-191816.784805298
td3	1050454.84450713	915325.953889763	-4.48619872746525	-5.90320513395877E-05	-309783.857719421
ppo	1009710.06648183	1027998.61852467	0.60014692299965 6	1.49151384884493E-05	-184124.06946183
sac	1026971.53194534	1031869.39014529	0.15872210292755 5	3.44759921219945E-06	-192925.196998211
mvo	1000730.08876205	928866.787238764	-2.45339499910148	-4.53984765280635E-05	-264967.362361946
dji	1041671.00043766	958126.440812903	-2.74825175463574	-5.22629744627238E-05	-235216.170014614

Train 2010-2020 Trade 2021-2022

Dataset: DSE top 39



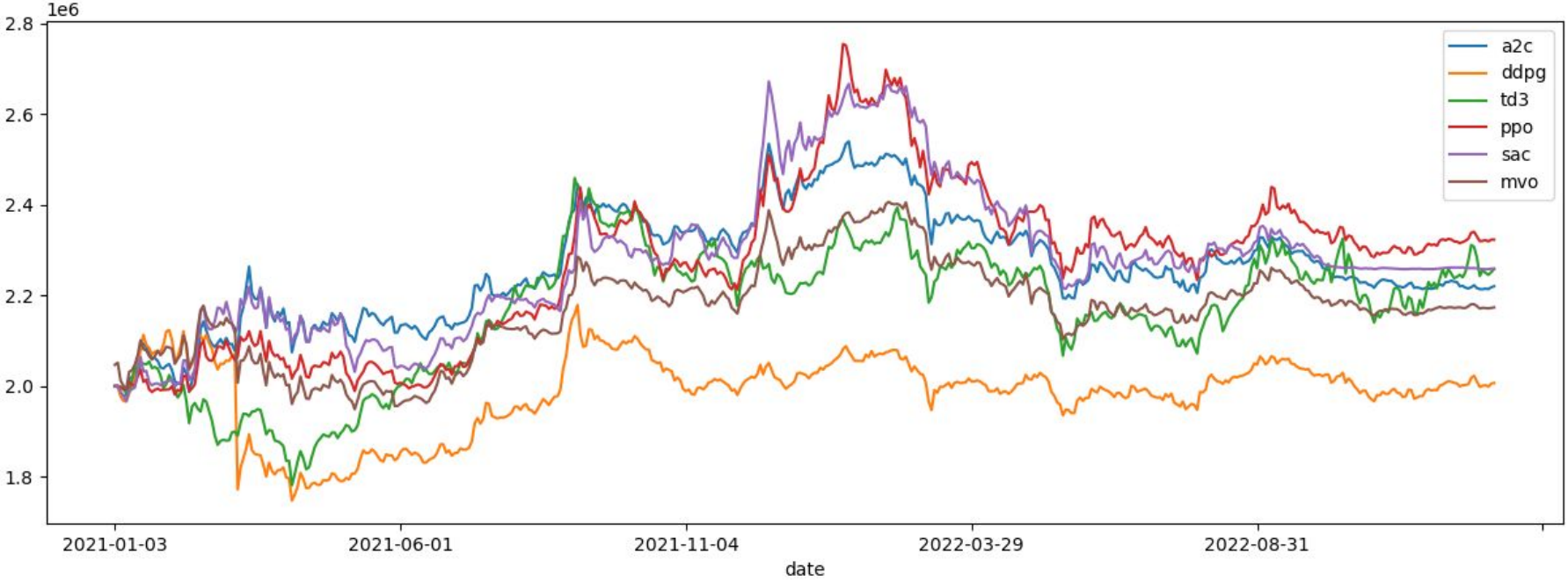
01. Train 2010-2020 Trade 2021-2022

Dataset: DSE top 39

	initial_value	final_value	annual_return	sharpe_ratio	max_drawdown
a2c	1.855317e+06	1.859139e+06	0.068613	0.000001	-289717.900000
ddpg	1.967413e+06	1.921226e+06	-0.788738	-0.000008	-374417.600000
td3	2.052115e+06	2.093511e+06	0.667931	0.000008	-343964.500000
ppo	1.968796e+06	1.956892e+06	-0.201946	-0.000002	-336006.133750
sac	2.056457e+06	2.147644e+06	1.456733	0.000016	-368834.300000
mvo	2.007220e+06	2.048120e+06	0.674651	0.000009	-359588.740402

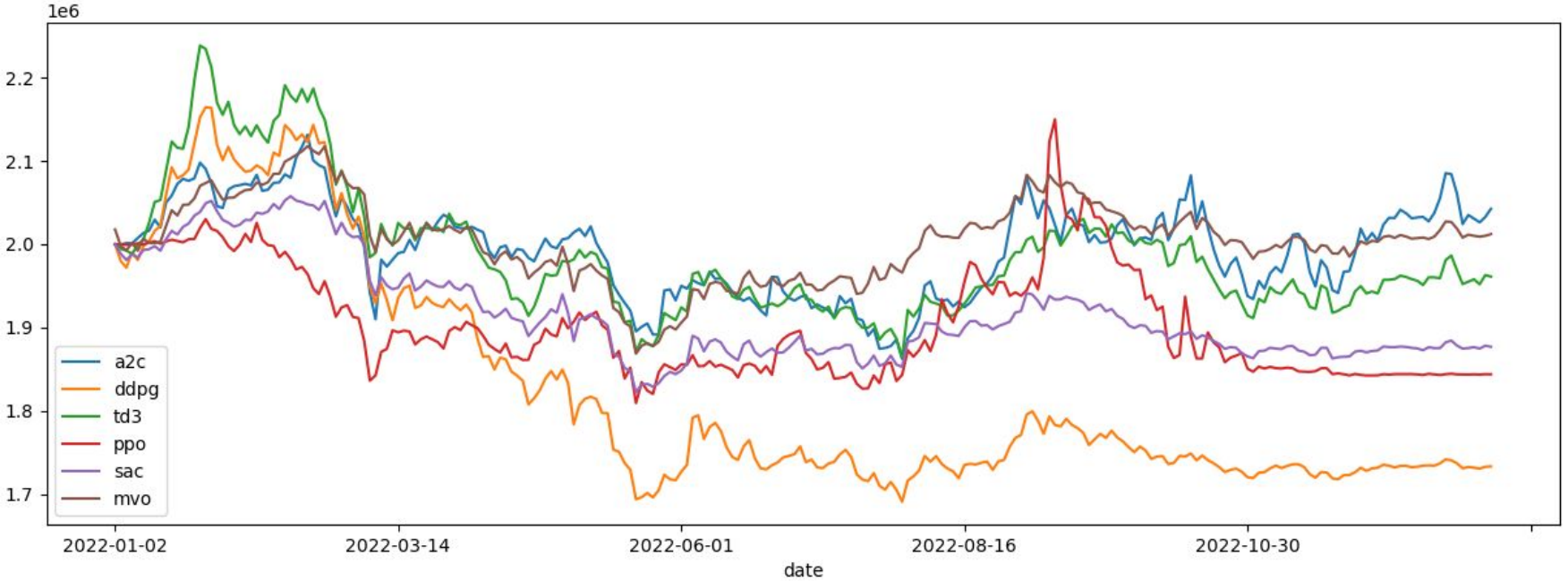
01. Train 2010-2020 Trade 2021-2022

Dataset: DSE top 39



02. Train 2012-2020 Trade 2021-2022

Dataset: DSE top 39



03. Train 2012-2021 Trade 2022

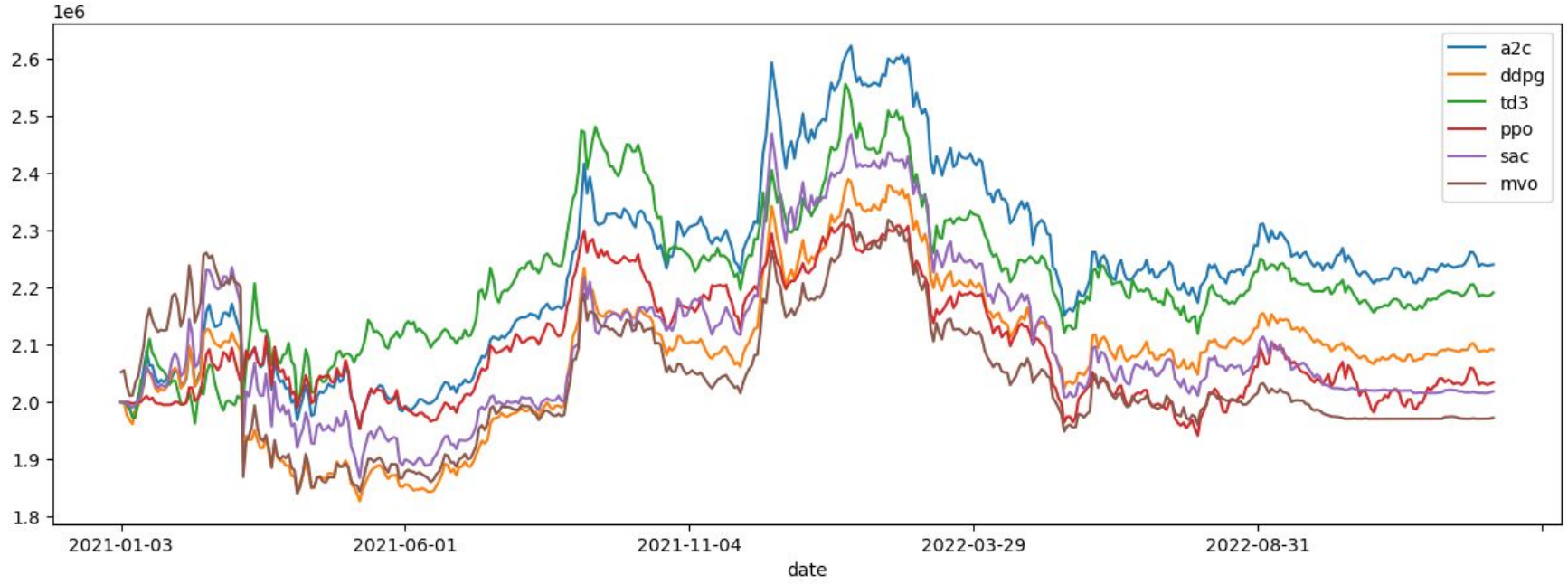
Custom Reward Functions

Cashless Penalty

- Absence of cash when a stock is doing good and it should be bought.
- Not having the ability to buy a stock at the right moment requires penalties.

```
self.reward = (end_total_asset -
begin_total_asset - max(0,
((self.cash+end_total_asset - self.state[0]) *
self.cash_penalty_percentage -
self.cash))*self.cash_penalty_percentage) *
self.reward_scaling
```

Dataset: DSE top 39



01. **Cash Penalty**

Dataset: DSE top 39

	initial_value	final_value	annual_return	sharpe_ratio	max_drawdown
a2c	2.000000e+06	2.239783e+06	3.846530	0.000024	-470314.400000
ddpg	2.000000e+06	2.091690e+06	1.505383	0.000012	-364009.500000
td3	2.000000e+06	2.190890e+06	3.085335	0.000025	-435628.700000
ppo	2.000000e+06	2.033836e+06	0.560780	0.000006	-370989.179550
sac	2.000000e+06	2.018565e+06	0.308466	0.000002	-459919.248350
mvo	2.052309e+06	1.972859e+06	-1.307439	-0.000012	-420029.816314

01. **Cash Penalty**

Custom Reward Functions

Stop Loss

- Risk-management system that automatically sells a security once it reaches a certain price.
- Designed to limit losses in case the security's price drops below that price level.

- This environment penalizes the model if exceeded the stop-loss threshold, selling assets with under expectation %profit, and also for not maintaining a reserve of cash.
- This enables the model to do trading with high confidence and manage cash reserves in addition to performing trading procedures.

Custom Reward Functions

Stop Loss

r_i = (sum(cash, asset_value) + additional_reward - total_penalty - initial_cash) / initial_cash / days_elapsed

, where total_penalty = cash_penalty + stop_loss_penalty + low_profit_penalty

cash_penalty = max(0, sum(cash, asset_value)*cash_penalty_proportion-cash)

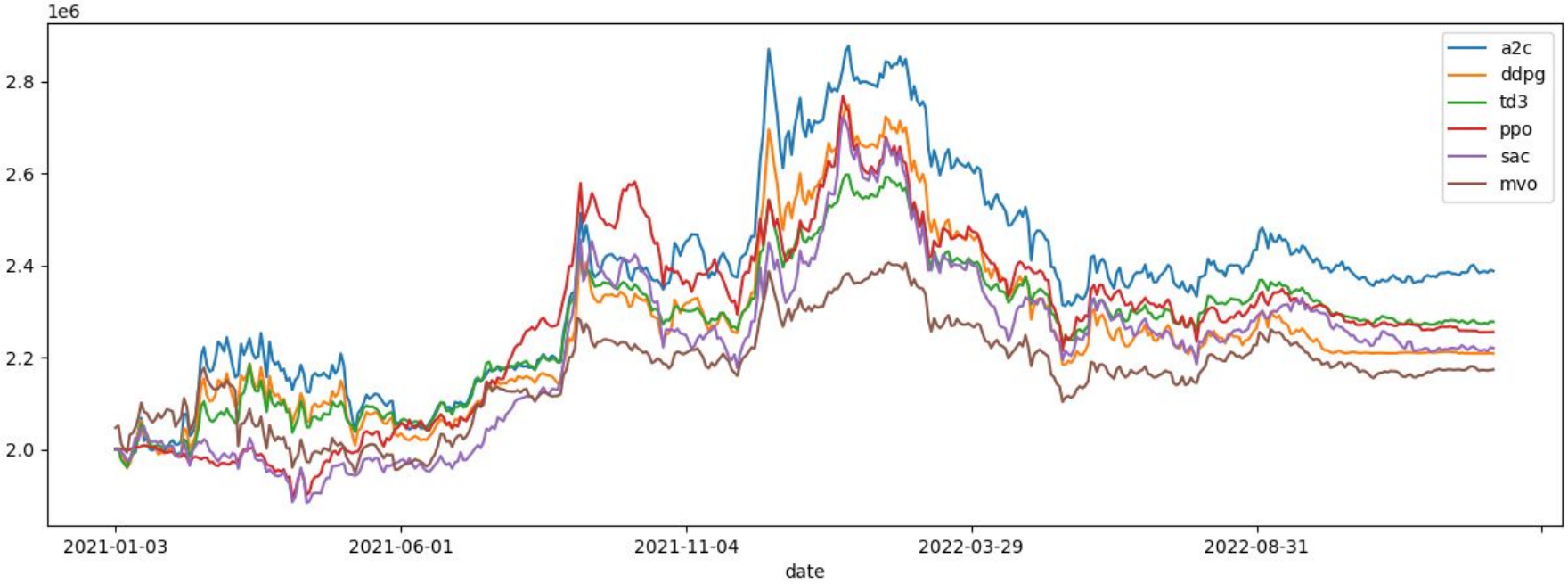
stop_loss_penalty = -1 * dot(holdings,negative_closing_diff_avg_buy)

low_profit_penalty = -1 * dot(holdings,negative_profit_sell_diff_avg_buy)

additional_reward = dot(holdings,positive_profit_sell_diff_avg_buy)

Cash = Exponential Average of last 25 days.

Dataset: DSE top 39



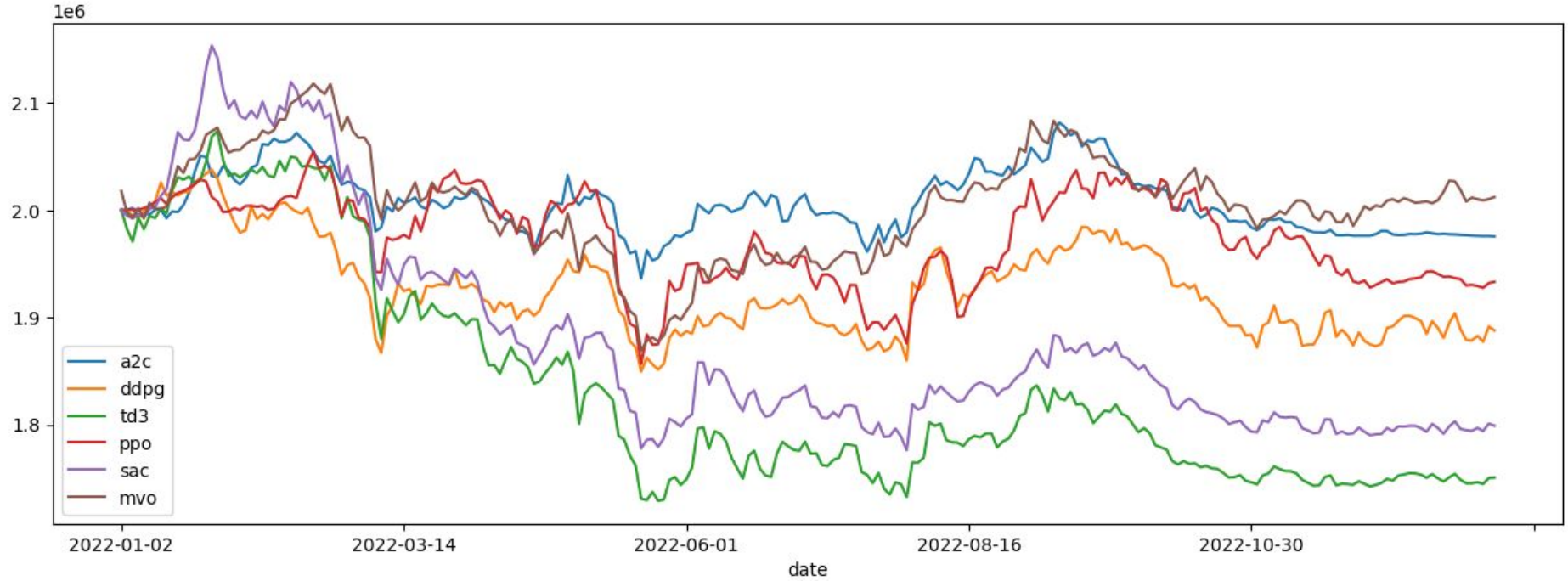
01. Train 2012-2020 Trade 2021-2022 **Stop Loss**

Dataset: DSE top 39

	initial_value	final_value	annual_return	sharpe_ratio	max_drawdown
a2c	2000000.0	2220054.90	3.54	2.91e-05	-346426.19
ddpg	2000000.0	2006954.28	0.11	1.37e-06	-376089.79
td3	2000000.0	2257929.22	4.12	2.83e-05	-391398.89
ppo	2000000.0	2322867.71	5.11	2.90e-05	-520390.64
sac	2000000.0	2259114.39	4.14	2.59e-05	-456789.89
mvo	2047001.23	2173538.62	2.01	1.91e-05	-302759.42

01. Train 2012-2020 Trade 2021-2022 **Stop Loss**

Dataset: DSE top 39



02. Train 2012-2021 Trade 2022 **Stop Loss**

Future Directions

- Incorporate other factor and indicators.
- Reward Functions in the FinRL library could be optimized so they match DSE's patterns.
- Guided train to minimize risk.