Sentiment-Driven LSTM Analysis of Bitcoin Price: Uncovering Insights from Tweets and Macroeconomics Data

Po-Han, Lu Department of Cross College Elite Program, Bachelor
Jyun-Siyan, Chu Department of Biomedical Engineering, Master
Chia-Chun Hsu, Department of Biomedical Engineering, Bachelor

Advisor: Ya-Ning Chang, Assistant Professor
yaningchang@gs.ncku.edu.tw
Miin Wu School of Computing,
National Cheng Kung University

National Cheng Kung University

Abstract

This project utilizes deep learning, specifically LSTM and NLP for sentiment analysis to analyze Bitcoin price with Bitcoin tweets, and Macroeconomics indexes such as S&P 500 index and USD INDEX. The primary goal is to leverage sentiment analysis and LSTM techniques to uncover patterns in the cryptocurrency market, particularly about Bitcoin price movements, which can provide financial insights and effective decision-making. By incorporating multiple data sources and leveraging deep learning techniques, the outcome of the project will provide valuable insights into the interplay between social media sentiment, financial indicators, and Bitcoin price dynamics, while showcasing the efficacy of deep learning models for time series analysis in the cryptocurrency domain.

Keywords: Bitcoin; Sentiment analysis; LSTM; Macroeconomic factor

Motivation

This research project is driven by the compelling need to unravel the complex dynamics that influence Bitcoin price, one of the most prominent cryptocurrencies. Bitcoin's decentralized characteristic and volatile price fluctuations have sparked significant interest among traders, investors, and financial analysts. However, understanding the factors that drive its price movements remains a challenging task. The motivation behind this project stems from the recognition that sentiment expressed on Twitter, can have a profound impact on market behavior. Analyzing the collective sentiment surrounding Bitcoin on these platforms can provide valuable insights into investor sentiment, market trends, and potential price volatility. By leveraging sentiment analysis techniques and employing LSTM models, renowned for their ability to process sequential data, this research project aims to uncover hidden patterns, sentiment-driven indicators, and correlations between sentiment and Bitcoin price, ultimately contributing to a deeper understanding of Bitcoin price dynamics.

Additionally, incorporating macroeconomic factors into the analysis further enriches our understanding of Bitcoin price movements. Economic indicators, such as S&P 500 index and USD INDEX can exert significant influence on cryptocurrency markets. By integrating macroeconomic data into our research framework, we aim to capture the broader and deeper contextual factors that shape Bitcoin's price. This integration allows us to explore the interplay between sentiment, social media, macroeconomic indicators, and Bitcoin price, enabling a more comprehensive analysis of the complex dynamics at play. The ultimate goal of this project is to empower cryptocurrency stakeholders with actionable insights, helping them make informed decisions, manage risks, and navigate the evolving landscape of Bitcoin and its associated markets.

Literature Review

Bitcoin, crafted in 2008 by Satoshi Nakamoto, stands as a decentralized digital currency. In the past decade, Bitcoin has shattered expectations, surpassing $68,000 per coin. However, this astonishing surge intertwines with heightened fluctuations, amounting to 3.36 times the volatility of the S&P 500. Thus, understanding Bitcoin's trends to mitigate the risks posed by its unpredictable nature remains a formidable challenge. Study [1] predicts Bitcoin prices using historical prices, tweet sentiment, volume, user following, and verification status, achieving a 9.06 MAPE on the FinBERT model. Study [2] establishes a strong correlation between Google Trends search volume, tweet volume, and cryptocurrency prices, and reveals that tweet volume, rather than sentiment, is a better predictor of price direction. Study [3] finds that random forest regression outperforms LSTM in price prediction when incorporating a wide range of macroeconomic data. Lastly, study [4] introduces the Twitter Financial Sentiment Index (TFSI), which captures credit and financial market sentiment through NLP analysis of Twitter data, showing similarities to traditional economic metrics like the U. Michigan Consumer Sentiment Index and Excess Bond Premium. We will explain the papers we review in the following.

The paper by Haritha and Sahana [1] discusses the cryptocurrency ecosystem and its impact on Twitter. It introduces an algorithm that aims to predict the Bitcoin price by analyzing historical prices, the sentiment of tweets, and various metrics such as tweet volume, user following, and verification status. The algorithm employs Bidirectional Encoder Representations from

Transformers (BERT) based Neural Network Models to forecast tweet sentiment and uses a Gated Recurrent Unit (GRU) to predict Bitcoin prices.

The paper by Abraham and colleagues [2] presents a method for predicting changes in Bitcoin and Ethereum prices by utilizing Twitter data and Google Trends data. It highlights the significant price swings experienced by both cryptocurrencies in daily and long-term valuations. The increasing use of Twitter as a news source influences purchase decisions and provides insights into the currencies' popularity. Understanding the impact of tweets on price direction can offer advantages to cryptocurrency users and traders. Their result demonstrates a high correlation between search volume index, tweet volumes, and cryptocurrency prices, indicating their use as proxies for general interest. The outcome reveals that tweet volume, rather than sentiment, is a better predictor of price direction, as tweet sentiment tends to be positive regardless of price changes (as in Figure 1).

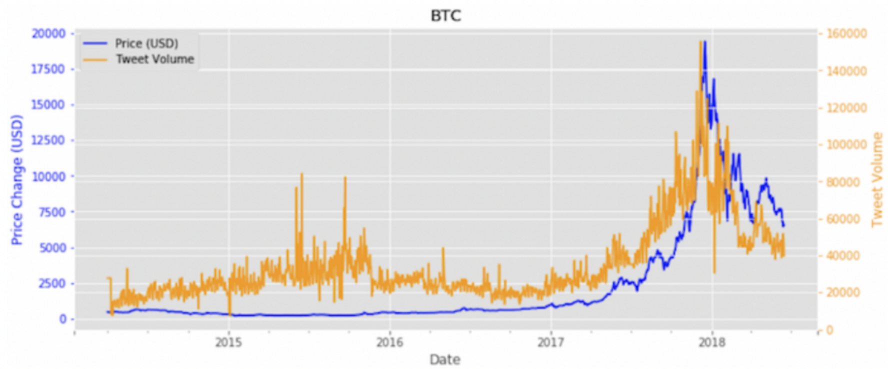

Figure 1. Bitcoin tweet volumes (yellow line) by day. Bitcoin price is shown by the blue line of [2]

The paper [3] aims to develop an algorithm that accurately predicts the price of Bitcoin for the next day. It utilizes random forest regression and LSTM techniques to achieve this goal. The research finds that three US stock market indexes (NASDAQ, DJI, and S&P 500), oil price, and ETH price influenced Bitcoin prices between 2015 and 2018 (as in Figure 2).

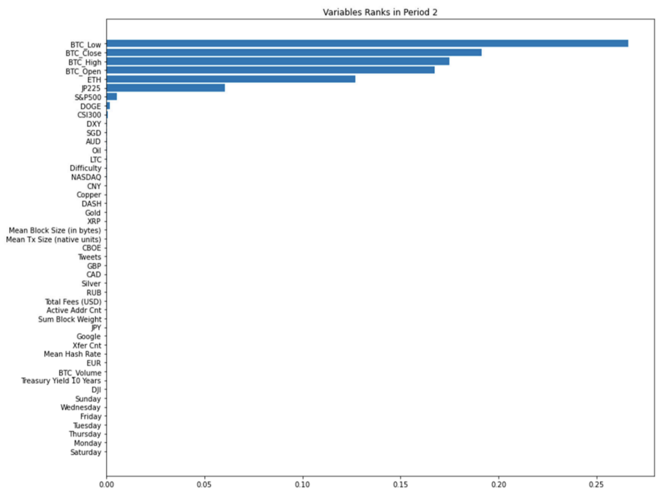

Figure 9. Explanatory variable importance ranks using random forest regression.

Figure 2. Explanatory variable importance ranks using random forest regression of [3]

In this paper [4], the authors built a new measure of credit and financial market sentiment using NLP on Twitter data, called the Twitter Financial Sentiment Index (TFSI). This index is a metric used to gauge the overall sentiment or mood of the financial markets by analyzing tweets related to various financial instruments, companies, or economic indicators. This index is derived from sentiment analysis techniques that examine the language, context, and emotions expressed in tweets. By analyzing the sentiment of tweets discussing specific stocks, commodities, currencies, or broader economic trends, the index aims to capture the general market sentiment at a given time.

To calculate the index, various sentiment analysis algorithms are applied to a large volume of tweets related to the financial markets. The sentiment scores are aggregated and normalized to generate the index value, which can range from negative to positive, with zero indicating neutral sentiment.

The TFSI presents a quantitatively and qualitatively similar picture to the U. Michigan Consumer Sentiment Index (as in Figure 3), which can offer insights into the overall mood of market participants, helping traders and investors to understand the prevailing sentiment and potential market trends.

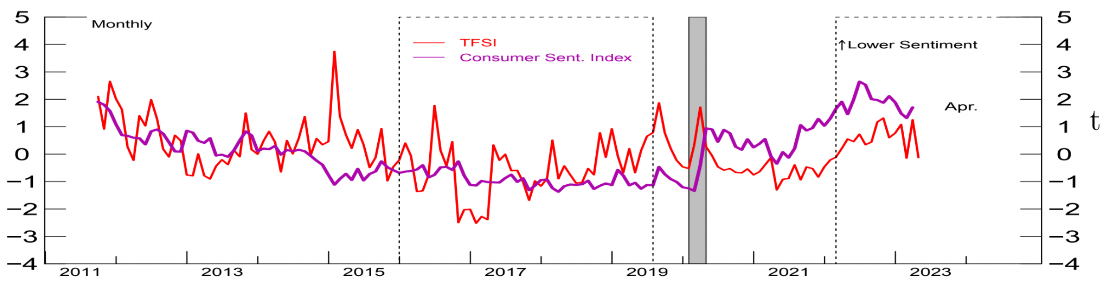

Figure 3, TFSI and U. Michigan Consumer Sentiment Index

Our project aims to pioneer an innovative approach: harmoniously merging sentiment analysis and macroeconomic data to construct a powerful "Sentiment-driven LSTM" model, specifically designed for predicting Bitcoin prices. Unlike existing studies [1-4] that commonly utilize these data sources in isolation, our integration unveils untapped potential and propels us toward a new era of forecasting accuracy. By combining the emotive intelligence of sentiment analysis with the economic underpinnings of macroeconomics data, our model establishes a holistic and robust prediction framework, empowering stakeholders with a comprehensive understanding of Bitcoin price dynamics.

Data Collection

We collect our Bitcoin tweets dataset from Kaggle: Bitcoin tweets - 16M tweets With Sentiment Tagged [5]. The dataset contains Bitcoin-related tweets in multiple languages, the tools the dataset used to collect the tweets are "twint", "tweepy". The dataset collects the tweets posted between 2014-09-18 to 2019-11-23, contained 18,452,917 tweets, and averaged 9,748 tweets a day. The sentiment tag is labeled by humans with three kinds of labels: *positive*, *negative*, and *neutral*, as the analyzed result between the Bitcoin price and the tweets. We average the sentiment score per day, positive as 1, negative as -1, and neutral as 0, and set it as a feature of the dataset. On the other hand, we also collect macroeconomic data from Yahoo! Finance. We use the Yahoo! Finance API, yfinance, to withdraw the historical data of the USD trend, S&P 5000 index, and Bitcoin trend. The Bitcoin trend is also used as the target label of our model. Every macroeconomic data is arranged in chronological order by day. The data for each day contains the open price, the close price, the highest price of the day, the lowest price of the day, the adjusted close price, and the trading volume of the day. So, in total, we have 30 features (6x5) for daily economic data, and the total data we get from yfinace between 2014-09-18 to 2019-11-23 are 1303x30 as there is a total of 1303 days that have a history data according to Yahoo! Finance.

The data is windowed by their orders, weekly, monthly, and trimonthly. The dataset is stored in three dimensions. The first dimension of the dataset represents how many days the dataset has. The second dimension represents how many days to look back, a week is 7, etc. The third dimension represents how many features we have, in our case 19. For example, if we want to predict the Bitcoin price on 2018-10-21, we look up the data (including macroeconomic data and sentimental data) one week before 2018-10-21 (i.e. 7 days to 2018-10-14). So the data that we put into our model would have 1296 in the first dimension because the total 1303 days minutes seven days (because we cannot predict the first day with no data before it), and 7 in the second dimension as we looked back seven days, and 19 in the third dimension as the 18 macroeconomic data features plus one sentiment feature. If we looked back on monthly orders, we would look back 30 days. So the first dimension of the data would be 1303 minus 30, which is 1273. The second dimension would be 30 days, and the third dimension is 19 features. The shape of input data would be (1273, 30, 19).

Methodology

The workflow of our project is illustrated in Figure 4. First, we collect Bitcoins-related tweets data via Twitter API and then feed it to the VADER (Valence-Aware Dictionary and Threaction Reasoner) sentiment analyzer [7] to get the sentiment score. And at the same time, using Yahoo! Finance API to collect historical macroeconomic data and go through the window process. We combine sentiment data and economic data into the LSTM model to make the final prediction of future Bitcoin price and also give out the feature weighting insight to see the importance of the features.

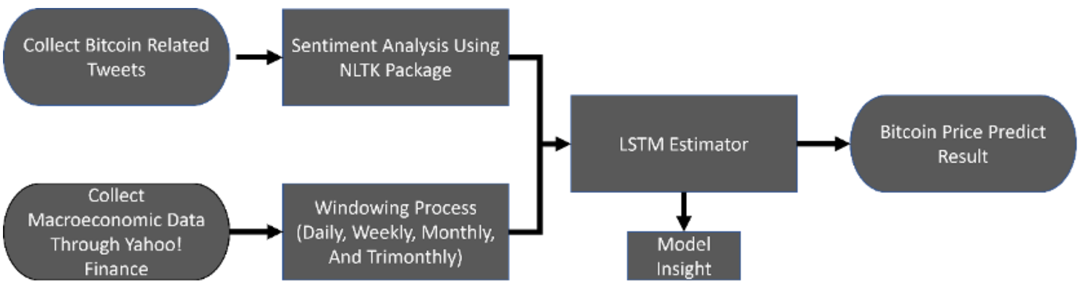

Figure 4. The work block diagram of our project

```
>>> from nltk.corpus import stopwords
>>> stopwords.words('english')
['i', 'me', 'my', 'myself', 'we', 'our', 'ours', 'ourselves', 'you', 'your', 'yours',
'yourself', 'yourselves', 'he', 'him', 'his', 'himself', 'she', 'her', 'hers',
'herself', 'it', 'its', 'itself', 'they', 'them', 'their', 'theirs', 'themselves',
'what', 'which', 'who', 'whom', 'this', 'that', 'these', 'those', 'am', 'is', 'are',
'was', 'were', 'be', 'been', 'being', 'have', 'has', 'had', 'having', 'do', 'does',
'did', 'doing', 'a', 'an', 'the', 'and', 'but', 'if', 'or', 'because', 'as', 'until',
'while', 'of', 'at', 'by', 'for', 'with', 'about', 'against', 'between', 'into',
'through', 'during', 'before', 'after', 'above', 'below', 'to', 'from', 'up', 'down',
'in', 'out', 'on', 'off', 'over', 'under', 'again', 'further', 'then', 'once', 'here',
'there', 'when', 'where', 'why', 'how', 'all', 'any', 'both', 'each', 'few', 'more',
'most', 'other', 'some', 'such', 'no', 'nor', 'not', 'only', 'own', 'same', 'so',
'than', 'too', 'very', 's', 't', 'can', 'will', 'just', 'don', 'should', 'now']
```

Figure 5. Stopword corpus from NLTK package

I. NLTK:

We use the Natural Language Toolkit (NLTK) package for the sentiment analysis task, which is an NLP toolbox based on Python. It was first created as part of a computational linguistics course in the Department of Computer and Information Science at the University of Pennsylvania [6]. We use some of the tools in NLTK, which we will introduce as follow:

A. *nltk.stopwords:*

Stopwords in the NLP field are high-frequency words in a sentence that have less correlation with the semantic meaning of the sentence itself. Taking the stopwords away from the context diminishes the presence of redundant words, thus enabling the distinctiveness of the context to emerge and be discerned from other texts. The NLTK package has a corpus of stopwords shown in Figure 5.

B. *nltk.steming:*

In a language system like English, one word could have multiple tenses, such as singular, plural, active, and passive, but its semantic meaning is the same. So we use the stemming process to fuse those words as one word. The process significantly reduces the number of words, which reduces the complexity of the data and speeds up the following process. For example, "studying" and "study" in the sentences "I am studying" and "we study a book in the class" are not so different Thus, both words would be seen as "study" after the stemming process. We import nltk.stem for the steming process.

C. *nltk.tokenize:*

A corpus that contains a large amount of data is not directly suitable to put into the NLP model. Tokenization helps us divide the corpus into word elements which can be processed in the following tasks. For example, "Arthur didn't feel very good." can be tokenized as "Arthur", "did", "n't", "feel", "very", "good", ".". In the NLTK package we import nltk.word_tokenize for the tokenization task.

D. *nltk.vader.sentiment:*

The NLTK package has a sentiment analysis tool named VADER. VADER is a rule-based system with a well-constructed and validated sentiment lexicon [7]. The system uses over 7500 vocabularies to identify the sentiment score. It generates a score that summarizes the emotional strength of the input text,by summing the scores for each feature in the lexicon, adjusted according to the rule, and then normalized to be between -1 (the most extreme negative) and +1 (the most extreme positive).

Although VADER can sometimes make mistakes in processing texts such as ironic corpus, the system can predict results word by word and its processing speed is generally fast because it is a rule-based structure.

II. The LSTM model:

We use the LSTM model as our estimator for the prediction task of the Bitcoin price. LSTM structure is capable of learning long-term dependencies, which is suitable for the windowing data form we have in our dataset. Following is the basic model structure that we choose for the task.

Vanilla LSTM:

The original structure was brought up by Sepp Hochreiter and Jürgen Schmidhuber in 1997[10]. The LSTM is a special RNN that contains three basic elements: forget gate, input gate, and output gate. The gates are shown in Figure 6. Forget gate is used to forget the unimportant and remember the important. We use the input gate to update the cell status. The output gate determines the value of the next hidden state. This state contains information on previous inputs.

To conclude, the forget gate determines which relevant information from the prior steps is needed. The input gate decides what relevant information can be added from the current step, and the output gates finalize the next hidden state [9].

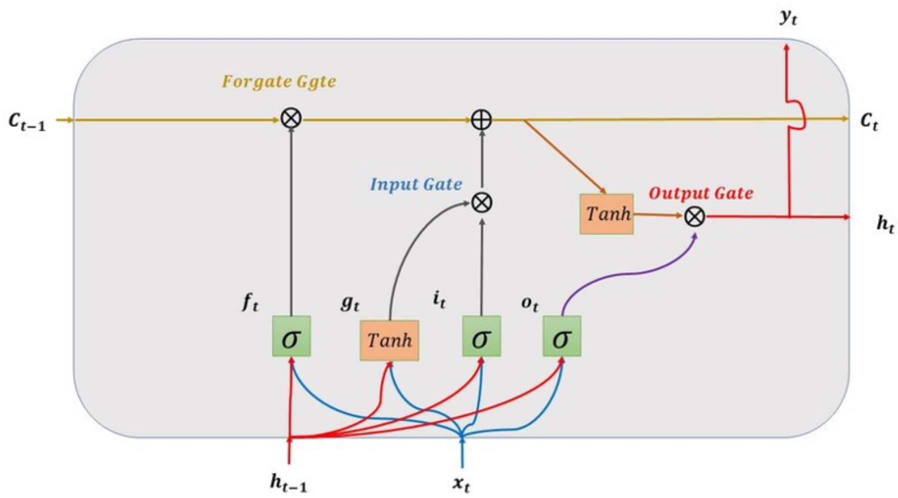

Figure 6. One Vanilla LSTM cell [8][9]

Results

In this study, we designed three experiments for observing the impact of different time windows including 1-week, 1-month, and 3-months. For each experiment, the LSTM model and CNN-LSTM model were used to train on different combinations of features. From our pilot explorations, we opted to use the macroeconomic features of BTC as the baseline feature set, and subsequently incorporated USD features, sentiment features, and SP500 features as additional features, as shown in Tables 1, 2, and 3. We then evaluated the regression results against the ground truth using Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) on both the training and validation datasets.

For the 1-week time windows experiments, the BTC + Sentiment feature set acquired the best results of the RMSE and MAE in the validation dataset, as indicated in Table 1**.** The results obtained by the LSTM model were 0.12 and 0.09, and the results obtained by the CNN-LSTM model were 0.16 and 0.12. Also, the CNN-LSTM model got the same results in the BTC + USD + Sentiment feature set. Similarly, for the 1-month time window experiments, the LSTM model yielded superior performance in the BTC + Sentiment feature set, as shown in Table 2, with the RMSE of 0.12 and the MAE of 0.09. The CNN-LSTM model obtained the best results in the BTC + USD + Sentiment feature set with the RMSE of 0.14 and the MAE of 0.11. However, when the time window was extended to 3 months, the baseline BTC feature set outperformed the other feature set in the validation dataset, as highlighted in Table 3. It achieved the RMSE of 0.11 and the MAE of 0.08 with both the LSTM and CNN-LSTM model.

Table 1. Regression evaluation for 1-week time window

[Time window: 1 week] BTC Price Prediction – Regression Evaluation								
Features (# of features)	LSTM				CNN-LSTM			
	Train		Validation		Train		Validation	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
BTC (6)	0.09	0.07	0.21	0.18	0.07	0.04	0.22	0.18
BTC + USD (12)	0.05	0.03	0.21	0.13	0.09	0.05	0.19	0.15
BTC + Sentiment (7)	0.06	0.04	**0.12**	**0.09**	0.06	0.05	**0.16**	**0.12**
BTC + SP500 (12)	0.05	0.03	0.21	0.18	0.07	0.04	0.20	0.15
BTC + USD + Sentiment (13)	0.08	0.04	0.17	0.14	0.07	0.05	**0.16**	**0.12**
BTC + USD + SP500 (18)	0.05	0.03	0.23	0.20	0.05	0.04	0.17	0.14
BTC + Sentiment + SP500 (13)	0.04	0.03	0.20	0.16	0.04	0.02	0.19	0.16
BTC + USD + Sentiment + SP500 (19)	0.04	0.03	0.19	0.17	0.03	0.02	0.19	0.15

Table 2. Regression evaluation for 1-month time window

[Time window: 1 month] BTC Price Prediction – Regression Evaluation								
Features (# of features)	LSTM				CNN-LSTM			
	Train		Validation		Train		Validation	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
BTC (6)	0.11	0.05	0.26	0.21	0.07	0.06	0.18	0.14
BTC + USD (12)	0.08	0.05	0.15	0.11	0.06	0.04	0.16	0.12
BTC + Sentiment (7)	0.08	0.04	**0.12**	**0.09**	0.04	0.03	0.15	0.11
BTC + SP500 (12)	0.06	0.04	0.35	0.31	0.08	0.06	0.20	0.17
BTC + USD + Sentiment (13)	0.07	0.04	0.16	0.11	0.08	0.07	**0.14**	**0.11**
BTC + USD + SP500 (18)	0.05	0.03	0.43	0.39	0.07	0.04	0.21	0.17
BTC + Sentiment + SP500 (13)	0.07	0.04	0.32	0.29	0.08	0.06	0.18	0.15
BTC + USD + Sentiment + SP500 (19)	0.07	0.06	0.34	0.30	0.09	0.05	0.19	0.15

Table 3. Regression evaluation for 3-months time window

[Time window: 3 months] BTC Price Prediction – Regression Evaluation								
Features (# of features)	LSTM				CNN-LSTM			
	Train		Validation		Train		Validation	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
BTC (6)	0.07	0.05	**0.11**	**0.08**	0.08	0.04	**0.11**	**0.08**
BTC + USD (12)	0.07	0.04	0.14	0.11	0.06	0.03	0.16	0.11
BTC + Sentiment (7)	0.06	0.03	0.15	0.10	0.05	0.04	0.13	0.09
BTC + SP500 (12)	0.08	0.05	0.18	0.13	0.09	0.05	0.13	0.09
BTC + USD + Sentiment (13)	0.06	0.03	0.15	0.11	0.10	0.06	0.12	0.09
BTC + USD + SP500 (18)	0.09	0.07	0.17	0.13	0.07	0.04	0.12	0.09
BTC + Sentiment + SP500 (13)	0.11	0.09	0.18	0.12	0.08	0.05	0.12	0.09
BTC + USD + Sentiment + SP500 (19)	0.04	0.02	0.16	0.12	0.08	0.06	0.15	0.12

Discussion

I. Different feature combinations

In the experiments with the time window of 1-week and 1-month, there are three key observations in Tables 1 and 2. Firstly, the sentiment feature effectively enhances the predictive capability of the both LSTM and CNN-LSTM models. This phenomenon is clearly illustrated in Figure 7, where the predictive results of the BTC + Sentiment feature set closely align with the actual price trends. The blue line represents the ground truth, while the yellow and green lines represent the predictions from the training and validation datasets, respectively. Secondly, in the LSTM models, the SP500 feature does not improve the predictive performance. Moreover, it generates significant negative effects during the 1-month time window, as shown in Table 2. The average RMSE and MAE of the feature set containing SP500 are 0.36 and 0.32. Thirdly, the CNN-LSTM model is effective in mitigating the adverse impact of the SP500 feature within the LSTM model. It is noteworthy that the CNN-LSTM model demonstrates the ability to extract meaningful information from complex feature sets, leading to better prediction outcomes.

In the 3-month time window experiments, there is one interesting observation in **T**able 3. Using only the baseline BTC feature set yields the best prediction performance. We attribute this phenomenon to the fact that when considering a longer time window of 3 months, the macroeconomic feature directly relates to the BTC price provided the most effective information. These observed phenomena indicate that the sentiment features can significantly enhance the predictive capability of models for the medium-term (1-week and 1-month) time window. However, for the long-term (3-month) time window, the sentiment features do not further improve the predictive performance. This implies that sentiment features, capturing immediate reactions and sentiments related to Bitcoin events, play a crucial role in short-term price prediction by reflecting market sentiment and impacting short-term price movements. The real-time nature of tweets and the influence of social media have substantial implications in this regard. Additionally, as the number of feature combinations increases, the CNN-LSTM model outperforms the LSTM model in terms of prediction performance, indicating that the CNN-LSTM model is better at extracting valuable information from complex feature sets.

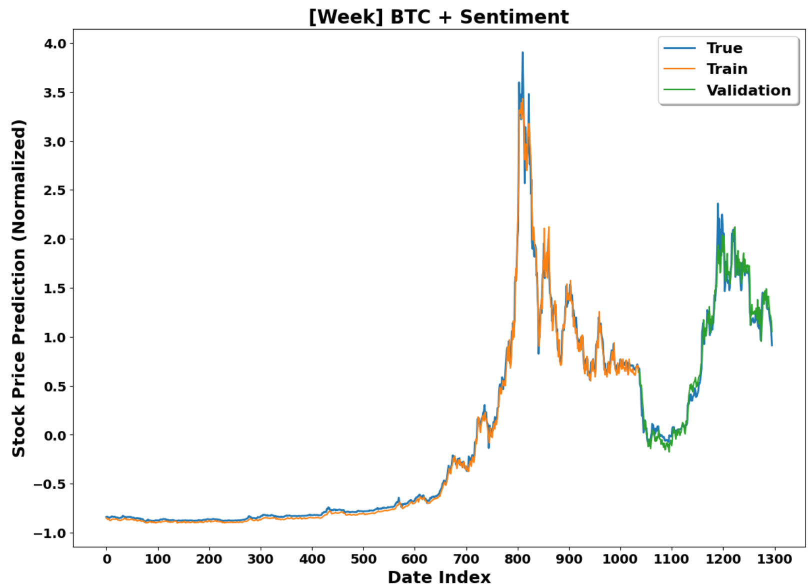

Figure 7. BTC price prediction with BTC + Sentiment feature set in a 1-week time window. The RMSE and MAE of the feature set are 0.12 and 0.09.

II. Social Impact

The observed experimental results have some significant social impacts. Firstly, given that most investors struggle to quantify the sentiment of social media, they can combine their own experiences with the findings of this study to design a comprehensive investment strategy for reducing risks and increasing returns. Secondly, based on the observed phenomena of sentiment features enhancing Bitcoin price prediction in this study, further research can be extended to systematically investigate the impact of social media sentiment on cryptocurrencies. Additionally, this observation also serves as a reference for researchers studying other financial markets, allowing them to investigate the influence of social media sentiment on different markets.

Thirdly, this research contributes to the advancement of the finance research field by demonstrating the efficacy of sentiment features in improving prediction results within medium-term (1-week and 1-month) time windows. This implies the influence of social media on market behavior and highlights the potential of specific feature combinations and deep learning models in the application of cryptocurrencies.

Conclusion and Future Work

In this study, we investigated the potential of using social media for predicting cryptocurrency prices by combining deep learning techniques with NLP sentiment analysis. The results showed that incorporating sentiment features improves the prediction performance of the deep learning models. Furthermore, the results of this study enable investors in the cryptocurrency field to enhance their investment strategies by integrating their own experiences, allowing them to better understand the influence of social media on market fluctuations.

In the future, combining raw textual information with macroeconomic data to design a multimodal model for extracting comprehensive features from diverse input data is a potential avenue of research, given that this study has shown the effectiveness of incorporating sentiment features from social media to enhance prediction performance. Additionally, it is necessary to gather new data to ensure that the size of the dataset meets the requirements of the models. Moreover, expanding the prediction timeframe to various scales such as three days, three weeks, and two months can help provide a comprehensive understanding of the model's applicability.

References

[1] Haritha, G. B., & Sahana, N. B. (2023, February). Cryptocurrency Price Prediction using Twitter Sentiment Analysis. In *CS & IT conference proceedings* (Vol. 13, No. 3). CS & IT Conference Proceedings.

[2] Abraham, J., Higdon, D., Nelson, J., & Ibarra, J. (2018). Cryptocurrency price prediction using tweet volumes and sentiment analysis. SMU Data Science Review, 1(3), 1.

[3] Chen, J. (2023). Analysis of Bitcoin Price Prediction Using Machine Learning. Journal of Risk and Financial Management, 16(1), 51.

[4] Adams, T., Ajello, A., Silva, D., & Vazquez-Grande, F. (2023). More than words: Twitter chatter and financial market sentiment. arXiv preprint arXiv:2305.16164.

[5] Gaurav, D. Bitcoin tweets - 16M tweets With Sentiment Tagged. Kaggle. https://www.kaggle.com/datasets/gauravduttakiit/bitcoin-tweets-16m-tweets-with-sentiment-tagged.
.
[6] Steven, B., Ewan, K., & Edward, L. Natural Language Processing with Python–Analyzing Text with the Natural Language Toolkit. nltk.org. https://www.nltk.org/book/.
[7] Hutto, C., & Gilbert, E. (2014, May). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In Proceedings of the international AAAI conference on web and social media (Vol. 8, No. 1, pp. 216-225).
[8] Safarzadeh, V. M., & Jafarzadeh, P. (2020, January). Offline Persian handwriting recognition with CNN and RNN-CTC. In 2020 25th international computer conference, computer society of Iran (CSICC) (pp. 1-10). IEEE.
[9] Gaurav, S. Introduction to LSTM Units in RNN Pluralsight. https://www.pluralsight.com/guides/introduction-to-lstm-units-in-rnn.
[10]Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.