Sirius University
of Science and Technology

MathAI 2025

SIRIUS AI LABORATORY

DEEP LEARNING MODELS MEET FINANCIAL DATA MODALITIES

Khubiev Y. K. , Semenov E. M.

Sirius 2025

Introduction

Algorithmic trading is a dynamic and continuously evolving field, that seeks new robust solutions.

To build robust and effective strategies we need diverse data, and we benefit from multimodality.

Deep Learning is on trend and there are various attempst to incorporate modern approach into quantitative finance [1, 2].

[1] Ho, T.-T.; Huang, Y. Stock Price Movement Prediction Using Sentiment Analysis and CandleStick Chart Representation. Sensors 2021, 21, 7957.
[2] Fazlija, B.; Harder, P. Using Financial News Sentiment for Stock Price Direction Prediction. Mathematics2022, 10, 2156.

Introduction

Multimodality

Data Source
- Candlestick data
- News flow

Financial Source
- Limit order book
- Candlesticks
- Trades statistics
- Orders statistics

[3] K.Y.Khubiev, M.E.Semenov, “Multimodal Stock Price Prediction: A Case Study of the Russian Securities Market,“ Program Systems: Theory and Applications 16, No.1, 83–130, (2025)

The Goal

1. Investigate *financially* grounded loss functions for algorithmic trading strategies (alphas) and portfolio optimization improvement

2. Introduce SotA approach for Limit Order Book (LOB) data treating

Data

Algorithmic Trading

- Moscow exchange
- Candlestick data
- Trades statistic
- Orders statistic
- LOB statistic
- 1 minute frequent data points

Limit Order Book

- ByBit exchange
- LOB state snapshots scrapped
- 200-300 ms frequent data points

Buy Order

Price (USDT)	Amount (BTC)	Total (USDT)	Sum (USDT)
87,992.78	3.62490	318,965.0282220	318,965.0282220
87,992.77	0.00058	51.0358066	319,016.0640286
87,992.39	0.00007	6.1594673	319,022.2234959
87,992.38	0.00007	6.1594666	319,028.3829625
87,992.37	0.00007	6.1594659	319,034.5424284
87,992.17	0.00258	227.0197986	319,261.5622270
87,992.08	0.00007	6.1594456	319,267.7216726
87,992.02	0.00007	6.1594414	319,273.8811140
87,992.01	0.00018	15.8385618	319,289.7196758
87,992.00	0.07067	6,218.3946400	325,508.1143158
87,991.92	0.00012	10.5590304	325,518.6733462
87,991.91	0.04680	4,118.0213880	329,636.6947342
87,991.90	0.00006	5.2795140	329,641.9742482
87,991.58	0.00064	56.3146112	329,698.2888594
87,990.81	0.00251	220.8569331	329,919.1457925

Sell Order

Price (USDT)	Amount (BTC)	Total (USDT)	Sum (USDT)
87,992.79	3.53713	311,241.9372927	311,241.9372927
87,992.80	0.00317	278.9371760	311,520.8744687
87,992.81	0.00258	227.0214498	311,747.8959185
87,992.82	0.00019	16.7186358	311,764.6145543
87,992.83	0.15164	13,343.2327412	325,107.8472955
87,992.84	0.00007	6.1594988	325,114.0067943
87,993.25	0.00014	12.3190550	325,126.3258493
87,993.63	0.00007	6.1595541	325,132.4854034
87,993.64	0.00254	223.5038456	325,355.9892490
87,993.65	0.00018	15.8388570	325,371.8281060
87,993.66	0.16255	14,303.3694330	339,675.1975390
87,993.82	0.00006	5.2796292	339,680.4771682
87,994.00	0.07073	6,223.8156200	345,904.2927882
87,994.42	0.10280	9,045.8263760	354,950.1191642
87,994.56	0.00006	5.2796736	354,955.3988378

Methodology

Algorithmic Trading

Let us define α to be a position vector, and $\boldsymbol{r}$ to be a return vector of length M equal the number of observed assets

Sharp Loss:

$$SharpLoss = -\frac{|\mathbb{E}(\mathbf{pnl})|}{\sigma(\mathbf{pnl}) + \epsilon} \quad (1)$$

Profit and Loss Loss

$$PnLLoss = -\alpha\mathbf{r} \quad (2)$$

where $pnl_i = \alpha_i r_i, \quad i = \overline{1, M}$

Modified Sharp Loss:

$$ModSharpLoss = \frac{\|\boldsymbol{\alpha} - \mathbf{r}\|_2^2}{|\mathbb{E}(\mathbf{pnl})| + 1} \quad (3)$$

Max Drawdown Loss

$$MDDLoss = -\min(cumsum(\mathbf{pnl}) - cummax(\mathbf{pnl})) \quad (4)$$

Turnover Regularization

$$\mathrm{TvrReg} = \lambda \cdot (\max(1, tvr - tb) + \max(1, bb - tvr)) \quad (5)$$

Methodology

Portfolio Optimization

Having a batch of low correlated alphas we need to construct a portfolio, that is a linear combination of the alphas:

- Equal weighted - each alpha positions contribute with the same fixed weight along the time interval
- Single weighted - a model generates a single weight for a given alpha broadcasting over its positions each day
- Point-wise - a model generates a vector of length M for each given alpha position each day

We use the introduced custom loss functions for optimization task

Methodology

Limit Order Book

LOB snapshot its instant state, we treat a sequence of LOB frames as input channels

Sampling: sliding window, gathering within a time interval
Embedding: scaling and image embodiment [4, 5]

- we propose to treat LOB as a) N dimensional and b) Grayscale image [4]
- we introduce Min-Max-Domain scaling that improves [4] and outperform [5] scaling techniques.

We also investigate the optimal target value: relative increment and price delta.

[4] Wuyi Ye, Yang Jinting and Chen Pengzhan, "Short-term stock price trend prediction with imaging high frequency limit order book data," International Journal of Forecasting, 40, 1189–1205 (2022).
[5] A. Ntakaris, M. Magris, J. Kanniainen, M. Gabbouj and A. Iosifidis, "Benchmark dataset for mid-price forecasting of limit order book data with machine learning methods," Journal of Forecasting (2017)

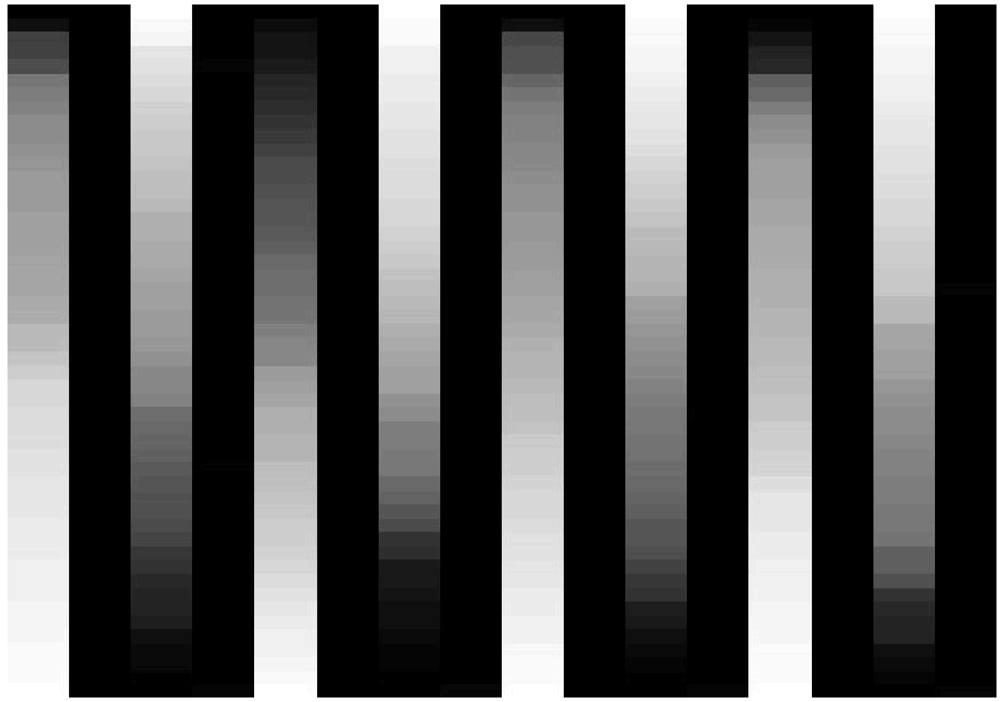
Methodology

Limit Order Book

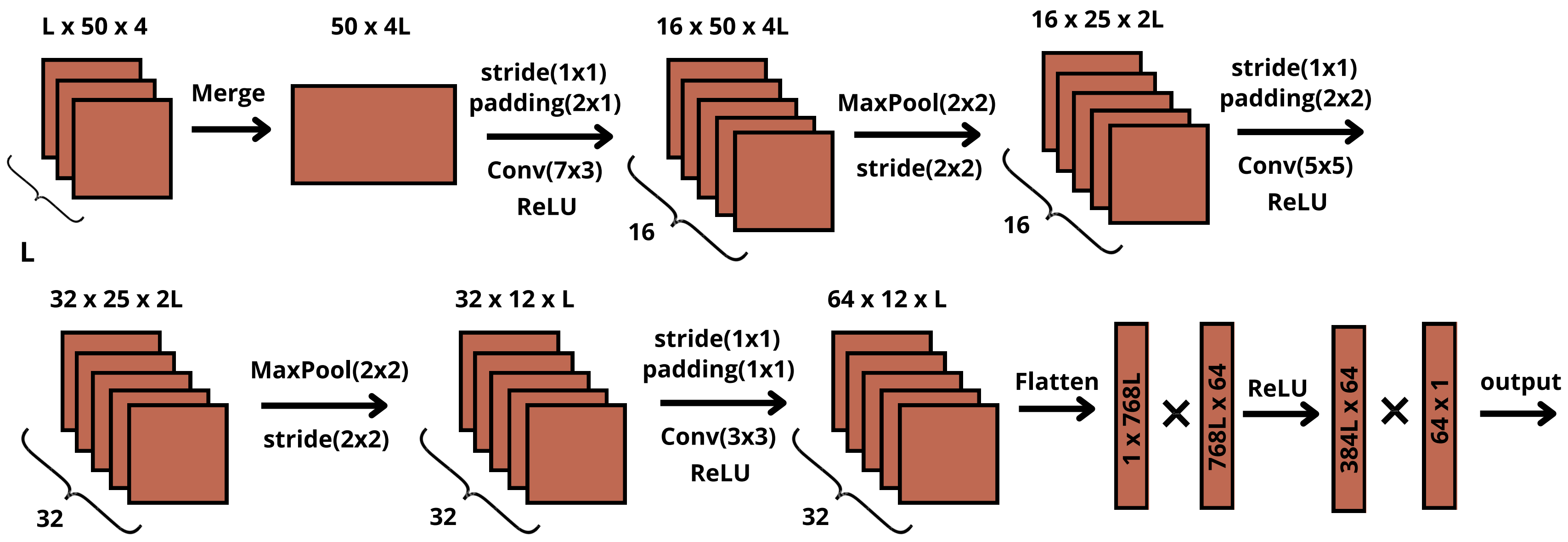
Min-Max-Global

Min-Max-Domain

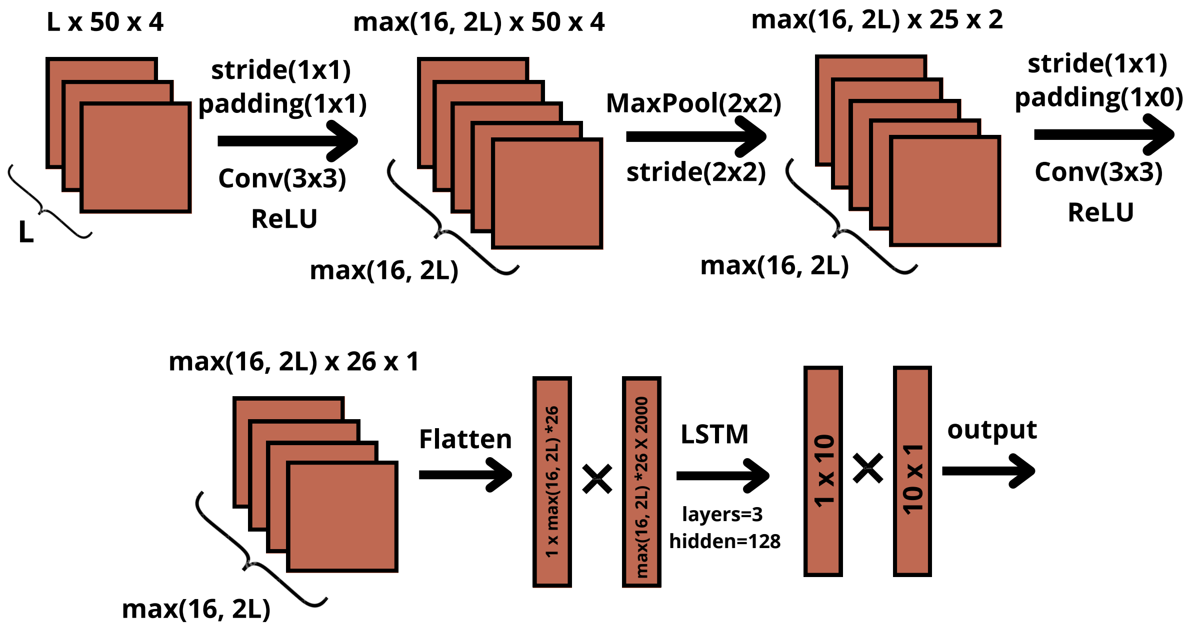
Z-Score



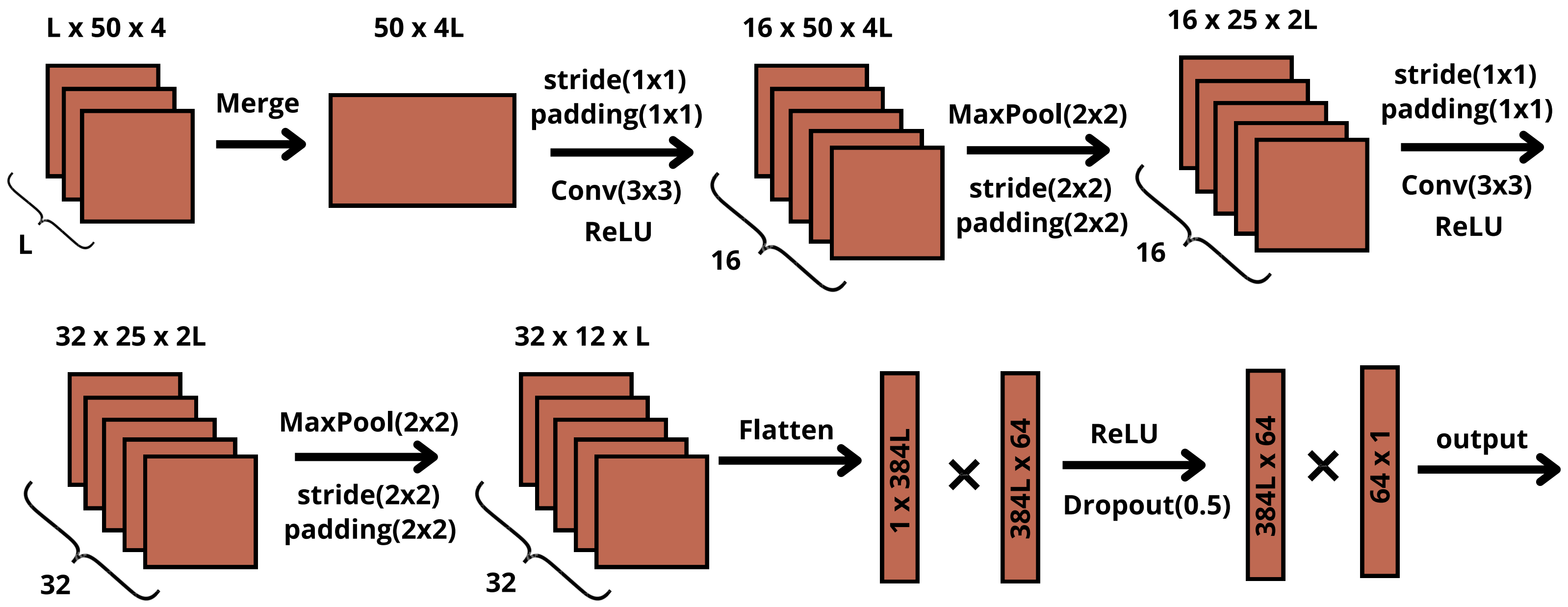
Models



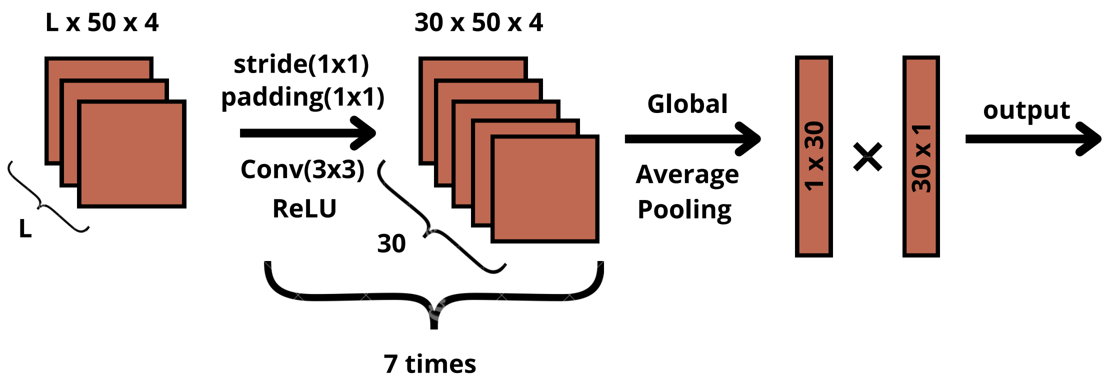
(a) CNNModel_2D



(a) CNN2LSTM



(a) SimpleCNN_2D

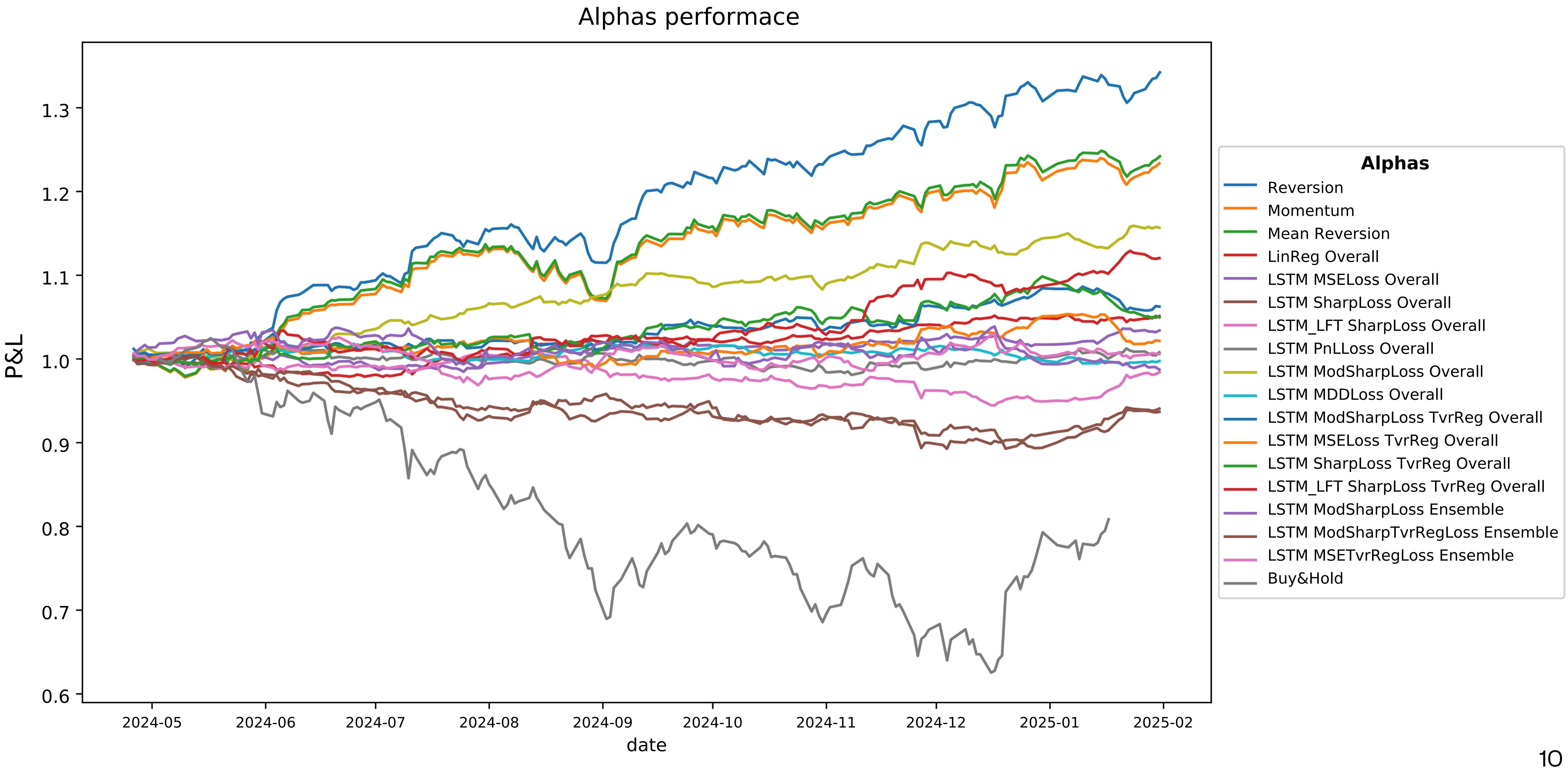


(d) SimpleCNN

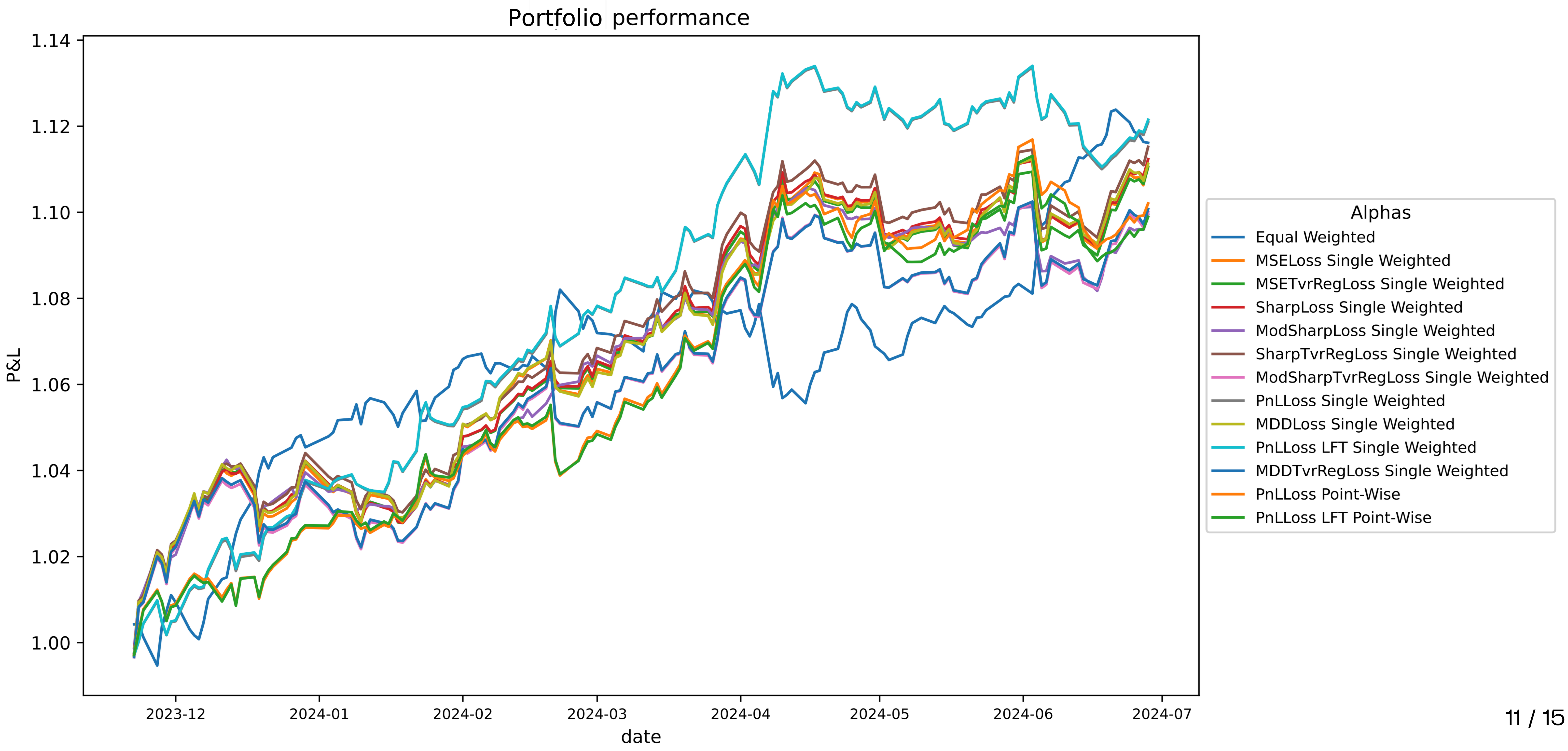
Sirius University
of Science and Technology

SIRIUS AI LABORATORY

Experiment and Results



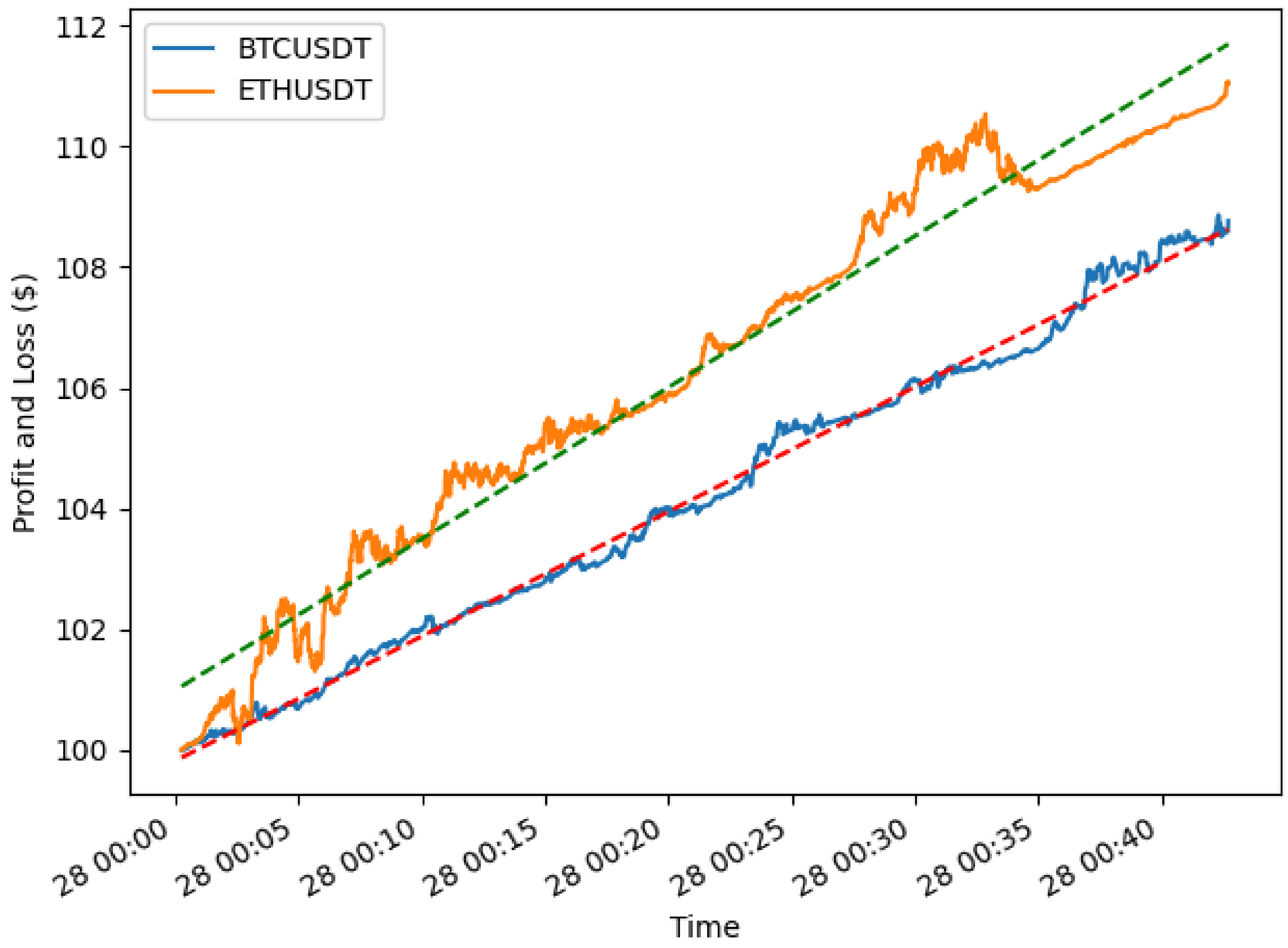
Experiment and Results



Sirius University of Science and Technology

SIRIUS AI LABORATORY

Experiment and Results



Model	Aggregation	MAPE, %
CNN2LSTM	staked	0.018244
SimpleCNN	stacked	0.103163
CNNModel_2D	merged	0.344536
SimpleCNN_2D	merged	0.377812

Aggregation	MAPE, %
stacked	0.060703
merged	0.361174

Line	Velocity (bps)
BTCUSDT	17.64
BTCUSDT Regression	17.61
ETHUSDT	22.26
ETHUSDT Regression	21.40

Conclusion & Further work

- The Custom loss functions help to fit deep learning model for designing alphas and porrtfolio optimization
- The Turnover regularization helps to control generated weight natively
- The proposed approach for LOB data performed SotA results

- The issue of the introduced regularization remains unaddressed
- The orders queue and execution limits are essential to investigate further
- The offered loss functions might be beneficial might improve reinforcement learning policies
- Design language agents for trading with financially grounded reward policies

[1] Ho, T.-T.; Huang, Y. Stock Price Movement Prediction Using Sentiment Analysis and CandleStick Chart Representation. *Sensors* 2021, *21*, 7957. https://doi.org/10.3390/s21237957

[2] Fazlija, B.; Harder, P. Using Financial News Sentiment for Stock Price Direction Prediction. *Mathematics* 2022, *10*, 2156. https://doi.org/10.3390/math10132156.

[3] K.Y.Khubiev, M.E.Semenov, "Multimodal Stock Price Prediction: A Case Study of the Russian Securities Market," Program Systems: Theory and Applications 16, No.1, 83–130, (2025), URL: https://psta.psiras.ru/2025/1 83-130.

[4] Wuyi Ye, Yang Jinting and Chen Pengzhan, "Short-term stock price trend prediction with imaging high frequency limit order book data," International Journal of Forecasting, 40, 1189–1205 (2022).

[5] A. Ntakaris, M. Magris, J. Kanniainen, M. Gabbouj and A. Iosifidis, "Benchmark dataset for mid-price forecasting of limit order book data with machine learning methods," Journal of Forecasting (2017) URL: https://api.semanticscholar.org/CorpusID:158644016

DEEP LEARNING MODELS MEET FINANCIAL DATA MODALITIES

Kasymkhan Khubiev

kasymkhankhubievnis@gmail.com

theGorgeousKing

Mikhail Semenov

semenov.me@talantiuspeh.ru

Semenov_ME