Informer In Algorithmic Investment Strategies on High Frequency Bitcoin Data

Filip Stefaniuk and Robert Ślepaczuk

Research Questions:

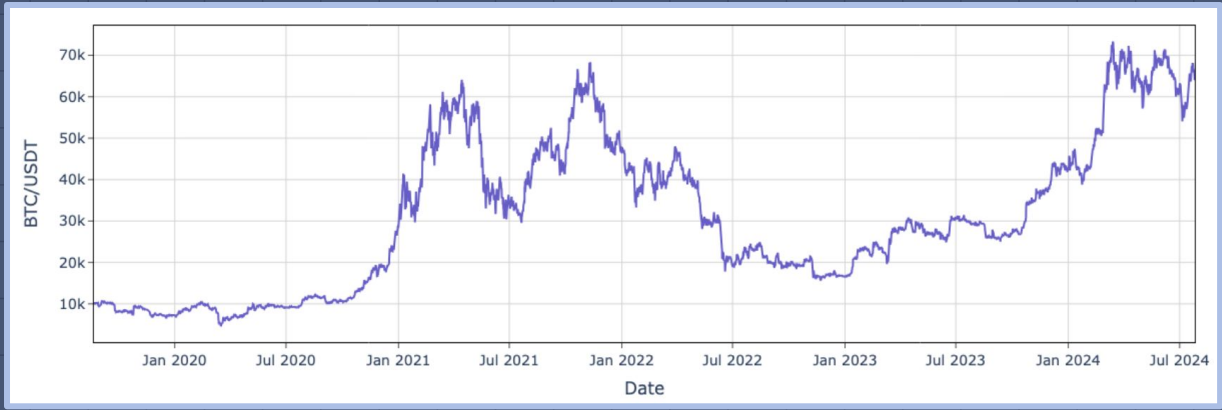
- *Is it possible to create an algorithmic strategy for trading Bitcoin, that is more efficient than Buy&Hold approach?*
- *Does signal from Informer model allow to create strategies that are more efficient on trading Bitcoin than strategies based on technical indicators?*
- *Does usage of higher frequency data allow to create more efficient strategies?*
- *How does selection of the machine learning model loss function influence the strategy performance?*

DATA

1

BTC/USDT

- Availability of high quality historical data, with fine-grained intervals, even up to 1 second.
- Cryptocurrency exchanges operate continuously, 24 hours a day, 7 days a week.
- Bitcoin stands as the most established cryptocurrency, with the BTC/USDT pair being one of the highest in trading volume



Price of BTC/USDT cryptocurrency pair in a period from 21.08.2019 to 24.07.2024 (5 years). The data was obtained from Binance.

https://www.binance.com/en-NG/landing/data

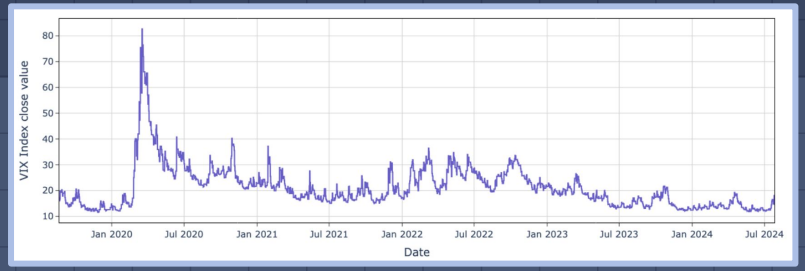
BTC/USDT Returns

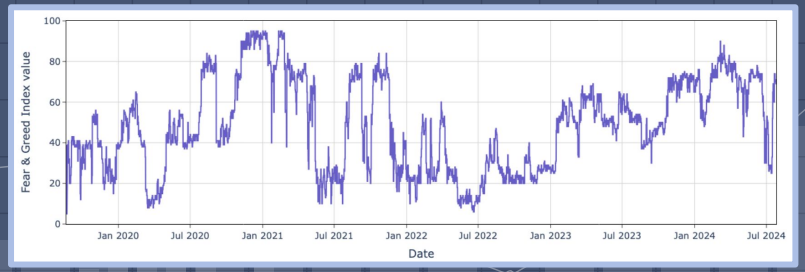
Statistic	5 min	15 min	30 min
Count	518 400	172 800	86 400
Mean	0.0000060	0.0000176	0.0000346
Std	0.0021843	0.0036712	0.0050768
Min	-0.1022537	-0.1191688	-0.1662924
25% percentile	-0.0007716	-0.0013018	-0.0017677
50% percentile	0.00	0 0.0000094	0.0000280
75% percentile	0.0007855	0.0013443	0.0018575
Max	0.1842885	0.2262878	0.1460125

```
[
  [
    1499040000000,      // Kline open time
    "0.01634790",       // Open price
    "0.80000000",       // High price
    "0.01575800",       // Low price
    "0.01577100",       // Close price
    "148976.11427815",  // Volume
    1499644799999,      // Kline Close time
    "2434.19055334",    // Quote asset volume
    308,                // Number of trades
    "1756.87402397",    // Taker buy base asset volume
    "28.46694368",      // Taker buy quote asset volume
    "0"                 // Unused field, ignore.
  ]
]
```

Additional Data

- The Cboe Volatility Index (VIX Index)
- The Federal Funds effective rates
- Crypto Fear/Greed index

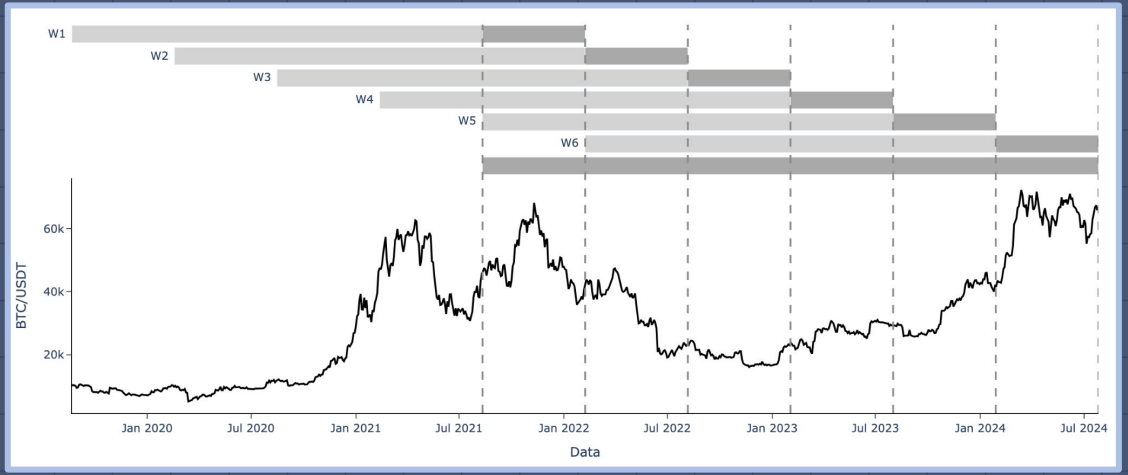




https://www.cboe.com/tradable_products/vix/vix_historical_data/
https://fred.stlouisfed.org/series/FEDFUNDS
https://alternative.me/crypto/fear-and-greed-index/

Data Windows

Part	5 min	15 min	30 min
Training	165888	55296	27648
Validation	41472	13824	6912
Test	51840	17280	8640



Michańków, Jakub, Pawel Sakowski, and Robert Ślepaczuk (2022). "LSTM in Algorithmic Investment Strategies on BTC and S&P500 Index". In: Sensors 22.3. issn: 1424-8220. doi: 10.3390/s22030917. url: https://www.mdpi.com/1424-8220/22/3/917.

METHODOLOGY

2

Assumptions

- During each time interval, one of the three possible positions can be taken: `long (1)`, `short(-1)` or `no position (0)`.
- When the strategy position changes, assets are always bought/sold at the close price of the previous interval.
- Each such position change incurs an exchange fee assumed to be 0.1%.
- Any fraction of the asset can be traded.
- No splitting of portfolio allocation, the asset is always bought/sold with the full portfolio value.
- At the end of the evaluation period all positions must be closed.

Metrics

- Annualized Return Compounded (ARC)
- Annualized Standard Deviation (ASD)
- Maximum Drawdown (MD)
- Information Ratio (IR*)
- **Modified Information Ratio (IR**)**
- Number of trades (N)
- Percentage of Long Position (LONG) / Short Position (SHORT)

$$IR^* = \frac{ARC}{ASD}$$

$$IR^{**} = IR^* \times \frac{|ARC|}{MD}$$

Kosc, Krzysztof, Pawel Sakowski, and Robert Ślepaczuk (2019). "Momentum and contrarian effects on the cryptocurrency market". In: Physica A: Statistical Mechanics and its Applications 523, pp. 691–701. issn: 0378-4371. doi: https://doi.org/10.1016/j.physa.2019.02.057. url: https://www.sciencedirect.com/science/article/pii/S037843711930216X.

Strategy

$$s_\theta([x_{t-\lambda}; ...; x_{t-1}]) = p_t$$

$$s_\theta : \mathbb{R}^{\lambda \times d} \rightarrow \{-1, 0, 1\}$$

$$s^{BuyAndHold}(\cdot) = 1$$

MACD Strategy

$$s_{\theta_{MACD}}^{MACD}(\cdot) = \begin{cases} 1 & if\ MACD_{t-1} \geq SIGNAL_{t-1} \\ -short & if\ MACD_{t-1} < SIGNAL_{t-1} \end{cases}$$

$$\theta_{MACD} = (fast, slow, signal, short)$$

RSI Strategy

$$s_{\theta_{RSI}}^{RSI}(\cdot) = \begin{cases} 1 & if\ RSI_{t-1}(\cdot) > enter\ long \\ 0 & if\ RSI_{t-1}(\cdot) < exit\ long \text{ and } p_{t-1} = 1 \\ -1 & if\ RSI_{t-1}(\cdot) < enter\ short \\ 0 & if\ RSI_{t-1}(\cdot) > exit\ short \text{ and } p_{t-1} = -1 \\ p_{t-1} & else \end{cases}$$

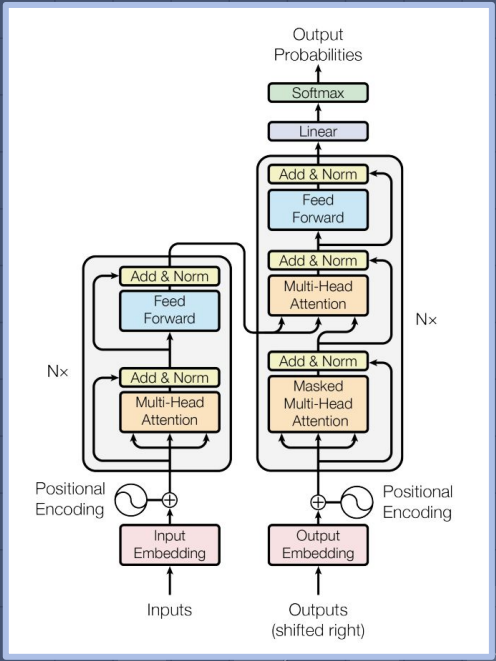
$$\theta_{RSI} = (window, enter\ long, exit\ long, enter\ short, exit\ short)$$

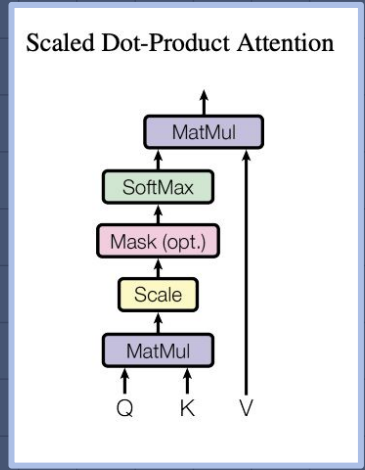
Informer Strategy

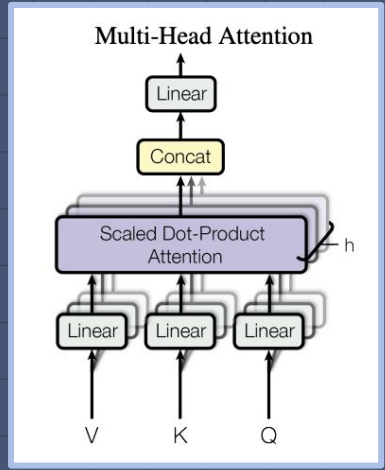
$$s_{\theta_{Informer}}^{Informer}(\cdot) = \begin{cases} 1 & if\ \hat{y}_t > enter\ long \\ 0 & if\ \hat{y}_t < exit\ long \text{ and } p_{t-1} = 1 \\ -1 & if\ \hat{y}_t < enter\ short \\ 0 & if\ \hat{y}_t > exit\ short \text{ and } p_{t-1} = -1 \\ p_{t-1} & else \end{cases}$$

$$\theta_{Informer} = (\Theta_{Informer},\ enter\ long,\ exit\ long,\ enter\ short,\ exit\ short)$$

Transformer



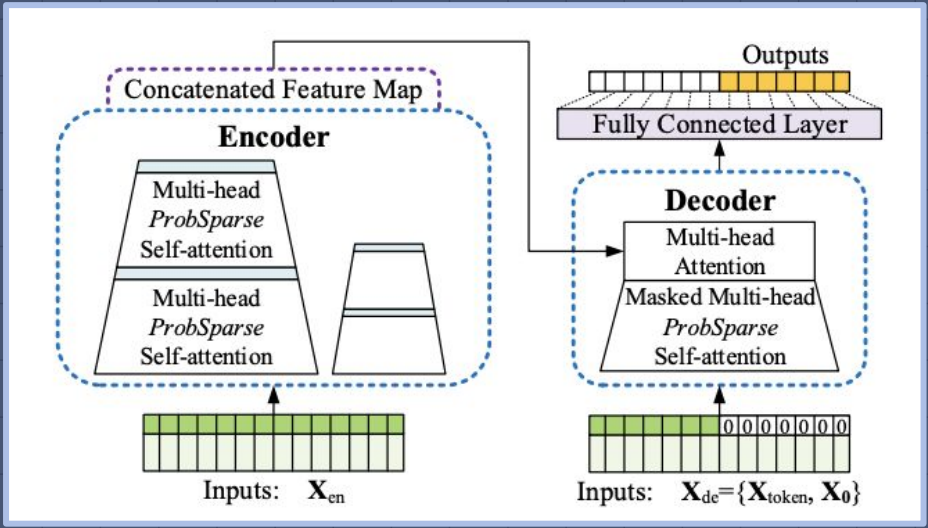




$$\mathcal{A}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = softmax\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d}}\right)\mathbf{V}$$

Vaswani, Ashish et al. (2017). "Attention Is All You Need". In: CoRR abs/1706.03762. arXiv: 1706.03762. url: http://arxiv.org/abs/1706.03762.

Informer



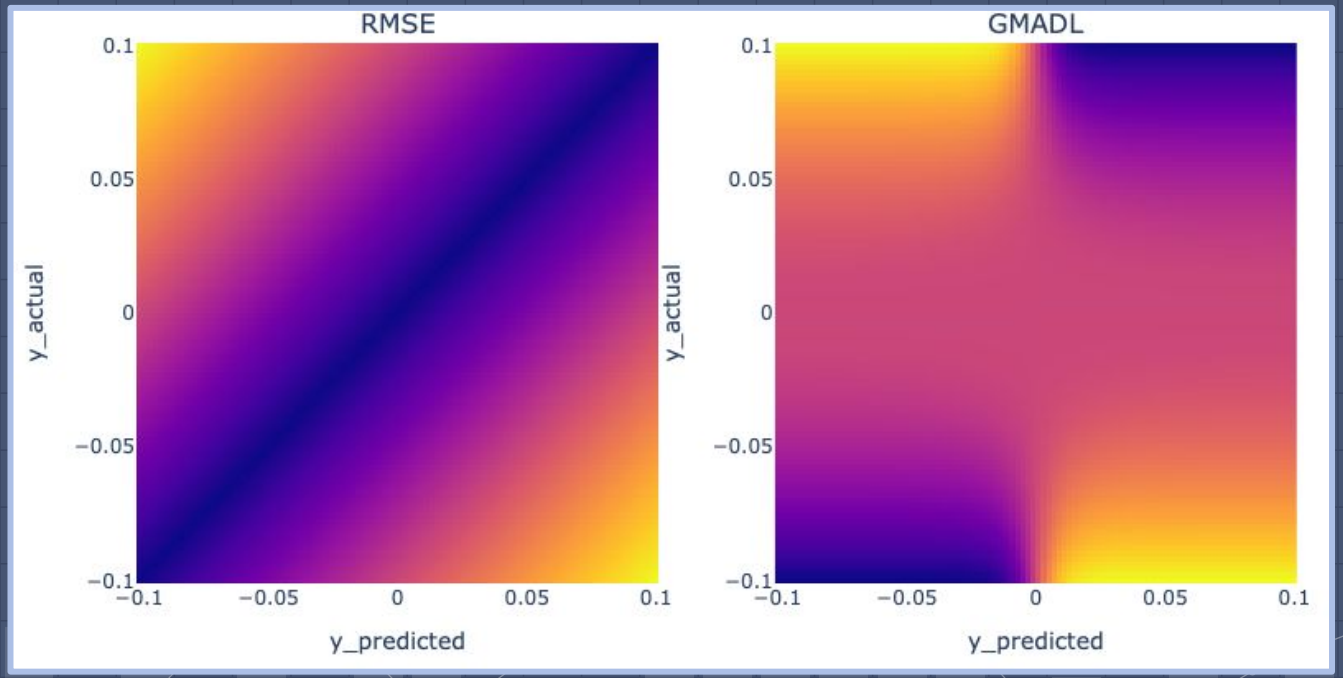
- Improved computation complexity

$$\mathcal{A}'(\mathbf{Q},\mathbf{K},\mathbf{V}) = softmax\left(\frac{\bar{\mathbf{Q}}\mathbf{K}^T}{\sqrt{d}}\right)\mathbf{V}$$

- Self-Attention Distilling
- Generative Inference

Zhou, Haoyi et al. (2021). Informer: Beyond Efficient Transformer for Long Sequence Time-Series Fore-
casting. arXiv: 2012.07436 [cs.LG]. url: https://arxiv.org/abs/2012.07436.

Loss function



Michańków, Jakub, Pawel Sakowski, and Robert Ślepaczuk (2024). "Mean Absolute Directional Loss as a new loss function for machine learning problems in algorithmic investment strategies". In: Journal of Computational Science 81, p. 102375. issn: 1877-7503. doi: https://doi.org/10.1016/j.jocs.2024.102375. url: https://www.sciencedirect.com/science/article/pii/S1877750324001686

Procedure

For each data window, independently:

1. Train model on a training part + search the model hyperparameter space using validation set (min loss)*
2. Evaluate the strategy on validation part, search the strategy hyperparameter space according to IR**
3. Evaluate the strategy with the selected set of the strategy hyperparameters on out-of-sample (test) part.

RESULTS

3

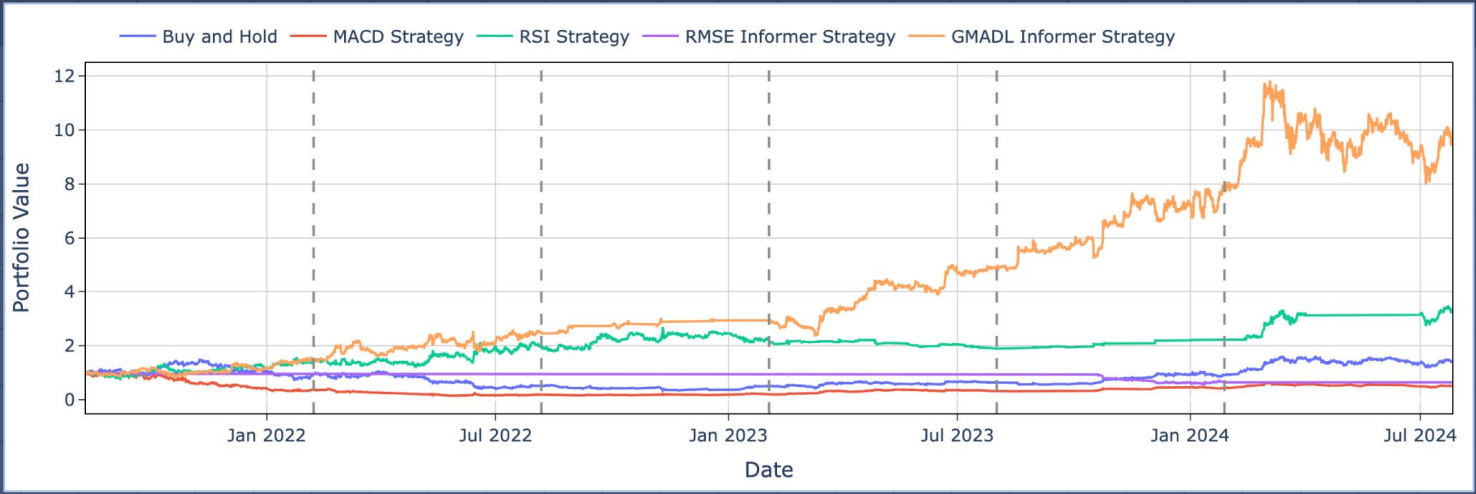
Results - 30 min

Strategy	VAL	ARC	ASD	IR*	MD	IR**	N	LONG	SHORT
Buy and Hold	1.440	13.12%	55.95%	0.235	77.20%	0.040	2	100.00%	0.00%
MACD Strategy	1.952	25.37%	52.36%	0.485	59.24%	0.207	327	52.30%	28.30%
RSI Strategy	**4.542**	**66.77%**	**46.25%**	**1.444**	**39.91%**	**2.415**	**377**	**30.79%**	**28.03%**
RMSE Informer	**2.727**	**40.37%**	**50.47%**	**0.800**	**51.75%**	**0.624**	**34**	**64.40%**	**24.67%**
GMADL Informer	**2.263**	**31.79%**	**36.70%**	**0.866**	**53.35%**	**0.516**	**811**	**35.51%**	**19.59%**

Results - 15 min

Strategy	VAL	ARC	ASD	IR*	MD	IR**	N	LONG	SHORT
Buy and Hold	1.440	13.12%	55.95%	0.235	77.20%	0.040	2	100.00%	0.00%
MACD Strategy	0.468	-22.64%	52.43%	-0.432	83.18%	-0.118	1311	51.80%	31.33%
RSI Strategy	0.800	-7.28%	55.66%	-0.131	66.67%	-0.014	1206	54.02%	43.08%
RMSE Informer	**1.509**	**14.93%**	**34.90%**	**0.428**	**45.54%**	**0.140**	**16**	**15.24%**	**27.60%**
GMADL Informer	**3.296**	**49.65%**	**52.70%**	**0.942**	**47.39%**	**0.987**	**362**	**49.37%**	**37.72%**

Results - 5 min



Strategy	VAL	ARC	ASD	IR*	MD	IR**	N	LONG	SHORT
Buy and Hold	1.440	13.12%	55.95%	0.235	77.20%	0.040	2	100.00%	0.00%
MACD Strategy	0.516	-20.04%	54.14%	-0.370	85.77%	-0.087	2535	50.39%	32.38%
RSI Strategy	**3.341**	**50.34%**	**50.41%**	**0.999**	**29.99%**	**1.676**	**846**	**28.29%**	**33.47%**
RMSE Informer	0.643	-13.88%	15.13%	-0.917	44.61%	-0.285	16	0.00%	9.58%
GMADL Informer	**9.747**	**115.88%**	**54.44%**	**2.129**	**32.66%**	**7.552**	**864**	**44.80%**	**41.51**

All Strategies

Strategy	VAL	ARC	ASD	IR*	MD	IR**	N	LONG	SHORT
Buy and Hold	1.440	13.12%	55.95%	0.235	77.20%	0.040	2	100.00%	0.00%
MACD Strategy (30min)	1.952	25.37%	52.36%	0.485	59.24%	0.207	327	52.30%	28.30%
MACD Strategy (15min)	0.468	-22.64%	52.43%	-0.432	83.18%	-0.118	1311	51.80%	31.33%
MACD Strategy (5min)	0.516	-20.04%	54.14%	-0.370	85.77%	-0.087	2535	50.39%	32.38%
RSI Strategy (30min)	**4.542**	**66.77%**	**46.25%**	**1.444**	**39.91%**	**2.415**	**377**	**30.79%**	**28.03%**
RSI Strategy (15min)	0.800	-7.28%	55.66%	-0.131	66.67%	-0.014	1206	54.02%	43.08%
RSI Strategy (5min)	**3.341**	**50.34%**	**50.41%**	**0.999**	**29.99%**	**1.676**	**846**	**28.29%**	**33.47%**
RMSE Informer (30min)	2.727	40.37%	50.47%	0.800	51.75%	0.624	34	64.40%	24.67%
RMSE Informer (15min)	1.509	14.93%	34.90%	0.428	45.54%	0.140	16	15.24%	27.60%
RMSE Informer (5min)	0.643	-13.88%	15.13%	-0.917	44.61%	-0.285	16	0.00%	9.58%
GMADL Informer (30min)	2.263	31.79%	36.70%	0.866	53.35%	0.516	811	35.51%	19.59%
GMADL Informer (15min)	3.296	49.65%	52.70%	0.942	47.39%	0.987	362	49.37%	37.72%
GMADL Informer (5min)	**9.747**	**115.88%**	**54.44%**	**2.129**	**32.66%**	**7.552**	**864**	**44.80%**	**41.51**

All Strategies - IR**

Strategy	W1	W2	W3	W4	W5	W6
Buy and Hold	-0.064	-1.087	0.007	2.753	14.265	15.103
MACD Strategy (30min)	1.845	1.943	-1.007	0.031	19.407	3.921
MACD Strategy (15min)	-1.559	-0.670	0.082	0.015	5.195	13.925
MACD Strategy (5min)	-1.467	-1.217	0.270	27.660	13.447	1.570
RSI Strategy (30min)	25.897	14.945	-0.005	7.410	14.265	0.000
RSI Strategy (15min)	-1.183	0.263	-0.302	-0.737	13.204	0.374
RSI Strategy (5min)	6.074	3.864	0.649	-2.150	21.237	24.493
RMSE Informer (30min)	-0.064	13.037	-0.009	10.584	14.265	0.069
RMSE Informer (15min)	13.827	0.000	-0.007	2.975	-3.662	0.000
RMSE Informer (5min)	-4.361	0.000	0.000	0.000	-1.989	0.000
GMADL Informer (30min)	0.194	0.322	-1.169	0.187	40.230	20.929
GMADL Informer (15min)	73.474	-0.153	2.085	-1.302	9.554	13.942
GMADL Informer (5min)	7.576	12.844	13.227	30.125	36.394	1.949

All Strategies – IR** ranked

Strategy	W1	W2	W3	W4	W5	W6	TOP 1	TOP 3	TOP 5
Buy and Hold	8	12	6	6	5	3	-	1	2
MACD Strategy (30min)	6	5	12	8	4	6	-	-	2
MACD Strategy (15min)	12	11	5	9	11	5	-	-	1
MACD Strategy (5min)	11	13	4	2	8	8	-	1	2
RSI Strategy (30min)	2	1	8	4	5	11	1	2	4
RSI Strategy (15min)	10	7	11	11	9	9	-	-	-
RSI Strategy (5min)	5	4	3	13	3	1	1	3	5
RMSE Informer (30min)	8	2	10	3	5	10	-	2	3
RMSE Informer (15min)	3	8	9	5	13	11	-	1	2
RMSE Informer (5min)	13	8	7	10	12	11	-	-	-
GMADL Informer (30min)	7	6	13	7	1	2	1	2	2
GMADL Informer (15min)	1	10	2	12	10	4	1	2	3
GMADL Informer (5min)	4	3	1	1	2	7	2	4	5

Sensitivity Analysis - GMADL 5 min

Length of the validation period

Strategy - length of the validation period	IR**
GMADL Informer (5min) - 3 months	0.458
GMADL Informer (5min) - 6 months	**7.552**
GMADL Informer (5min) - 9 months	**3.048**
GMADL Informer (5min) - 12 months	1.088

Choosing top n-th set of strategy parameters

Strategy - Top hyperparameter combinations	IR**
GMADL Informer (5min) - 1	**7.552**
GMADL Informer (5min) - 2	**2.758**
GMADL Informer (5min) - 3	2.331
GMADL Informer (5min) - 4	2.179
GMADL Informer (5min) - 5	2.150

Research Questions:

- *Is it possible to create an algorithmic strategy for trading Bitcoin, that is more efficient than Buy&Hold approach?*
- *Does signal from Informer model allow to create strategies that are more efficient on trading Bitcoin than strategies based on technical indicators?*
- *Does usage of higher frequency data allow to create more efficient strategies?*
- *How does selection of the machine learning model loss function influence the strategy performance?*

THANKS!

Any questions?

You can find me at:

- filip.stefaniuk@gmail.com
- linkedin.com/in/filipstefaniuk