

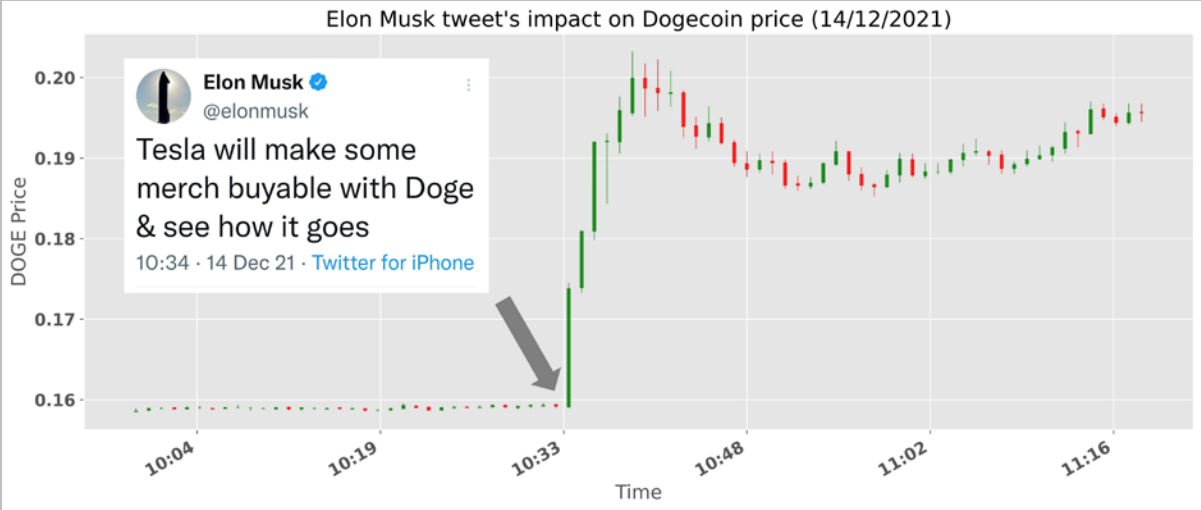
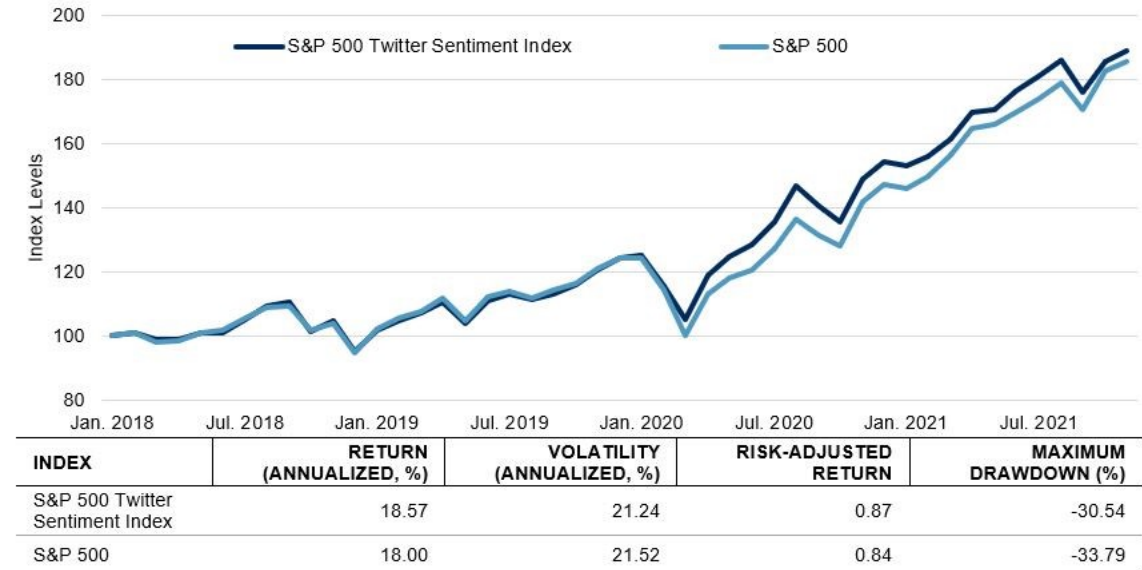
Neural Networks for Sentiment Analysis in Cryptocurrency Market

Baptiste PROVENDIER
CID: 01553706

Introduction

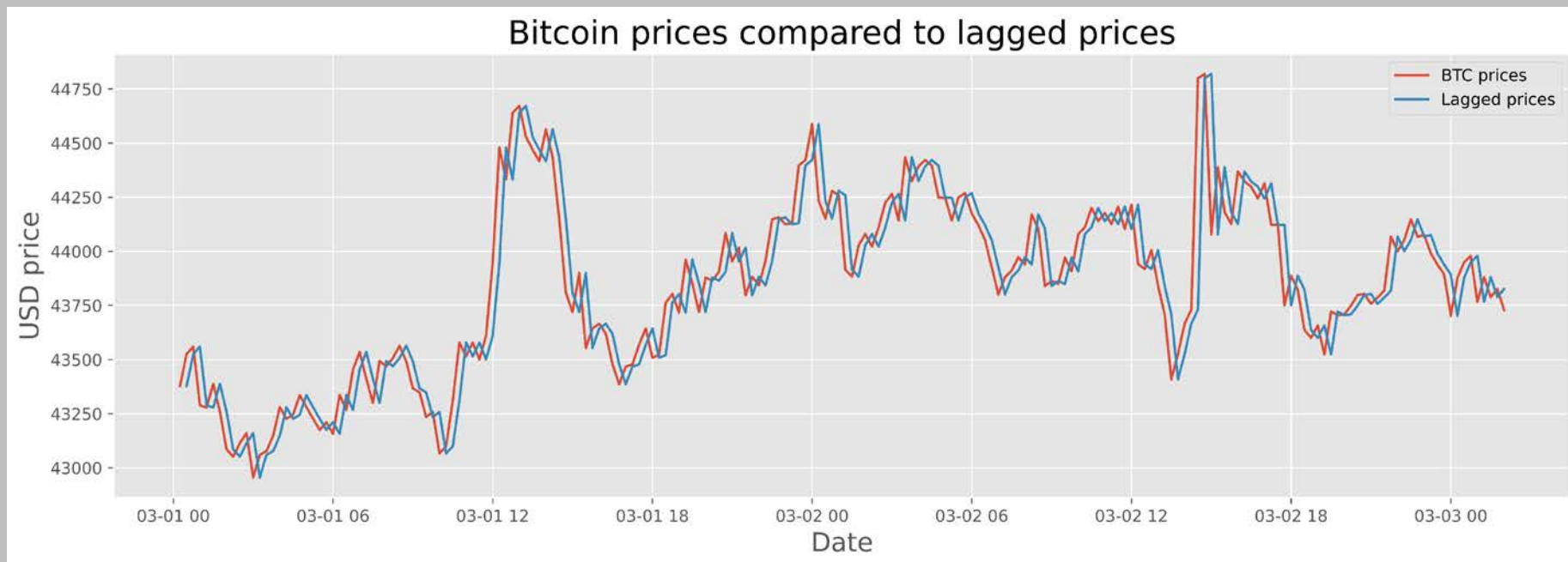
“the question is no longer [...] whether investor sentiments affect stock prices but rather how to measure investor sentiments and quantify its effects” – Malcolm Baker

Exhibit 2: Back-Tested Performance of the S&P 500 Twitter Sentiment Index versus the S&P 500



Motivations & Objectives

- Autonomous system with:
- **Data collection**
 - **Sentiment analysis**
 - **Price prediction**
 - **Trading strategy**



MSE: 0.3%
Correlation: 0.99

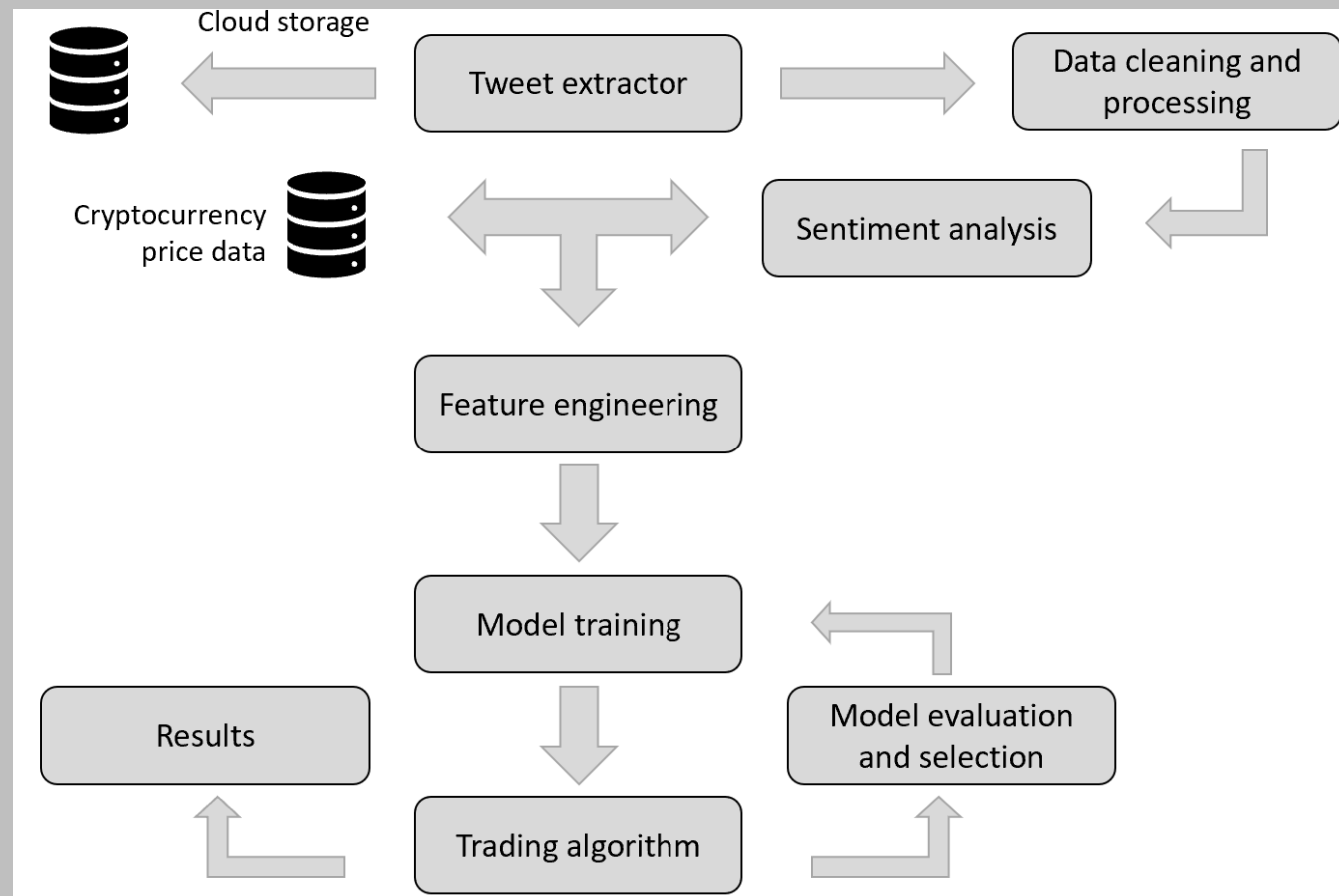
Return on investment using trading strategy: -13% in two days
Accuracy: 44%

System design

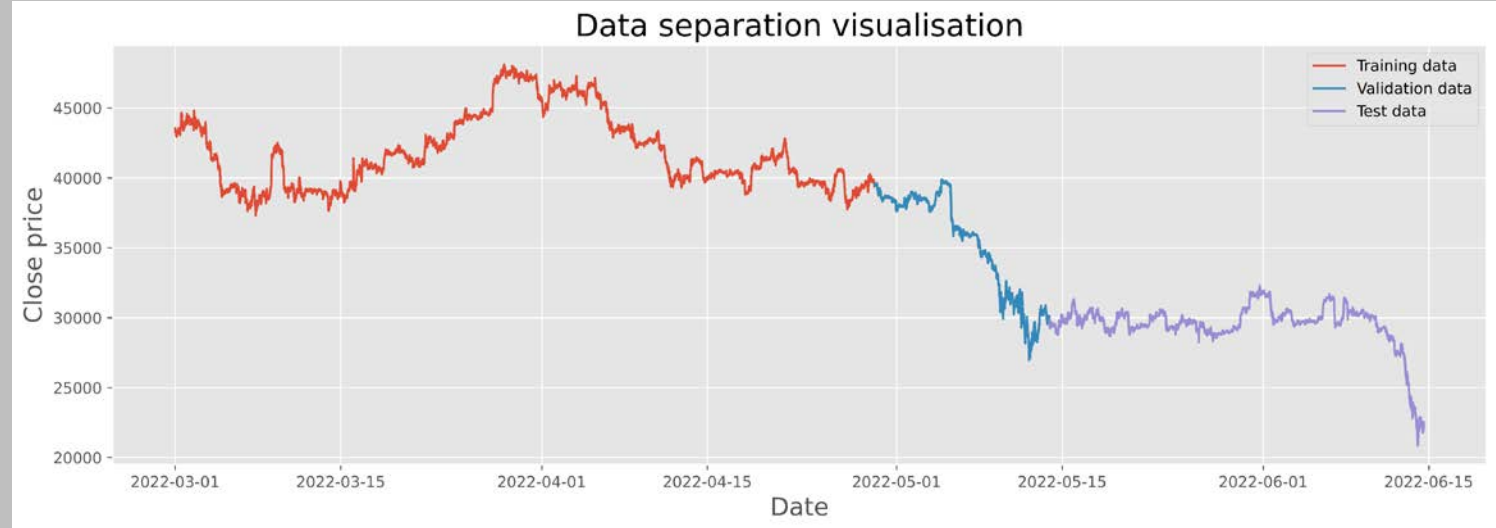
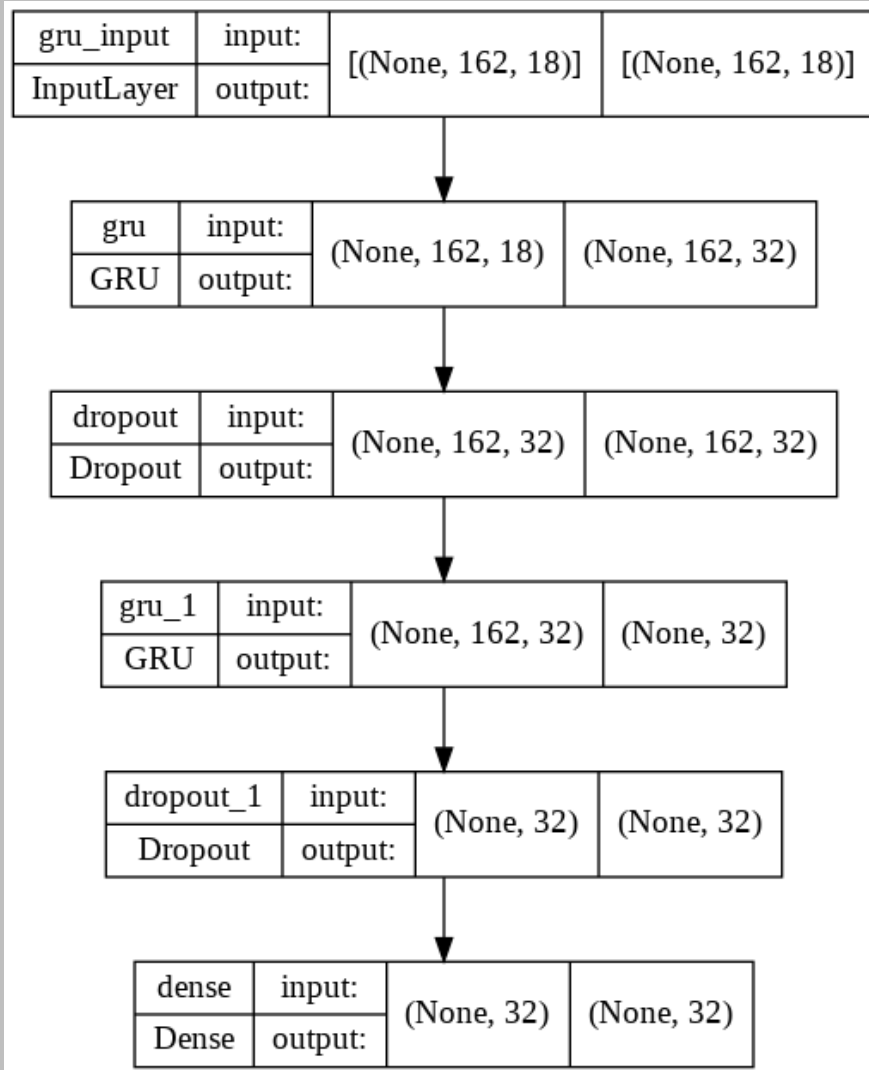
Sentiment features:

- Number of tweets
- Additive sentiment score
- Mean sentiment score
- Polarity sentiment score
- Number of positive tweets
- Percentage of positive tweets

Raw tweet	ðŸ†šðŸ†š· Rio de Janeiro will allow residents to pay property taxes in #Bitcoin & #cryptocurrency beginning in 2023.
Cleaned tweet	rio de janeiro will allow residents to pay property taxes in bitcoin cryptocurrency beginning in 2023



Implementation

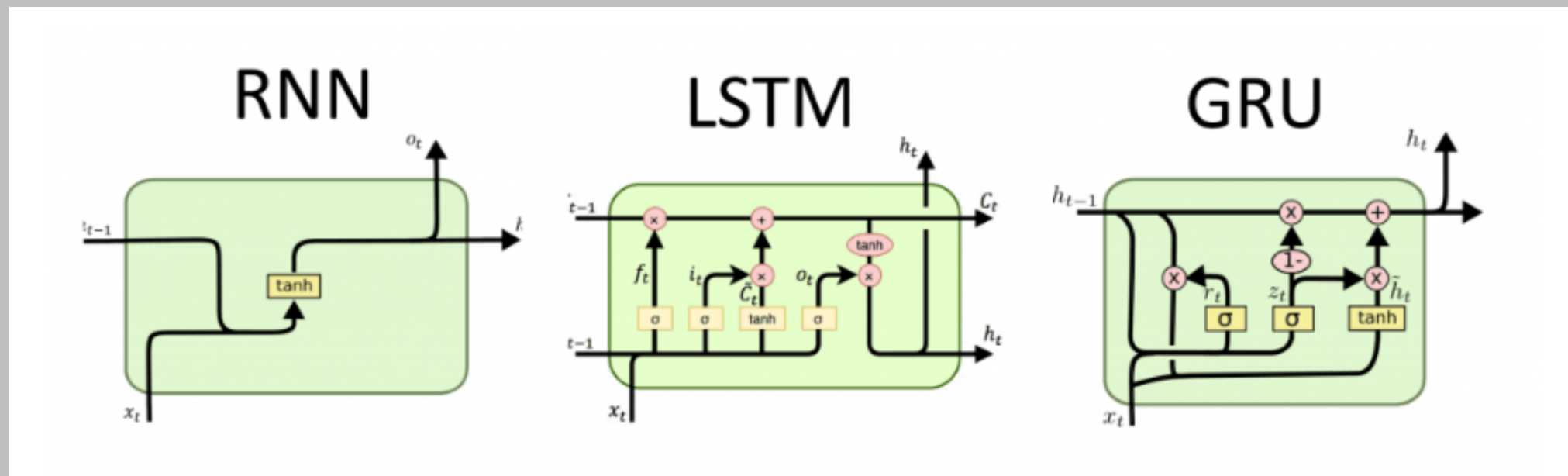
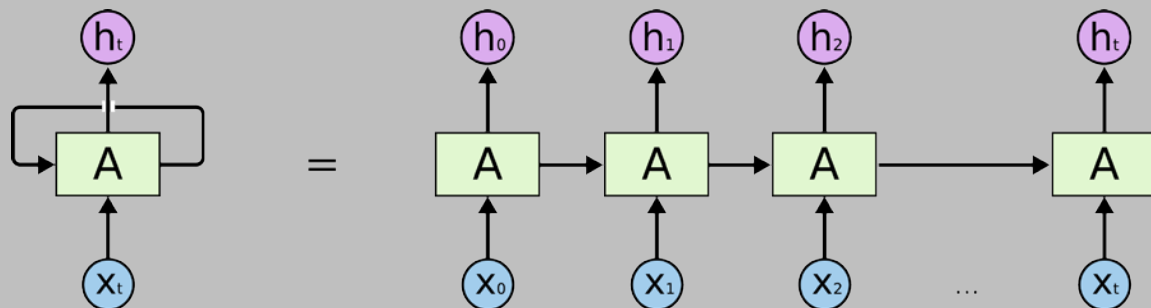


Hyperparameters:

- Sequence length: 24
- Batch size: 32
- Hidden size: 32
- Number of layers: 2
- Dropout rate: 0.1
- Learning rate: 0.01
- Type of model: GRU

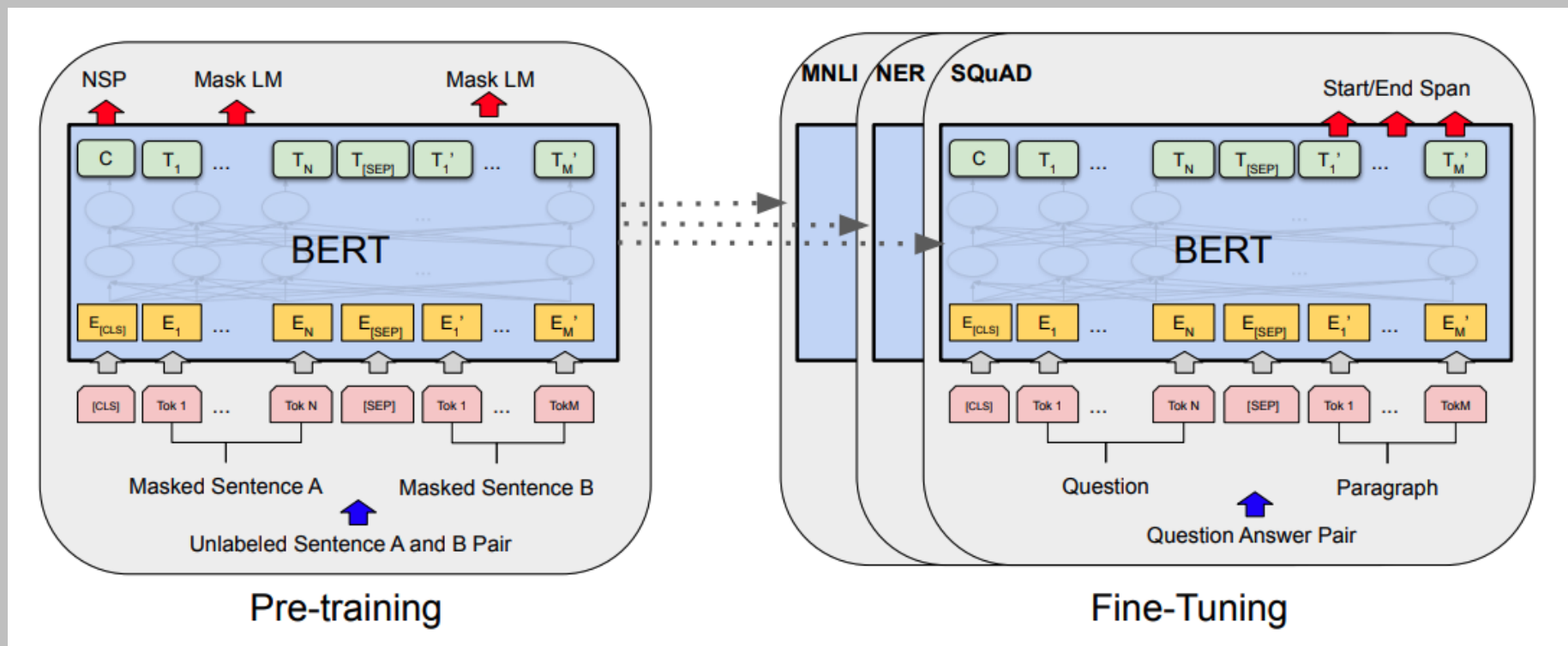
Background: Neural networks

Unrolled recurrent neural network



Background: Sentiment analysis

BERT: Bidirectional Encoder Representations from Transformers – Google, 2018



Background: Performance metrics

Technical Indicators used:

- **Relative Strength Index**
- Stochastic Oscillator
- Williams Percentage Range
- Moving Average Convergence Divergence
- **On Balance Volume**

Traditional error metrics:

- **Mean squared error**
- Root mean squared error
- Mean average error
- Mean average percentage error
- **Accuracy**
- F1-score

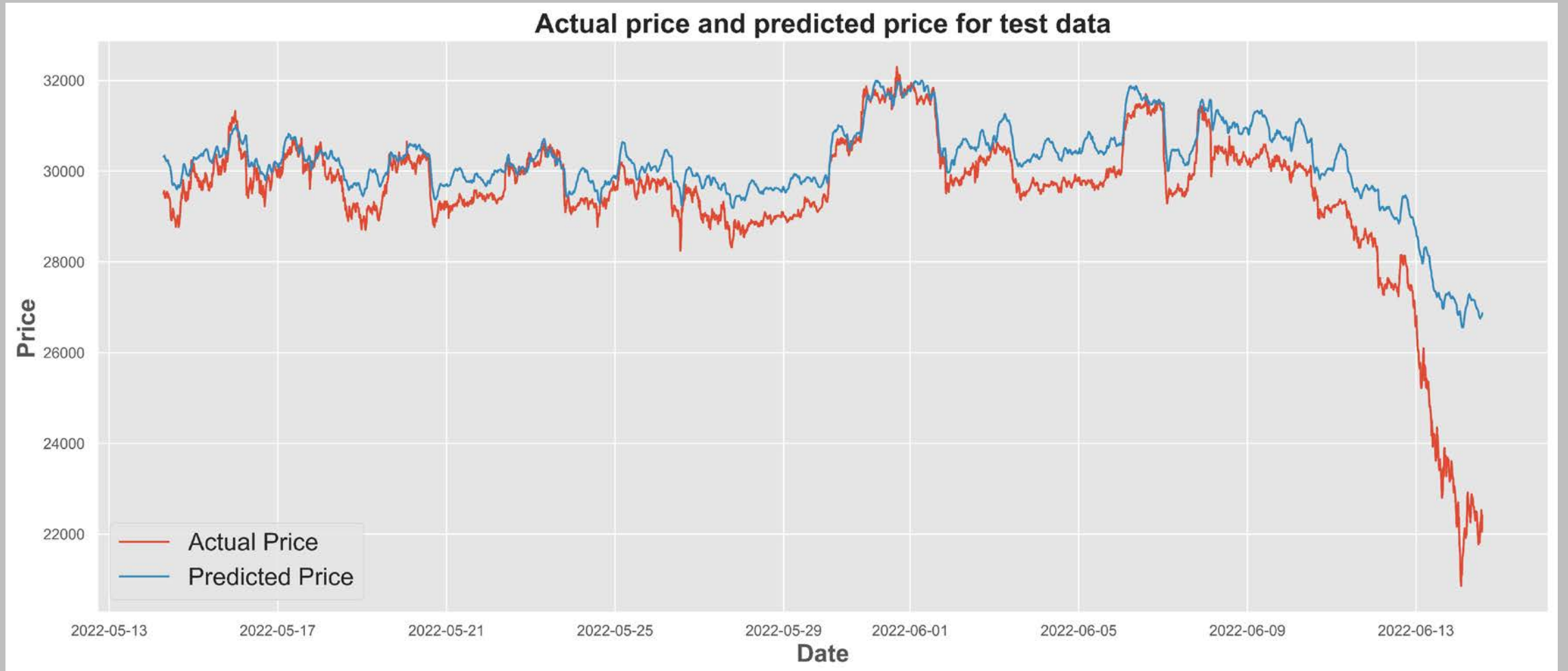
Financial evaluation metrics:

- **Return on investment**
- **Sharpe ratio**
- Value at risk





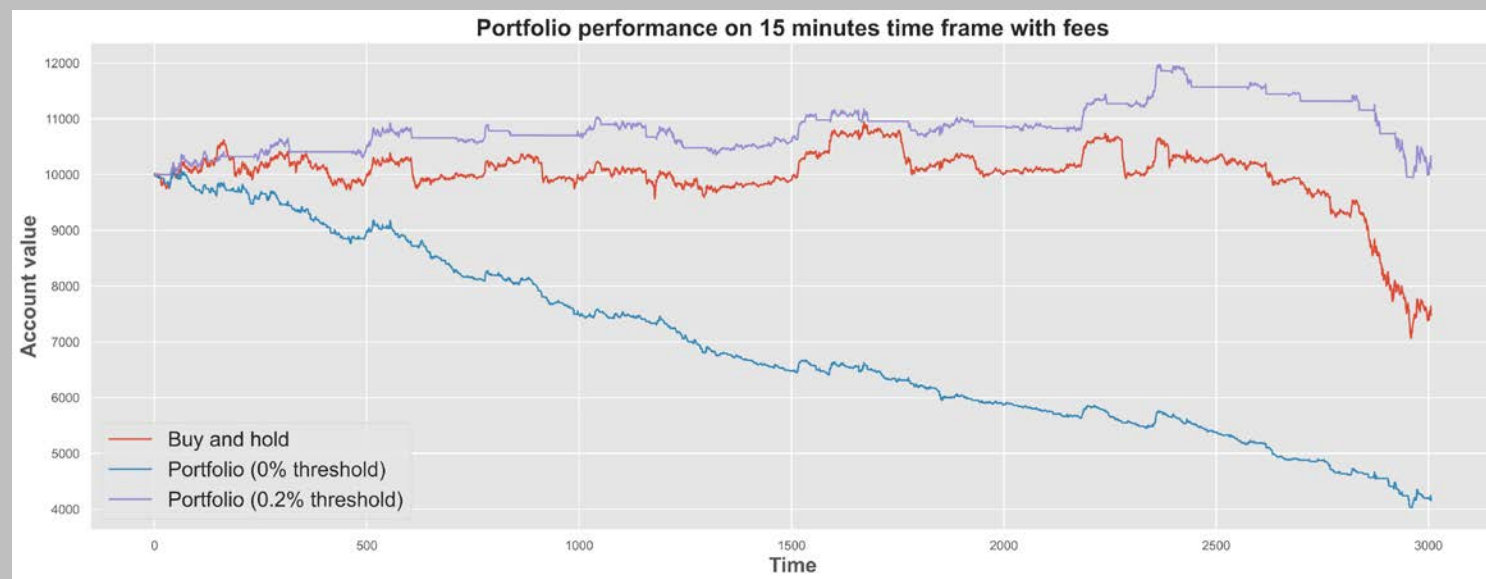
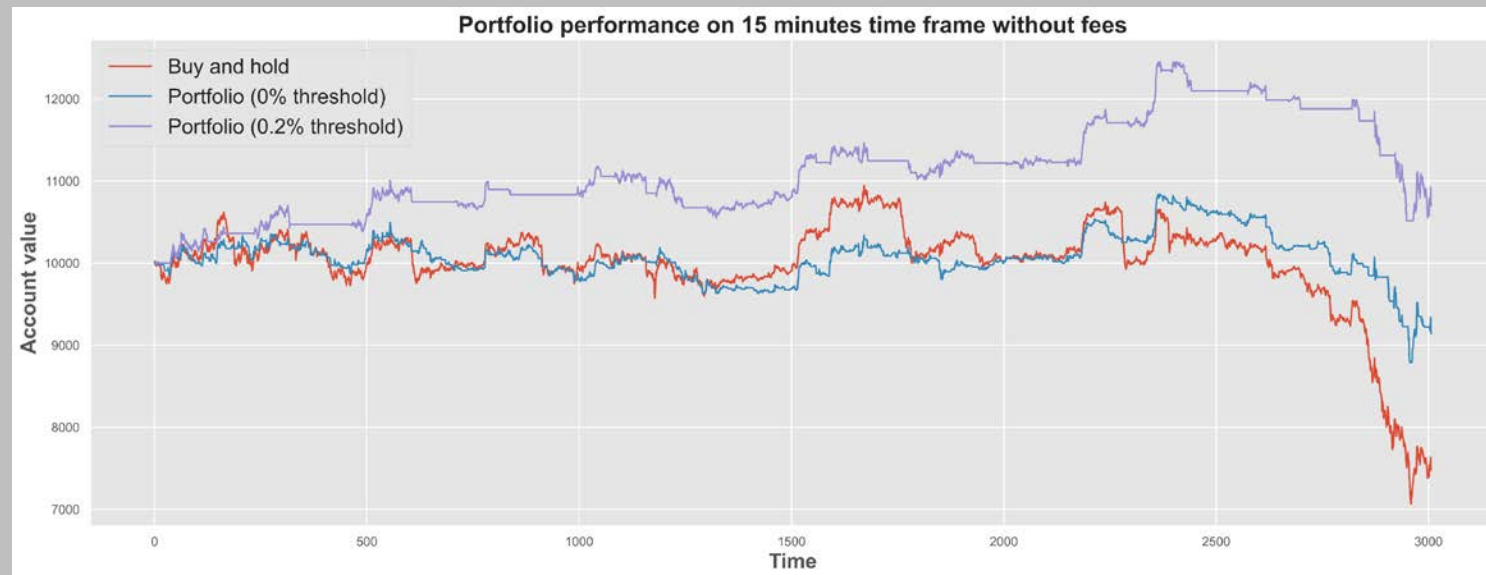
Results



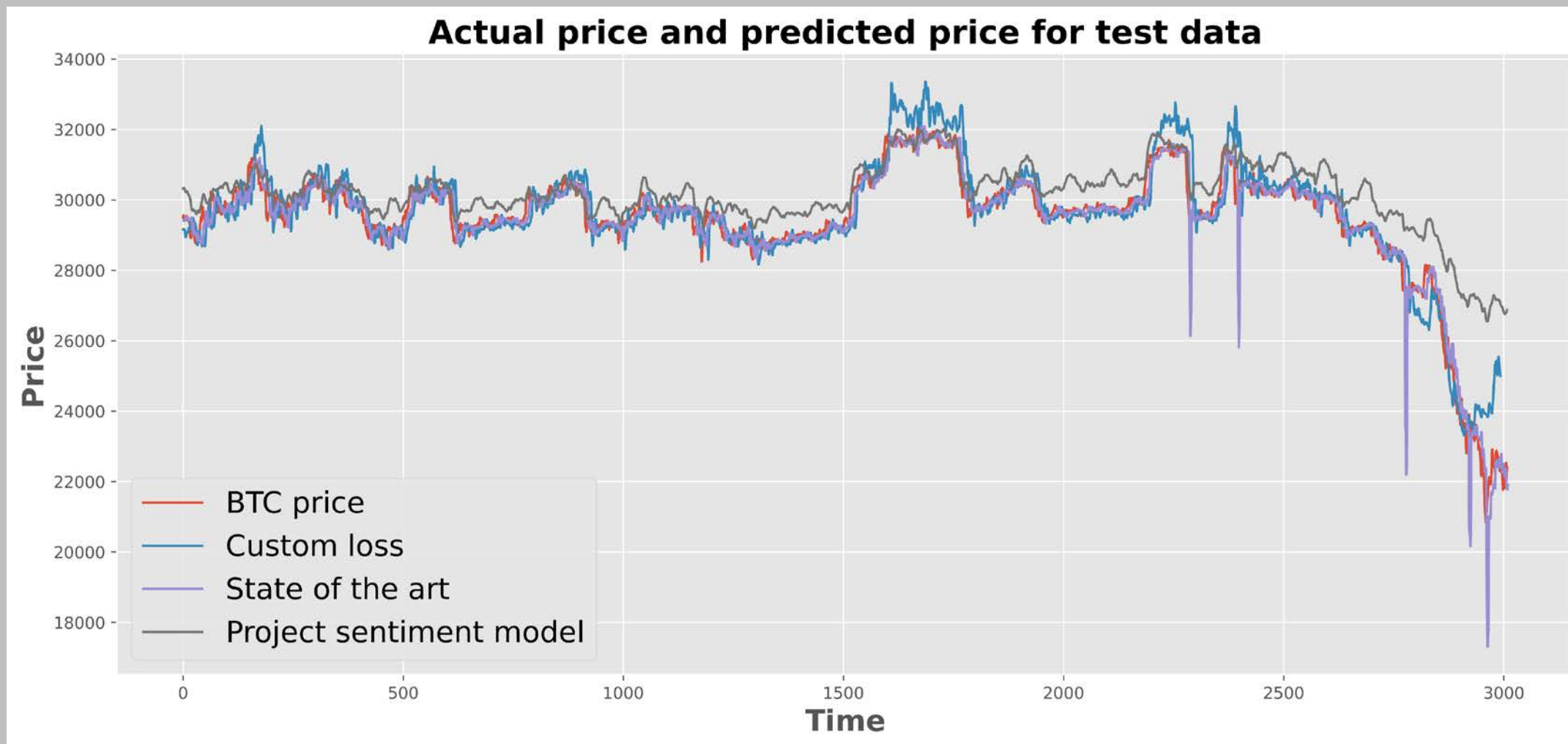
Results

	ROI	Sharpe ratio	Number of trades
Buy and hold	-26%	-0.36	-
0% threshold	-7.1%	-0.17	396
0.2% threshold	+8.6%	+0.16	28

	ROI	Sharpe ratio	Number of trades
Buy and hold	-26%	-0.36	-
0% threshold	-58%	-1.72	396
0.2% threshold	+2.8%	+0.04	28



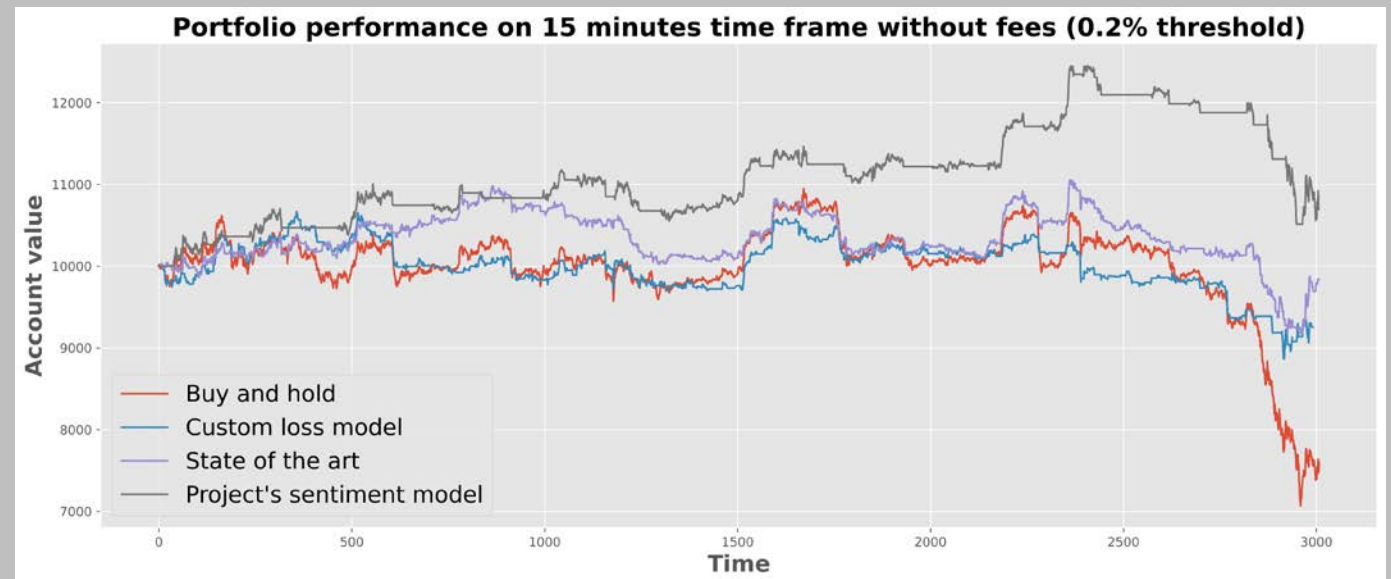
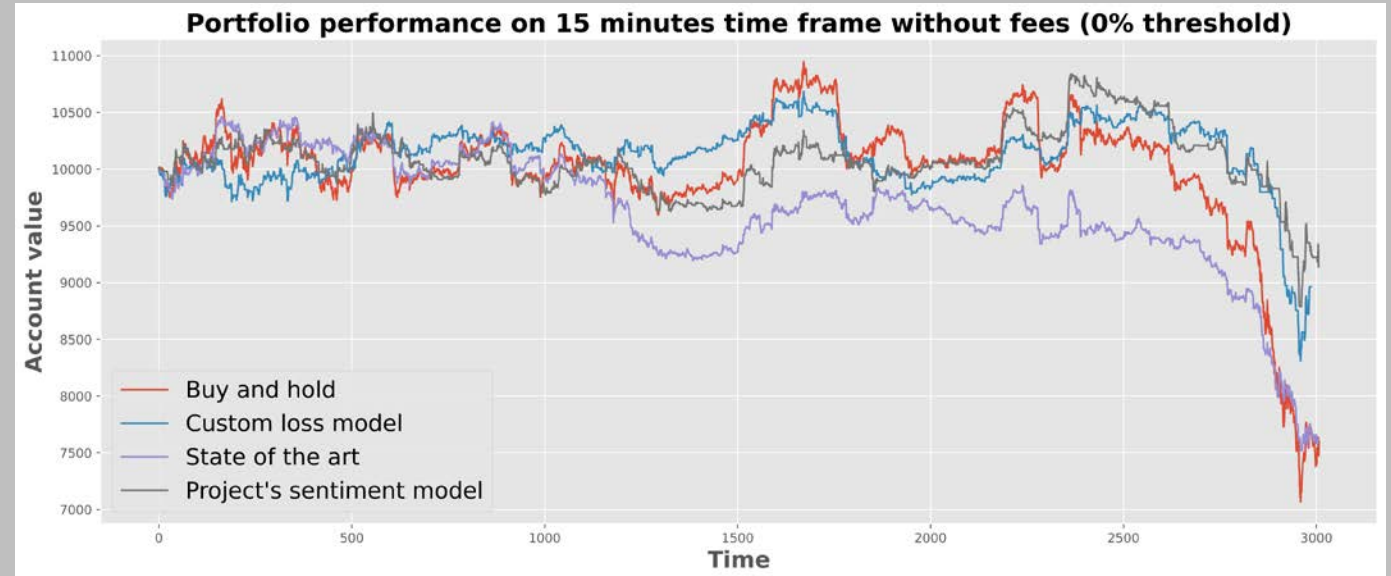
Evaluation to previous methods



Evaluation to previous methods

Model	0% threshold	0.2% threshold
Custom loss	443	237
State of the art	840	372
Project's sentiment model	396	28

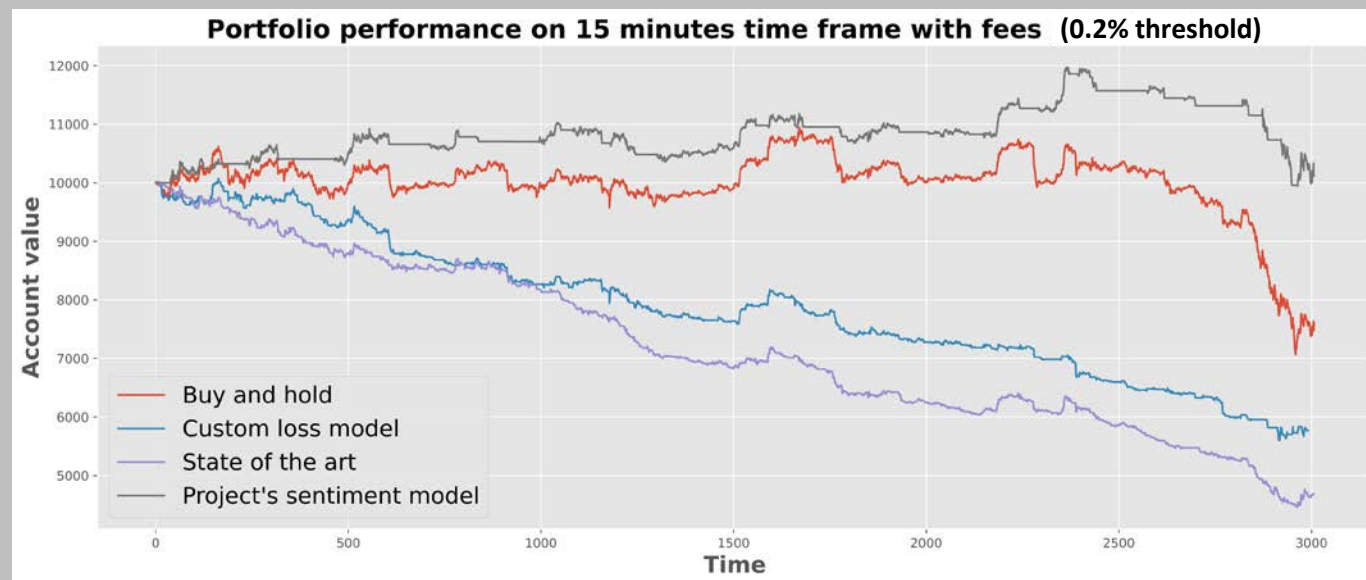
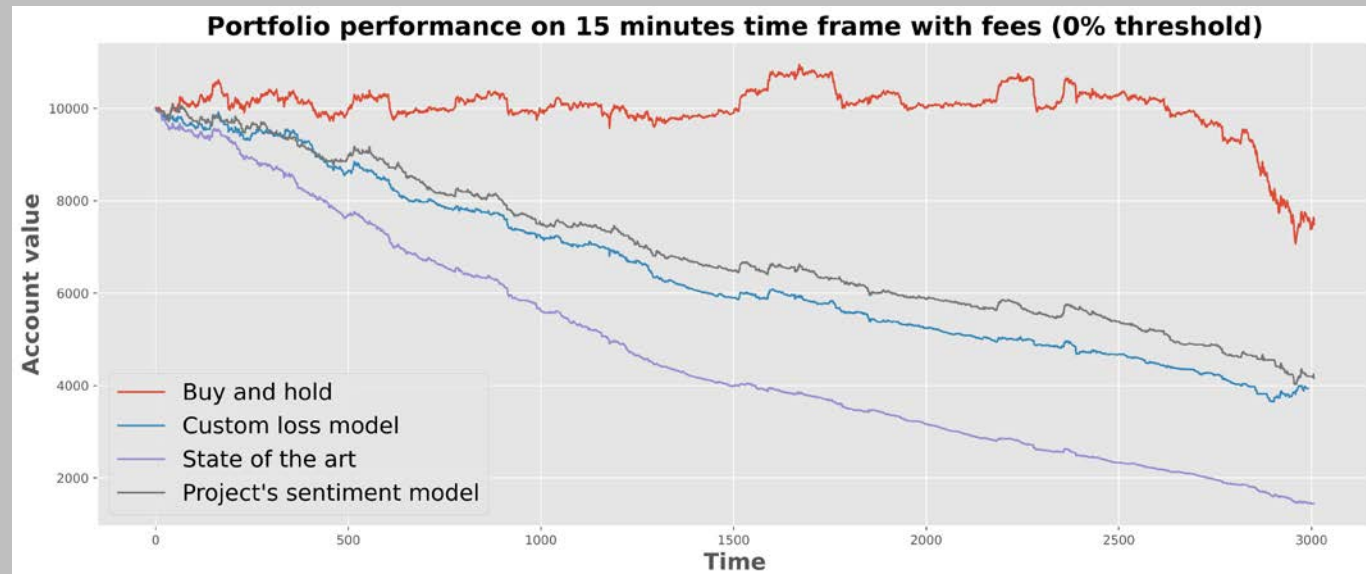
Number of trades made for different models and threshold values



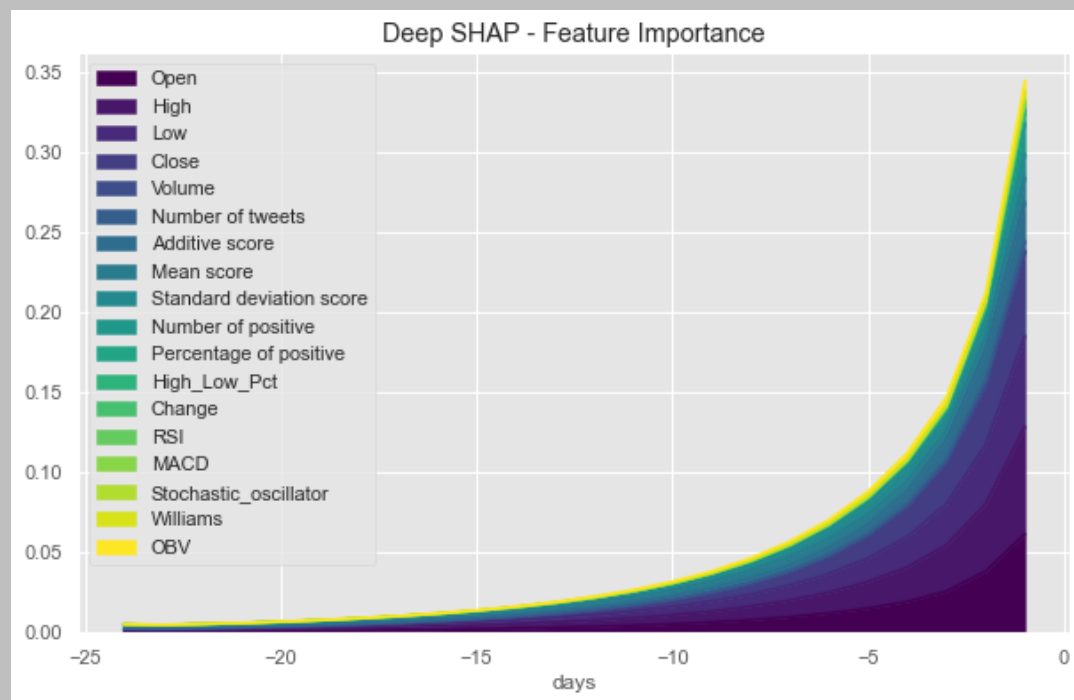
Evaluation to previous methods

Model	0% threshold	0.2% threshold
Custom loss	443	237
State of the art	840	372
Project's sentiment model	396	28

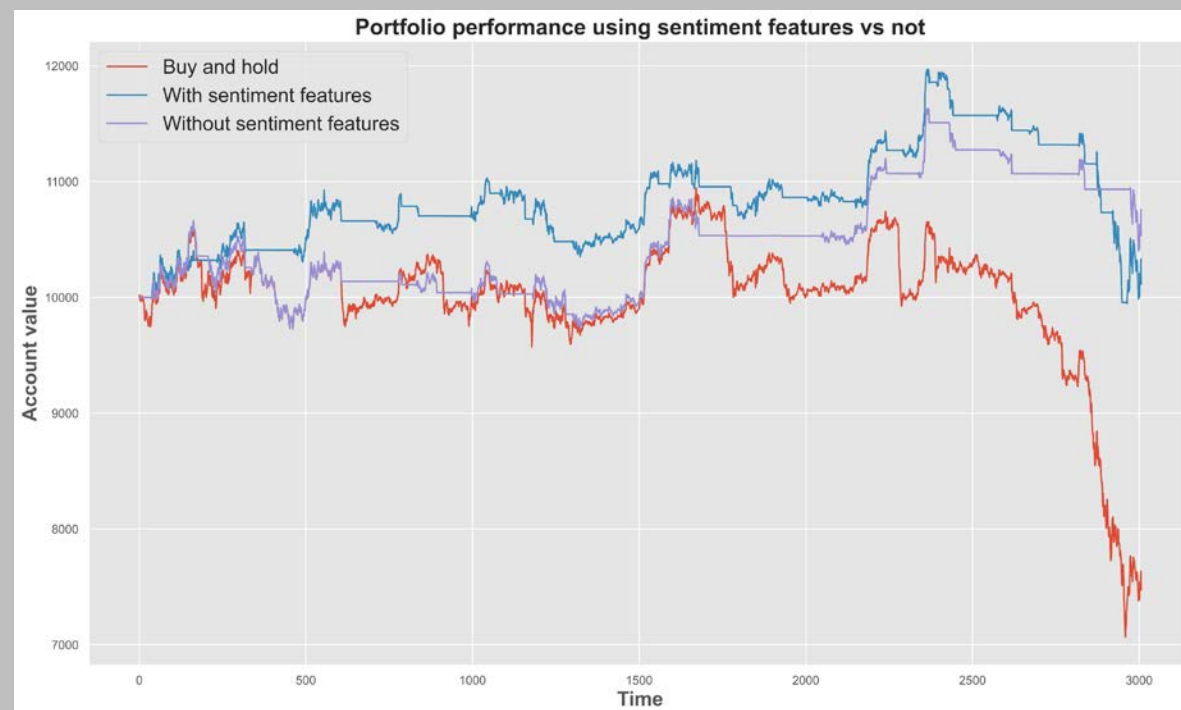
Number of trades made for different models and threshold values



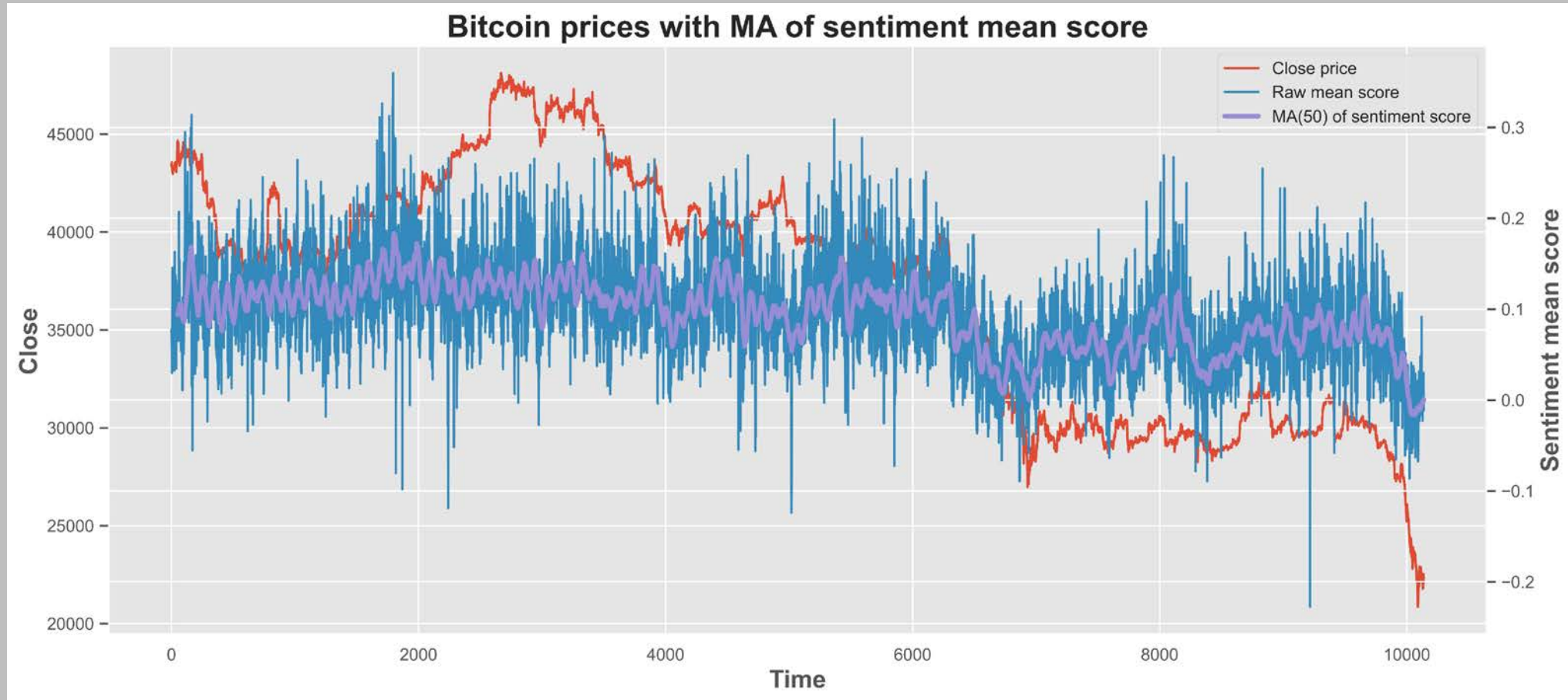
Sentiment importance



Sentiment features total 26%
of the prediction's importance



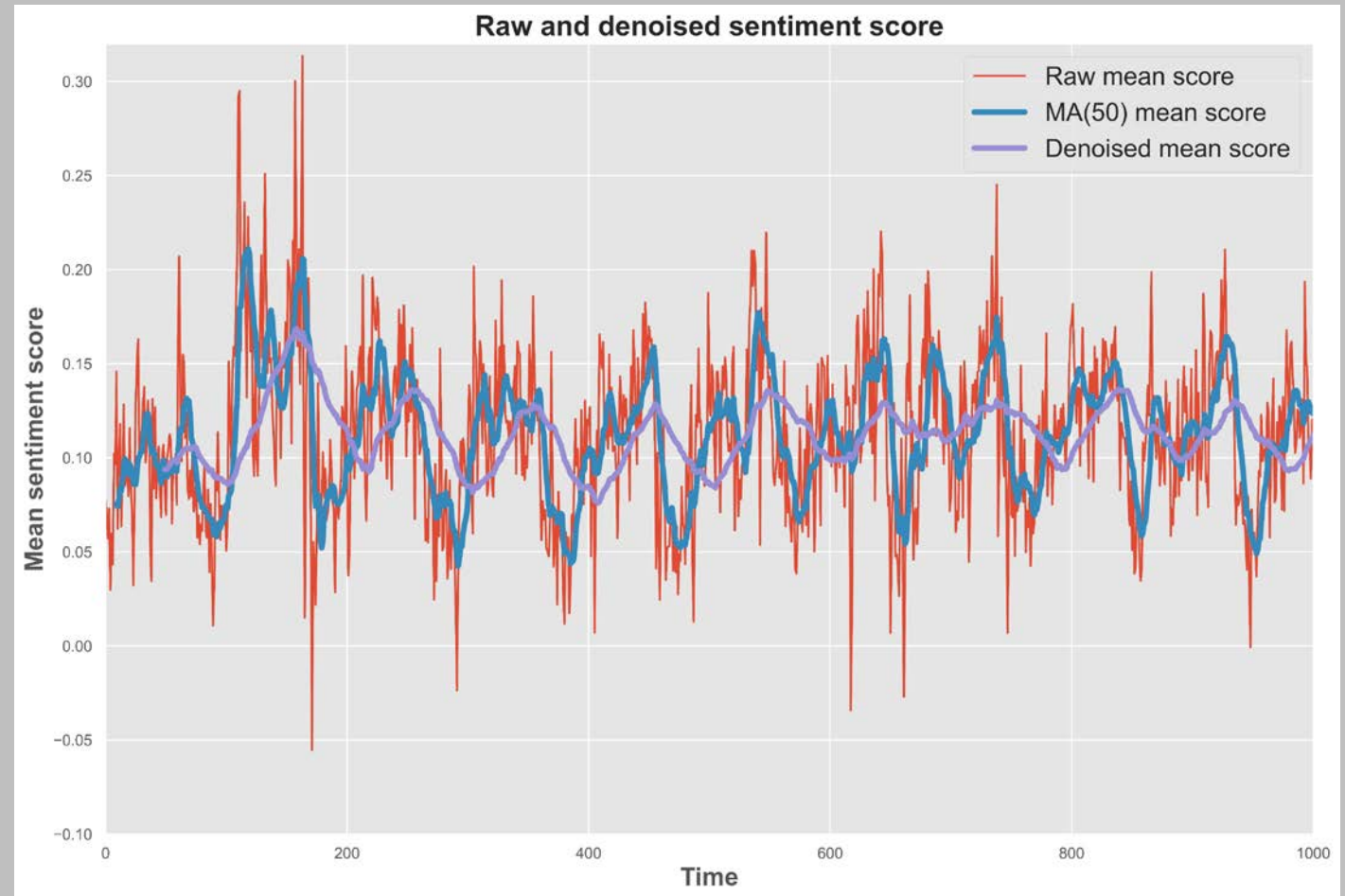
Conclusion



Future work

Improvements:

- More robust trading strategy
- Improve quality and quantity of the collected sentiment data
- Denoise sentiment data
- Loss function designed for trading purposes
- Real-time implementation



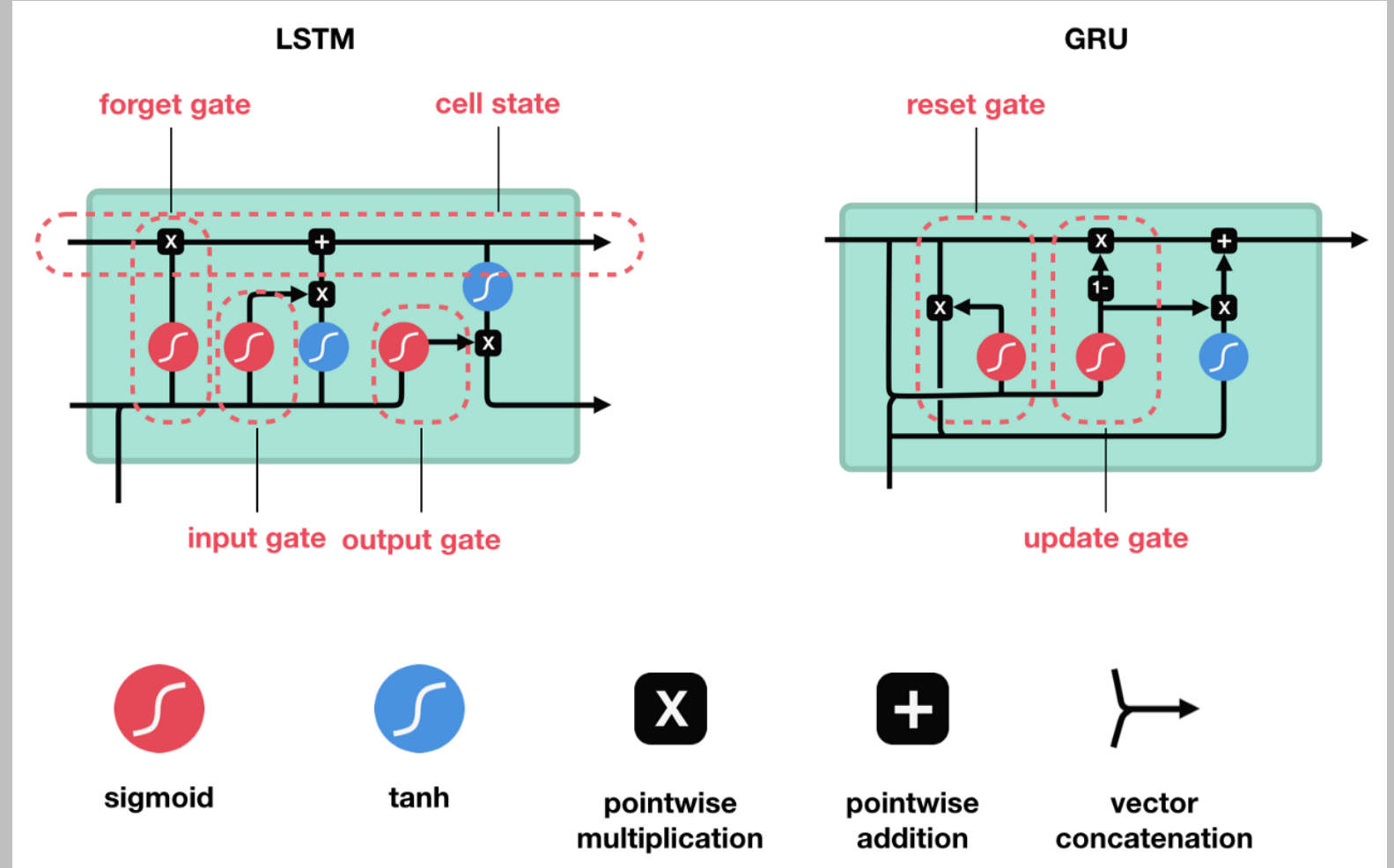
Questions

Code and report available on:
<https://github.com/bprovendier/NN-for-Sentiment-Analysis>



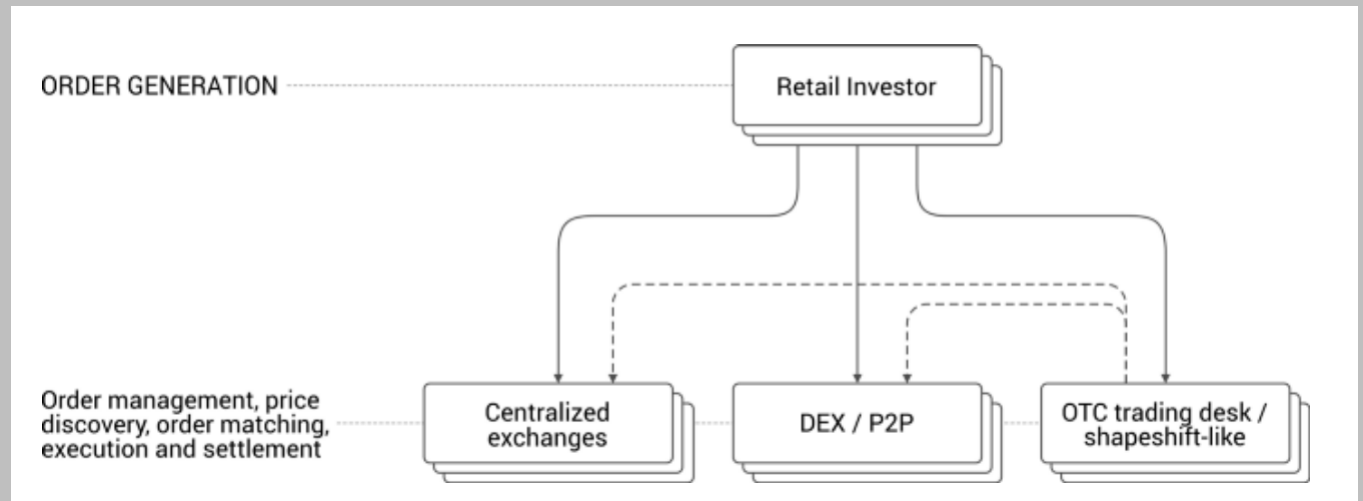
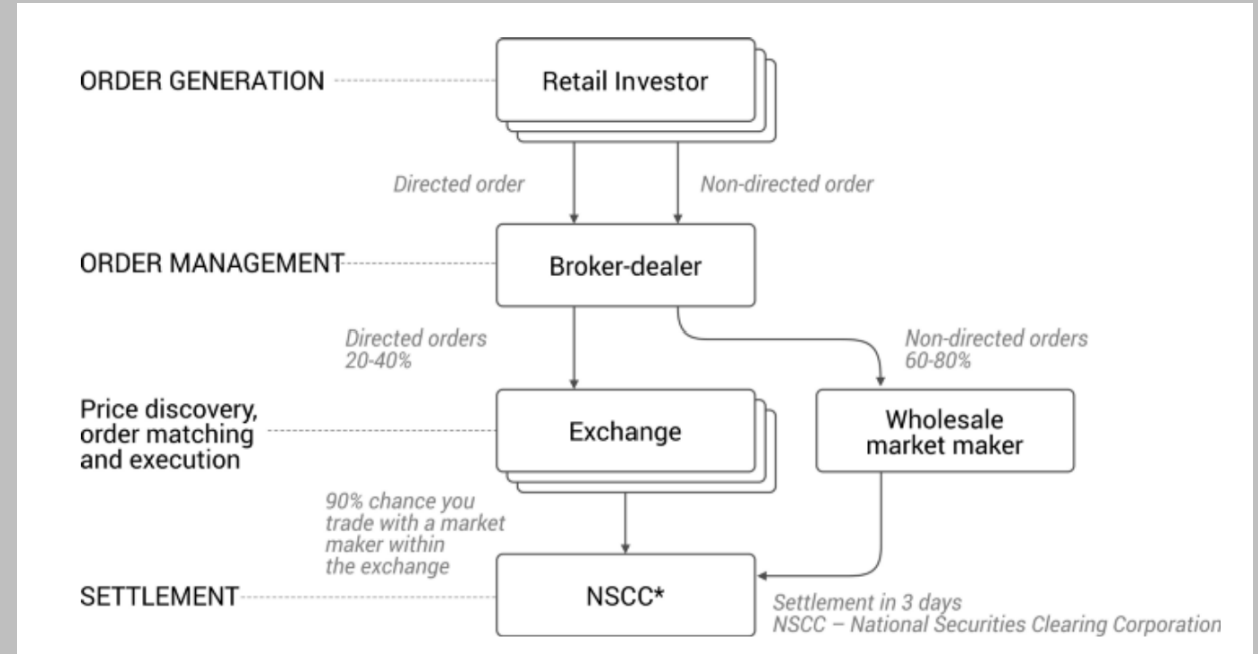
LSTM and GRU architecture

Forget gates (LSTM) and reset gate (GRU) help regulate the learning gradient and get rid of the unnecessary information in order to prevent fast decaying gradient.



Why use crypto?

- Inefficient markets (news travel slower, less participants, and more illiquid)
- More volatile
- No intrinsic value, purely driven by speculation



Hyperparameter selection

