

Supplementary Materials

Causal-Structure-Based Cryptocurrency Price Direction Prediction Model

Yuantai Cui¹ and Hiroaki Fukunishi^{1,2,*}

(*Co-corresponding author)

¹ Graduate School of Bionics, Computer Science and Media Sciences, Tokyo University of Technology,
Technology,

Tokyo, 192-0982, Japan

² School of Computer Science, Tokyo University of Technology, Tokyo, 192-0982, Japan

***Corresponding author details:**

H. Fukunishi

e-mail: fukunishiha@stf.teu.ac.jp

Contents:

Supplementary Material S1. Return-based performance using various causal filters.

Supplementary Material S2. Predictive performance of LSTM across different timesteps.

Supplementary Material S3. Causal graphs inferred in the backtesting experiment.

Supplementary Material S1. Return-based performance using various causal filters.

Table S1 presents the return-based performance of RW trials conducted under filter conditions, defined by a directed edge from each feature to the target variable. The No-Filter setting (baseline) represents performance across all 48 RW trials. Hereafter, each filter is denoted as “feature name → target.”

The Open → target filter (Open filter) achieved the highest win rate (59.26%), with a monthly average return of 3.34%, based on 27 trading RW trials. Although the EMA_200 → target filter recorded the highest monthly average return (5.01%), it was applied in only two RW trials, making the result potentially dependent on specific market conditions. Overall, the Open filter showed the most favorable return-based performance among the examined candidate filters in the exploratory phase for trading strategy design and was therefore selected for the subsequent backtesting experiments.

Table S1. Return-based performance under a 0.1% transaction fee in RW trials where a specific directed edge was detected.

Filter	Trading RW Trials (Months)	Win Rate (%)	Monthly Average Return (%)	Monthly Average Number of Trades
No-Filter (Baseline)	48	41.67	0.36	20.21
RSI_7 → target	46	41.30	0.21	19.89
RSI_14 → target	39	46.15	1.26	19.23
High → target	35	40.00	0.06	21.49
Close → target	34	47.06	1.67	21.88
Open → target	27	59.26	3.34	21.48
Low → target	22	50.00	3.08	21.50
EMA_50 → target	22	50.00	2.52	21.64
Bollinger_u → target	21	38.10	0.62	22.71
Bollinger_l → target	20	35.00	-1.22	22.80
Bollinger_m → target	19	36.84	-1.18	23.47
EMA_12 → target	11	45.45	3.10	22.73
EMA_26 → target	11	45.45	3.10	22.73
MACD_diff → target	3	33.33	-2.47	23.00
EMA_200 → target	2	50.00	5.01	21.50

Bold values indicate values that exceeded those of the No-Filter (baseline) strategy.

Supplementary Material S2. Predictive performance of LSTM across different timesteps.

The performance of LSTM strongly depends on the input sequence length (timesteps) hyperparameter, which determines how far back past observations are referenced at each prediction point. The hyperparameter defines the temporal dependency range captured by the prediction model. To identify the optimal timesteps, four configurations were compared: 6 (1 day; six 4-hour intervals), 18 (3 days), 30 (5 days), and 42 (7 days). The fixed-length window (15 months), the number of RW trials (48), and the trading rules were kept identical to those used for NLBPM.

Table S2.1 presents the return-based performance of LSTM across different timesteps. Figure S2 visualizes the impact of transaction fee rates on monthly average returns for each timesteps setting. The results show that, regardless of the transaction fee rate, a timesteps of 18 yields relatively higher monthly average returns. Under the 0% fee condition, positive monthly average returns were observed for timesteps of 18 and 30, whereas negative returns were recorded for timesteps of 6 and 42. Once the transaction fee rate reached 0.06% or higher, monthly average returns became negative for all timesteps. This suggests that LSTM does not possess sufficient predictive power to overcome transaction costs. Moreover, the average monthly number of trades tended to decrease as the timesteps increased. This may be because excessively long input timesteps reduce the model's responsiveness to recent trend changes, thereby missing profitable trading opportunities.

Table S2.2 presents the price direction classification performance of LSTM across different timesteps. The timesteps of 18 achieved the highest values in terms of monthly average accuracy, precision, and F1-score, which is consistent with the return-based performance results. However, the fact that the monthly average accuracy remained below 0.5 reflects the well-known challenges of applying deep learning models to financial time series forecasting, such as overfitting and sensitivity to noise.

Based on these findings, the timesteps of 18, which achieved the highest monthly average return, was adopted for backtesting experiments.

Table S2.1. Return-based performance of LSTM across different timesteps.

Timesteps	Monthly Average Return (%)				Win Rate (%)	Monthly Average Number of Trades
	Fee 0%	Fee 0.06%	Fee 0.08%	Fee 0.1%		
6 (1 day)	-0.40	-3.20	-4.11	-5.02	25.00	23.29
18 (3 days)	2.14	-0.89	-1.88	-2.86	33.33	24.48
30 (5 days)	1.09	-1.51	-2.36	-3.20	39.58	21.15
42 (7 days)	-0.71	-2.95	-3.68	-4.41	27.08	18.52

Bold values indicate the highest values across timesteps for each metric.

Table S2.2. Price direction classification performance of LSTM across different timesteps.

Timesteps	Monthly Average	Monthly Average	Monthly Average	Monthly Average
	Accuracy	Precision	Recall	F1-score
6 (1 day)	0.492	0.508	0.516	0.499
18 (3 days)	0.499	0.519	0.512	0.506
30 (5 days)	0.493	0.504	0.479	0.476
42 (7 days)	0.492	0.508	0.510	0.489

Bold values indicate the highest values across timesteps for each metric.

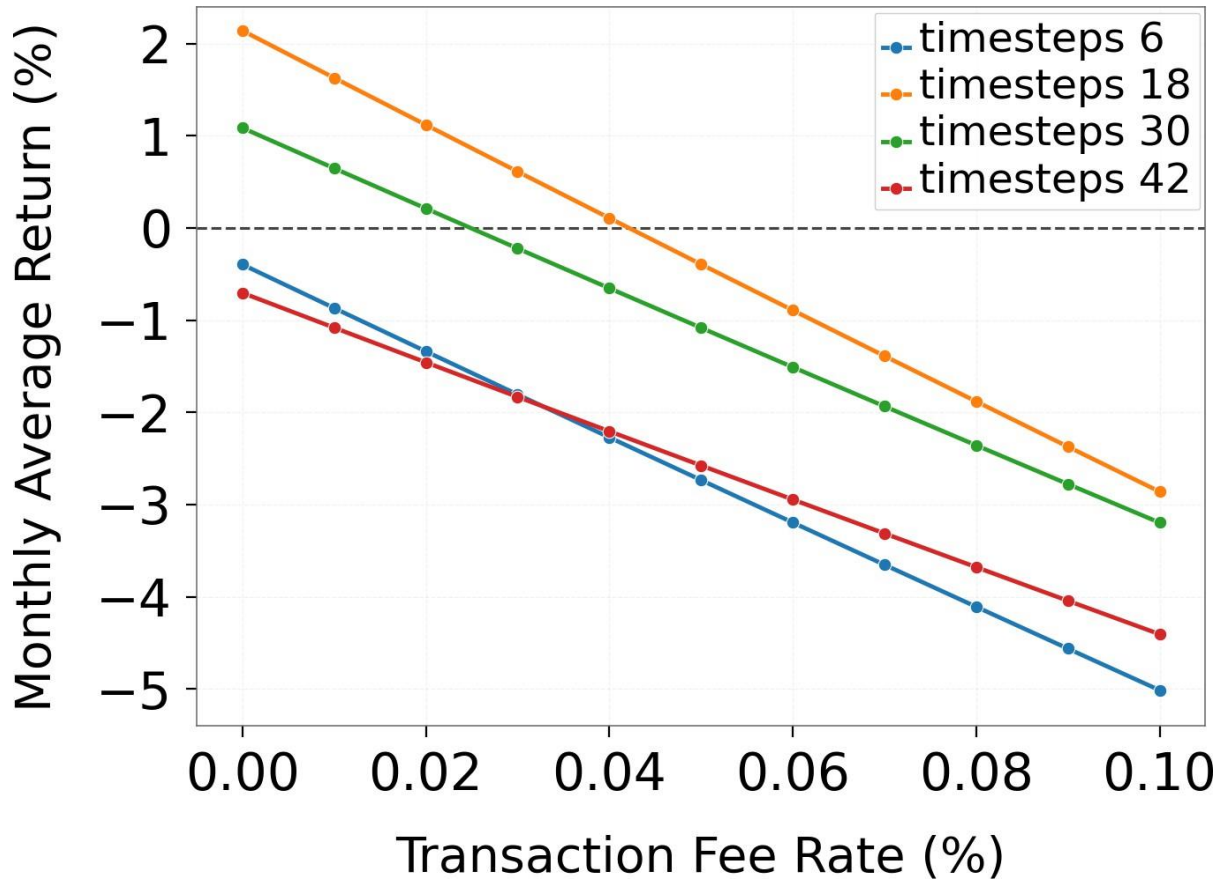


Figure S2. Monthly average returns of LSTM under different transaction fee rates across different timesteps.

Supplementary Material S3. Causal graphs inferred in the backtesting experiment.

Figure S3 presents the causal graphs inferred in the four RW trials selected by the Open filter out of the total 12 RW trials in the backtesting experiments, where a directed edge from Open (the opening price) to the target variable was identified. In all these RW trials, the corresponding edge weights were negative. In the backtesting period, this negative weight suggests that higher opening prices were associated with a higher likelihood of predicting a downward price movement four hours later. Under the trading rules adopted in this study, in such situations, new buy orders are suppressed and existing positions are liquidated. Therefore, this functions as a risk-avoidance signal that helps prevent potential losses in advance.

Furthermore, regarding the Relative Strength Index (RSI), the short-term indicator RSI_7 (7-period) exhibited a negative weight, whereas the medium-term indicator RSI_14 (14-period) exhibited a positive weight. This pattern is consistent with the high-return causal graphs observed in the exploratory phase for trading strategy design. These findings suggest that an increase in the short-term RSI_7 reflects overbought conditions that generate selling pressure, whereas an increase in the medium-term RSI_14 indicates trend-driven buying pressure.

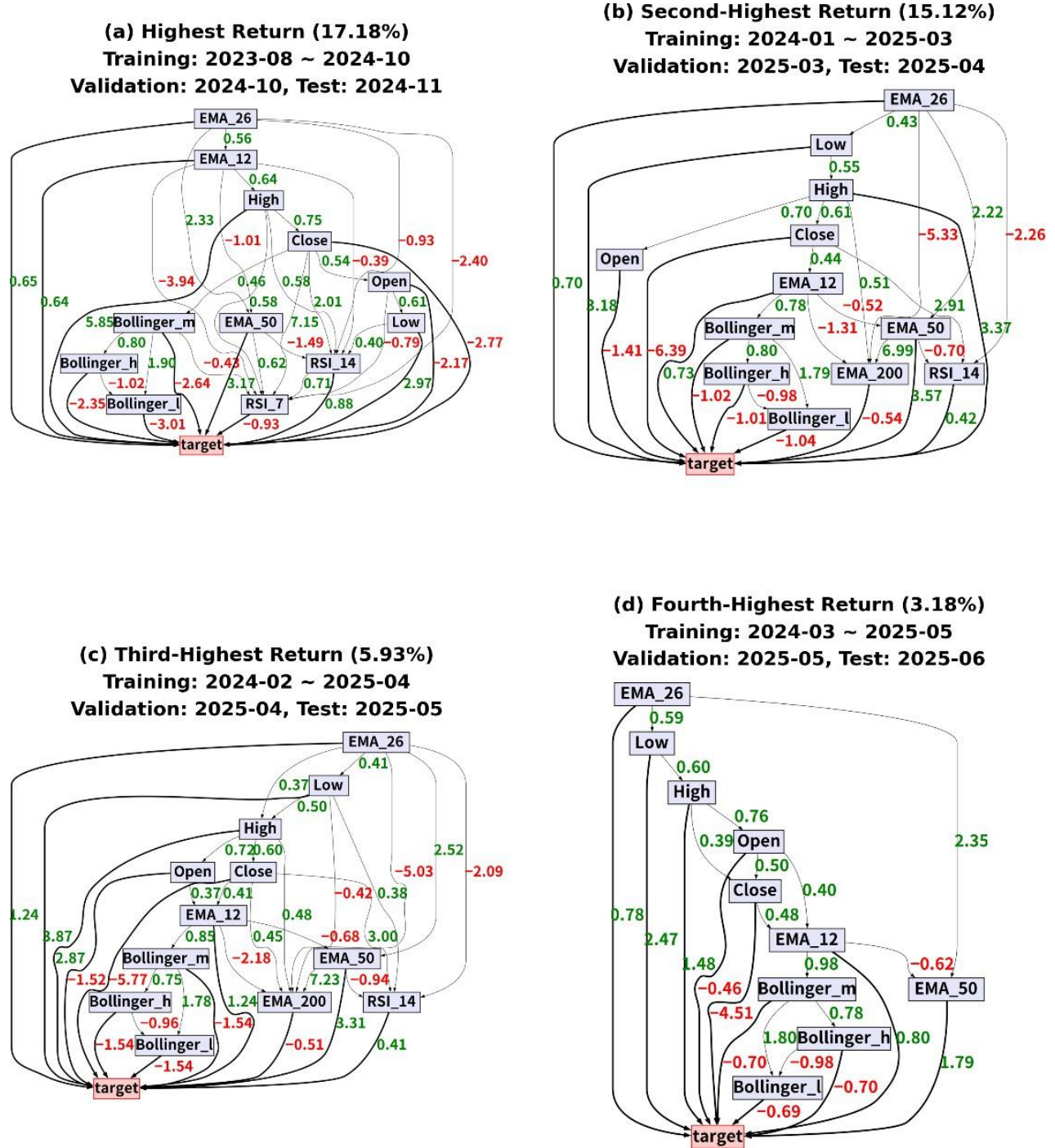


Figure S3. Causal graphs inferred in the four RW trials selected by the Open filter in the backtesting experiments. The red node indicates the target variable. Green and red values represent positive and negative weights. Thick edges denote the directed edge between the target variable and its parent nodes.