The International Workshop on Big Data in Emergent Distributed Environments (BiDEDE 2024)

PREDICTING BITCOIN PRICE MOVEMENT THROUGH SENTIMENT ANALYSIS: A COMPREHENSIVE STORY

Hanh Nguyen Phuong
Asefeh Asemi
Mutaz Alshafeey
Institute of Information Technology
Corvinus University of Budapest

AGENDA

INTRODUCTION

- Supply and demand interactions
- Volatility and Investor Behavior
- The Role of Twitter
- Sentiment Analysis (SA)

RESEARCH QUESTIONS

Main Aim: The study focuses on exploring the effect of sentiment derived from Twitter on the price movement of Bitcoin and find the optimal time to lead the best prediction.

01

To what extent does sentiment expressed in Tweets significantly impact on Bitcoin price movement?

02

Which prediction model demonstrates the most accurate prediction for Bitcoin price movement?

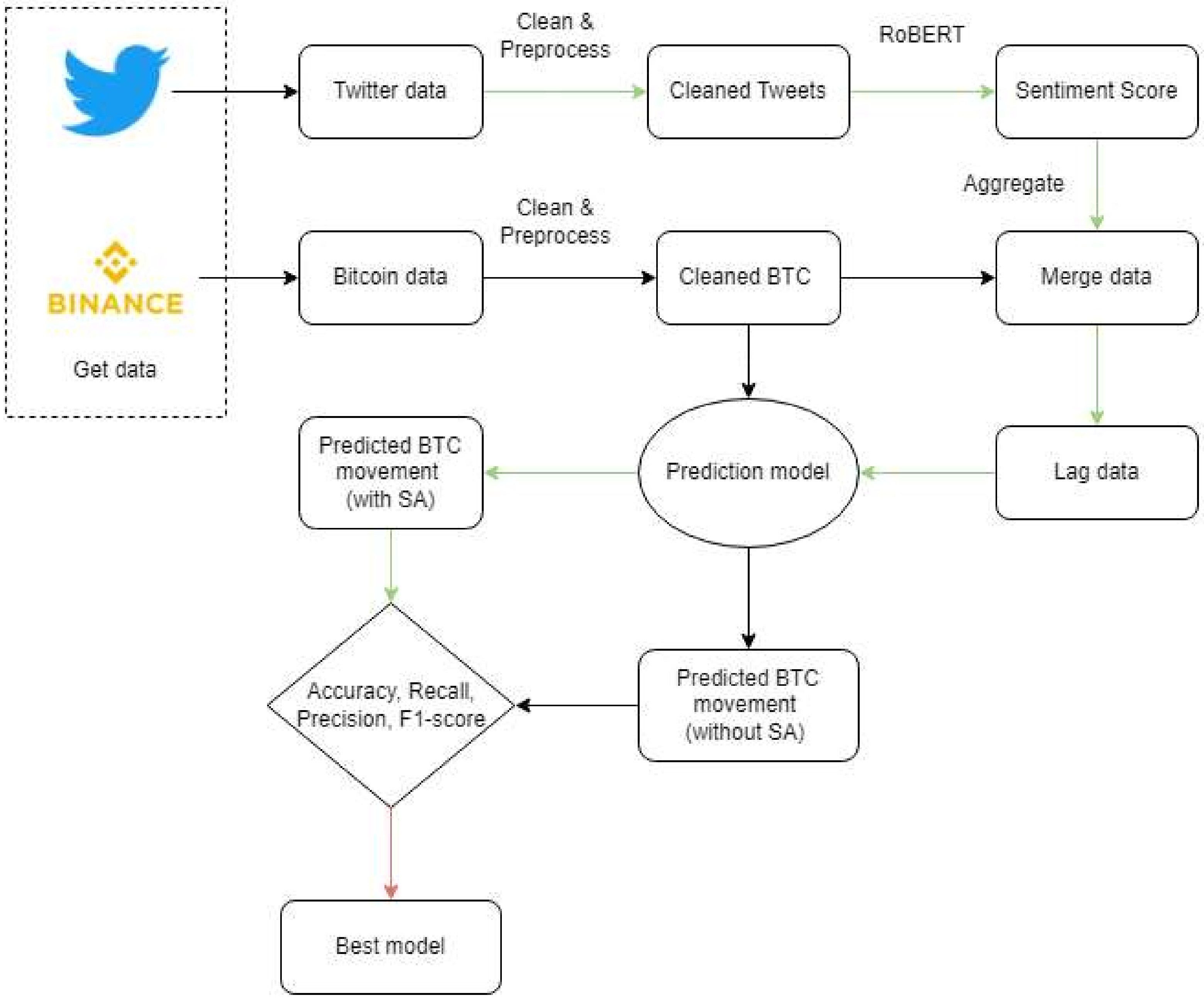
03

What is the optimal time lag that significantly influences the accuracy of Bitcoin price movement?

LITERATURE REVIEW

Research	Key Findings and Research gap
Ye et al, 2022; Low et al, 2024	Inclusion of sentiment data enhances prediction model performance.
Bidirectional Encoder Representations from Transformers (BERT) (Devlin et al, 2018)	BERT method enhances context understanding and offers a pre-trained model adaptable for various downstream tasks.
RoBERTa Variants with different domain-specific knowledge	Two RoBERTa variants trained on large Twitter datasets introduced by Jose et al, 2022, and Juan et al, 2021. There's a lack of direct effectiveness comparison in previous studies.
Trend in Cryptocurrency Price Prediction (Murray et al, 2023; Jaquart et al, 2022)	DL models, particularly LSTM, GRU, and BiLSTM, are favored due to their capacity to capture complex data relationships.
Time lags concept (Critien et al, 2022)	Previous studies used defined time lags. But no studies investigate how time lags affect the prediction models

METHODOLOGY (1/3)



Data Preparation to Forecast by Twitter Sentiment (Source: Author)

Problem Framing: Predicting Bitcoin price movement as binary classification.

SA and non-SA scenarios: Utilizing sentiment analysis (SA) as a multivariate forecasting method, non-SA as a univariate forecasting method.

Sentiment Scores: Extracted from RoBERTa models introduced by Jose et al. (2022) and Juan et al. (2021) for Twitter data.

Statistical Methods: Employing Augmented Dickey-Fuller test (ADF) and Autocorrelation for analysis.

Prediction Models: Three neural network (NN) models utilized - LSTM, GRU, and BiLSTM - after data preparation.

METHODOLOGY (2/3)

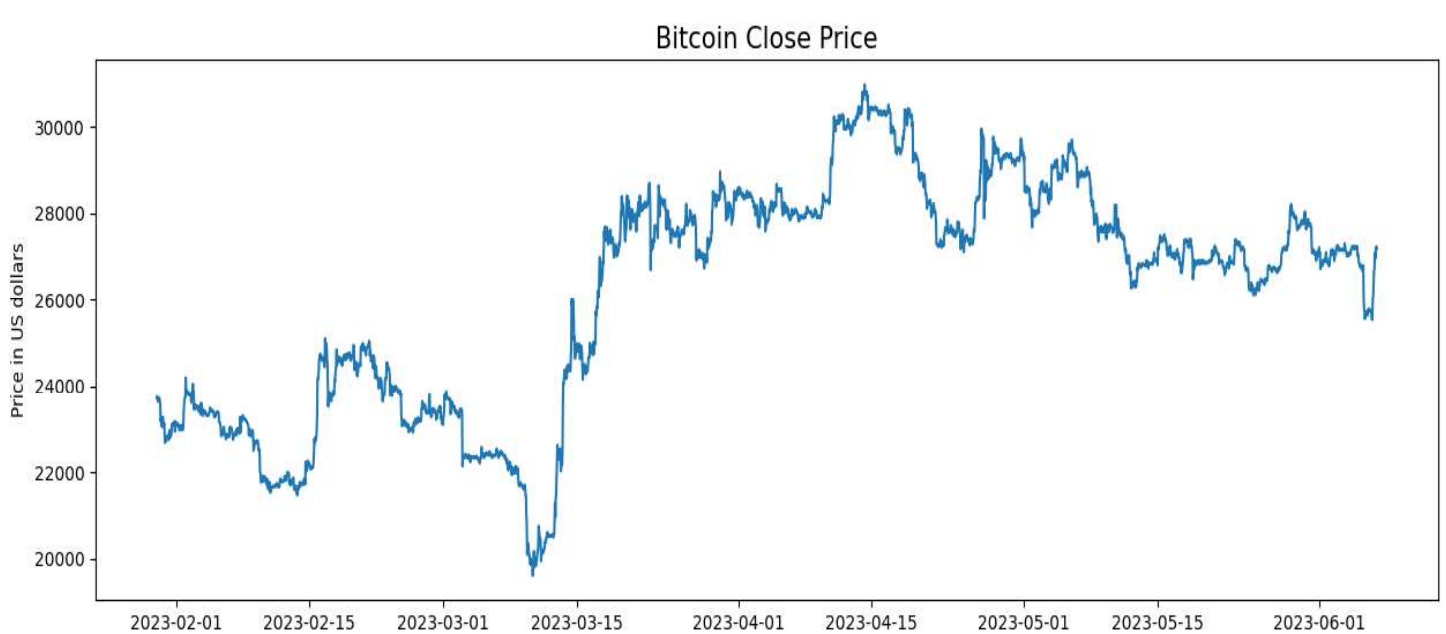
DATA ACQUISITION AND PROCESSING

Bitcoin Price: Gathered from Binance API at 1-hour intervals.

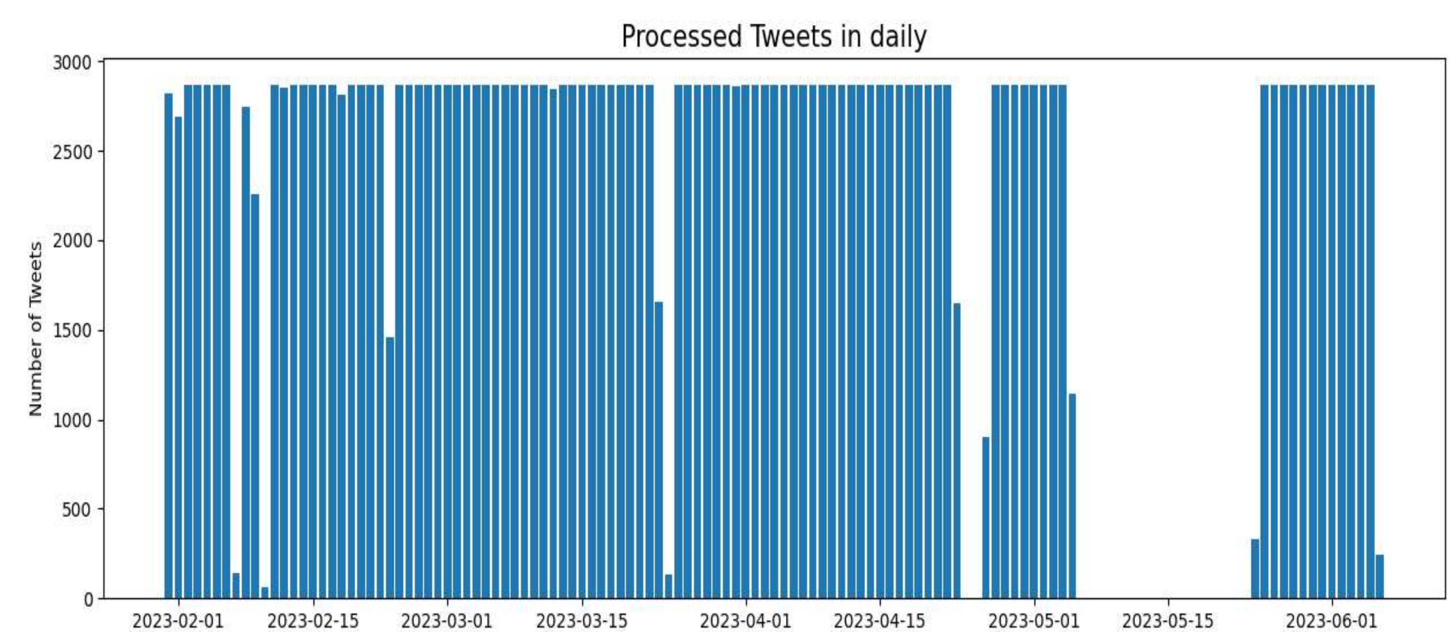
Real-time Twitter Data: Retrieved via tweepy Python library from Twitter API v2, filtering for English tweets containing "btc" or "bitcoin".

Twitter Dataset: Initially 285,405 tweets, reduced to over 285,000 after cleaning.

Timeframe: Data spans January 31, 2023, to June 6, 2023.



Bitcoin close price per day (Source: Author)



Number of processed tweets per day (Source: Author)

METHODOLOGY (3/3)

INTRODUCING LAG

Continuous Time Lags: Investigation spans from 2 to 17 hours to determine optimal lag for prediction model performance in both scenarios.

Lagged Datasets: Training instances encompass observations from preceding days, corresponding to the lag duration. For instance, a lag of 2 includes data from the previous 2 days.

ORIGINAL

Hour	F1 (target)	F2	F3
H1	1	0.1	0.2
H2	0	0.3	0.4
H3	1	0.5	0.6
H4	0	0.7	0.8
H5	1	0.9	0.1

MODIFIED

Hour	f1(t-2)	f2(t-2)	f3(t-2)	f1(t-1)	f2(t-1)	f3(t-1)	f1(t)
H3	1	0.1	0.2	0	0.3	0.4	1
H4	0	0.3	0.4	1	0.5	0.6	0

Dataset with lagging feature of 2-hour lag (Source: Author)

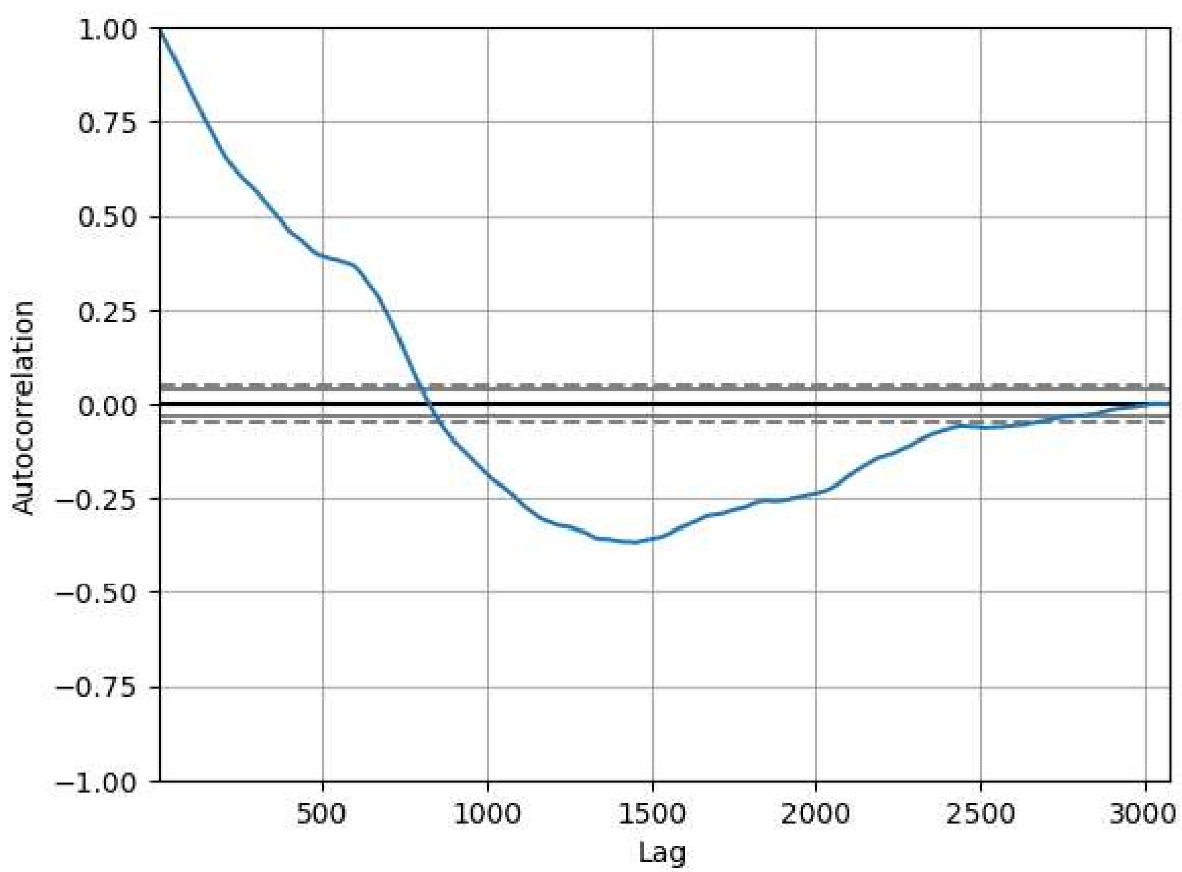
F1: target movement (up:1, down: 0) lag is 2 so all observations from last 2 days are variables for prediction. At H3, f1(t) is the actual target movement. We train all observations within 2 last days to predict the movement at H3.

RESULTS (1/4)

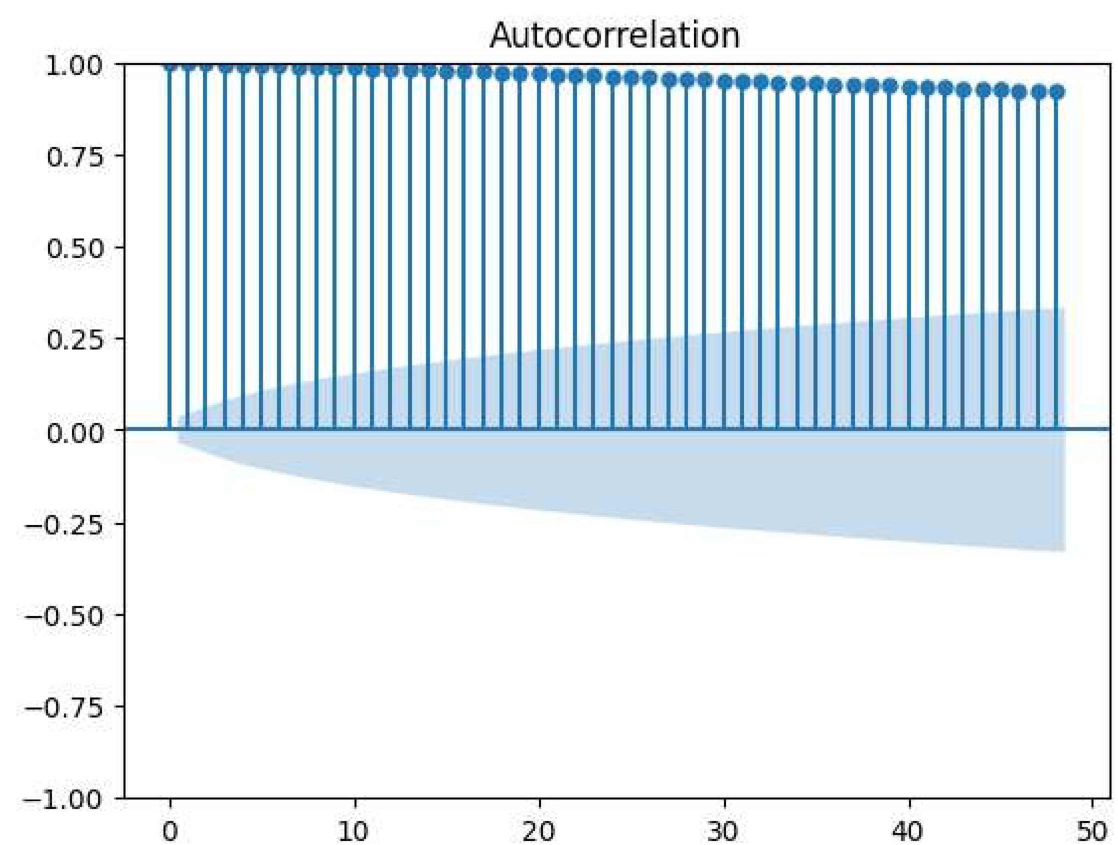
RESULTS OF STATISTICAL ANALYSIS

Non-Stationary Time Series: ADF test indicates Bitcoin price as non-stationary, with an ADF statistic of -1.4954 and p-value of 0.535780, failing to reject the null hypothesis at 1% critical value.

Strong Positive Correlation: Strong positive correlation observed for lags up to 500 hours, indicating a high degree of autocorrelation in the data.



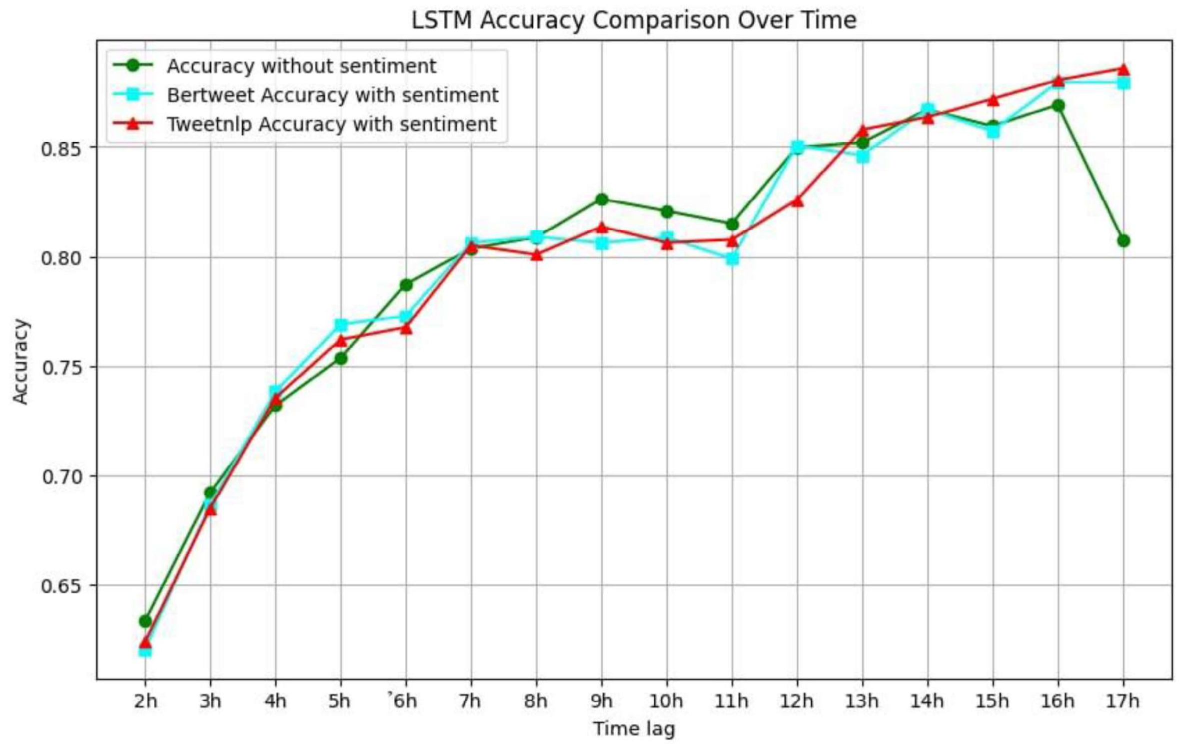
Autocorrelation in whole period (Source: Author)

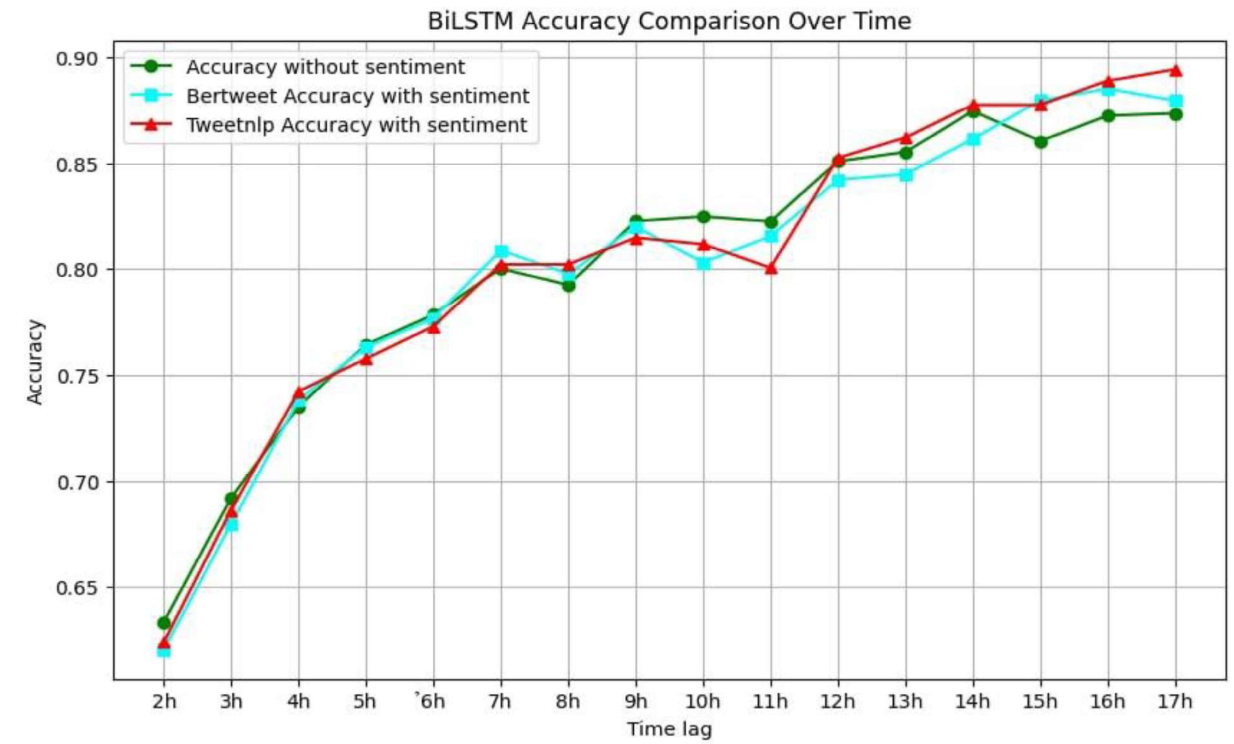


Autocorrelation in 50-hour interval (Source: Author)

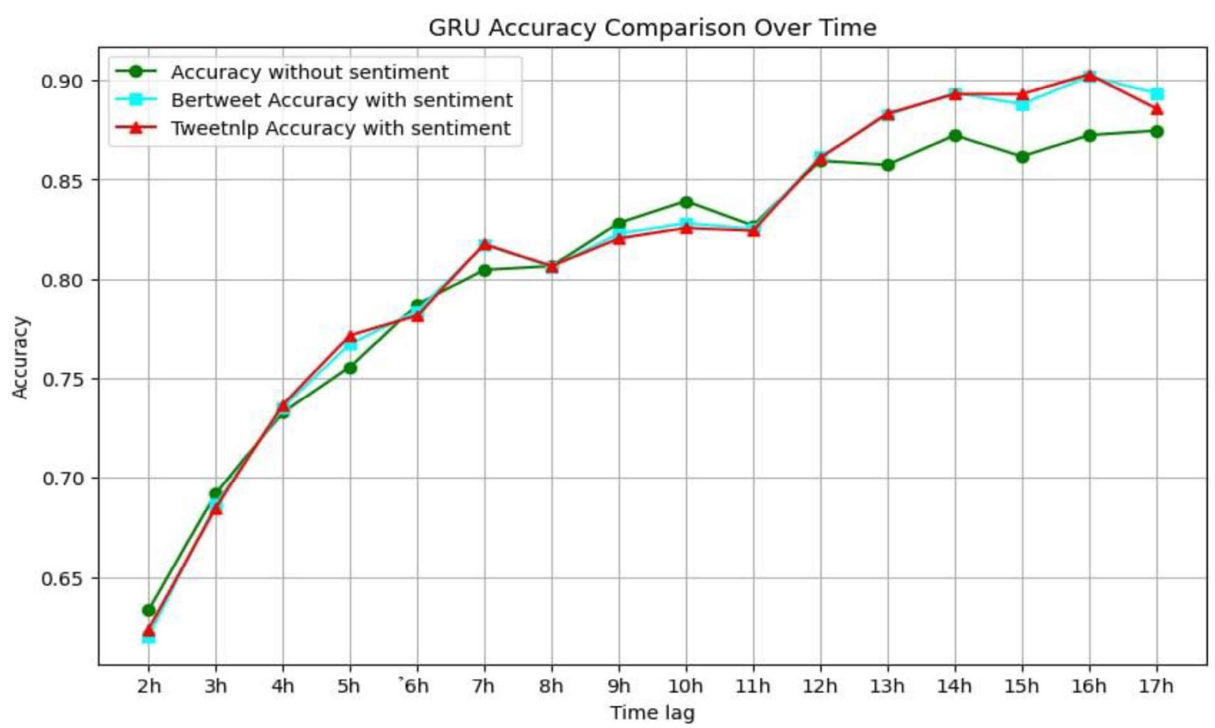
RESULTS (2/4)

RESULTS OF PRICE MOVEMENT PREDICTION





Accuracy results in each model (Source: Author)

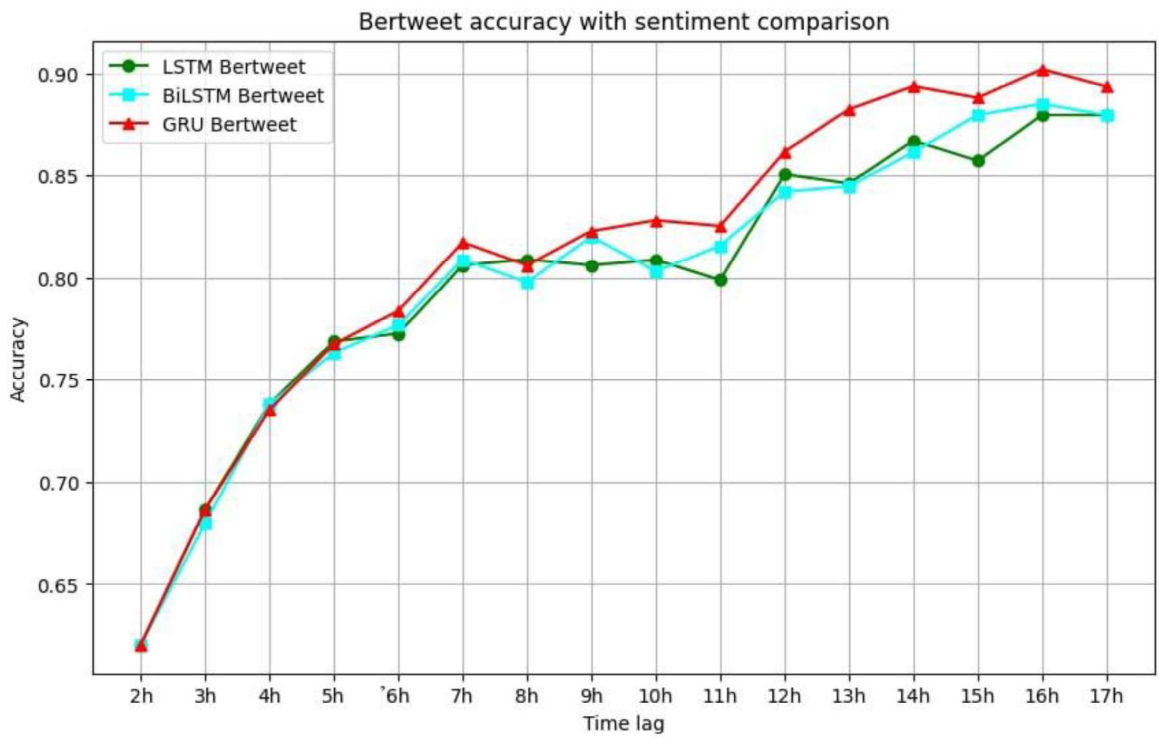


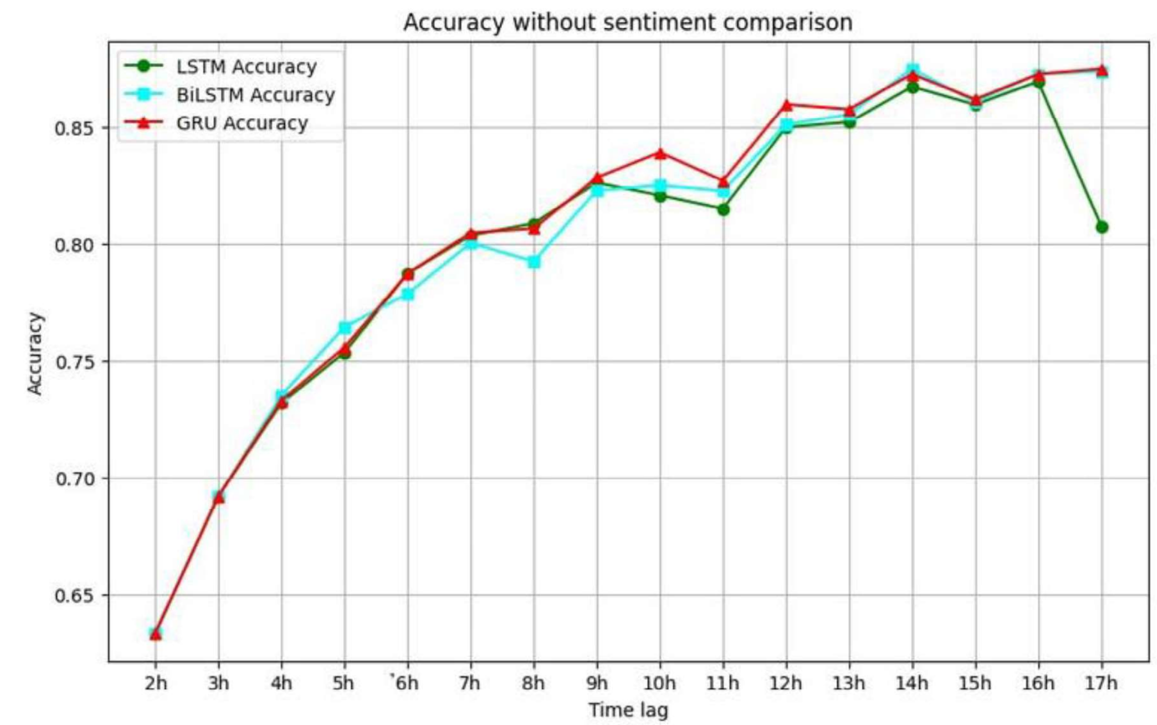
Balanced Data: Our dataset evenly represents increase and decrease classes, yielding similar accuracy and F1-scores, with focus on accuracy.

Model Performance: Initially, SA-lacking models outperform, but sentiment-inclusive models surpass them in accuracy for the last 5 hours.

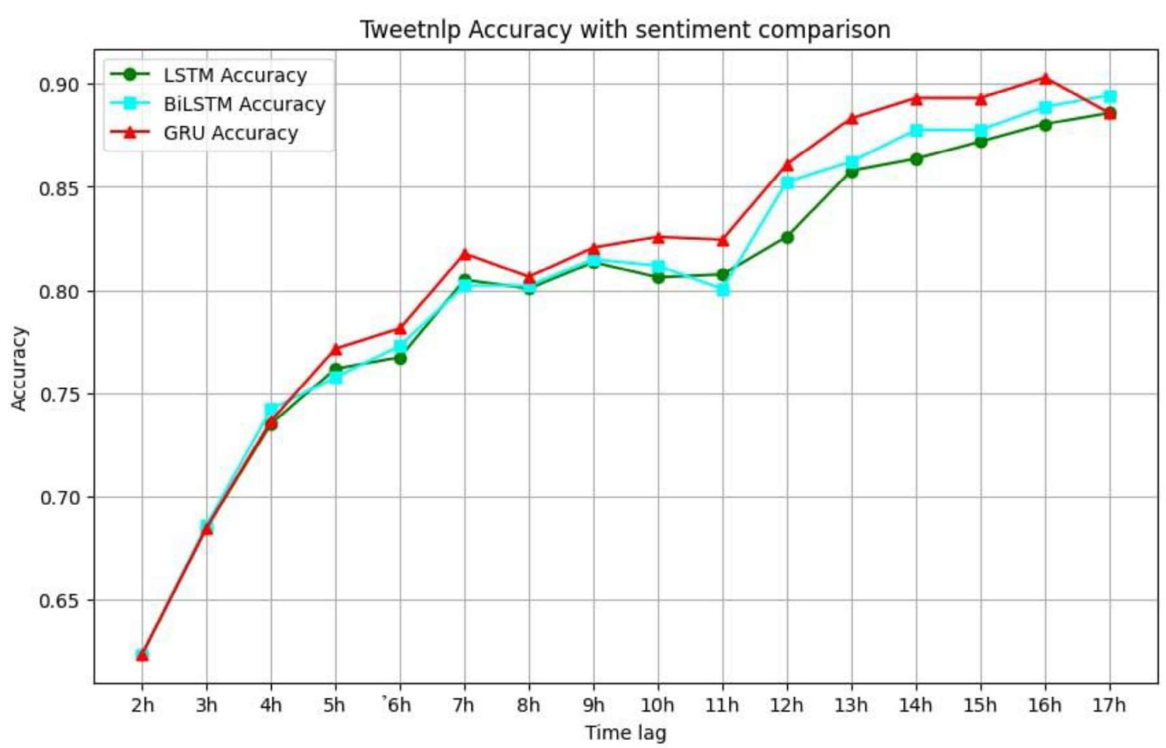
RESULTS (3/4)

RESULTS OF PRICE MOVEMENT PREDICTION





Accuracy results in each model (Source: Author)



GRU Superiority: GRU outperforms LSTM and BiLSTM in both scenarios.

SA Impact: Including SA notably enhances model performance for longer time periods (12 hours and beyond).

Top Accuracy: GRU with TweetNLP sentiment achieves the highest accuracy at 90.3%, closely followed by GRU with Bertweet sentiment at 90.2%.

RESULTS (4/4)

RESULTS OF PRICE MOVEMENT PREDICTION

Consistent Accuracy: All models maintain accuracy above 80% after 12 hours, with the inclusion of SA surpassing non-sentiment cases.

SA Impact: Highest accuracies observed with SA inclusion, notably with GRU leading, followed by BiLSTM with TweetNLP sentiment achieving 89.44% accuracy.

Small Accuracy Difference: The accuracy gap between sentiment and non-sentiment scenarios remains relatively small, below 2%.

GRU Superiority: Significant accuracy gap observed for GRU model from 13 to 16 hours, surpassing other models by 2% to 3%.

Consistent SA Performance: Accuracy difference between SA datasets remains around 1%.

LIMITATION

- **Limited Feature Scope:** Focus solely on Twitter sentiment attributes.
- **English Tweets Only:** Collecting exclusively English Tweets.
- **Missing Criteria:** Unable to incorporate user interaction due to API restriction.
- **Timeframe Restriction:** Analysis confined to specific time intervals.

CONCLUSION

Impact of Sentiment on Bitcoin Price: Sentiment expressed in Tweets notably influences Bitcoin price changes, particularly over 12 hours or longer.

Model Performance with Sentiment Analysis: Incorporating sentiment analysis (SA) led to the highest accuracy (90.3%) using a 2-layer GRU model with TweetNLP sentiment at a 16-hour lag.

Comparative Analysis of Prediction Models: GRU outperformed LSTM and BiLSTM models in both sentiment and non-sentiment scenarios, achieving a maximum accuracy of 87.47% (non-sentiment) and 90.3% (with SA) at specific time lags.

Optimal Time Lag Influence: Results suggest optimal lag times of 16 hours and 17 hours within the limited timeframe analyzed, but longer datasets may yield different optimal lag times.

REFERENCES

Ladislav, K. (2015). What Are the Main Drivers of the Bitcoin Price? Evidence from Wavelet Coherence Analysis.

Georgoula et al. (2015). Using Time-Series and Sentiment Analysis to detect the determinants of Bitcoin Prices.

Serafini et al. (2020, July). Sentiment-Driven Price Prediction of the Bitcoin based on Statistical and Deep Learning Approaches. International Joint Conference on Neural Networks (IJCNN), 19 - 24. doi:10.1109/IJCNN48605.2020.9206704

Ye et al. (2022). A Cryptocurrency Price Prediction Model Based on Twitter Sentiment Indicators. Communications in Computer and Information Science. doi:10.1007/978-981-19-0852-132

Low et al. (2024). Deep Learning and Sentiment Analysis-Based Cryptocurrency Price Prediction. Advances in Visual Informatics. doi:10.1007/978-981-99-7339-24

Critien et al. (2022). Bitcoin price change and trend prediction through twitter sentiment and data volume. Financial Innovation.

Murray et al. (2023). On Forecasting Cryptocurrency Prices: A Comparison of Machine Learning, Deep Learning, and Ensembles. Forecasting. 196-209. doi:10.3390/forecast5010010

Jaquart et al. (2022). Machine learning for cryptocurrency market prediction and trading. The Journal of Finance and Data Science, 331-352. doi:10.1016/j.jfds.2022.12.001

Jose et al. (2022). TweetNLP: Cutting-Edge Natural Language Processing for Social Media. the 2022 Conference on Empirical Methods in Natural Language Processing: System Demonstrations. doi:10.18653/v1/2022.emnlp-demos.5

Juan et al. (2021). pysentimiento: A Python Toolkit for Opinion Mining and Social NLP tasks. Computation and Language. doi:10.48550/arXiv.2106.09462

THANK YOU