Does Social Media Sentiment Predict Bitcoin Trading Volume?

Short Paper

Jayit Saha
Vellore Institute of Technology
Chennai, India
jayitsaha@gmail.com

Smit Patel
Indian Institute of Technology
Indore, India
smitpatel2360@gmail.com

Frank Xing
National University of Singapore
Singapore
xing@nus.edu.sg

Erik Cambria
Nanyang Technological University
Singapore
cambria@ntu.edu.sg

Abstract

Social media sentiment is proven to be an important feature in financial forecasting. While the effect of sentiment is complex and time-varying for traditional financial assets, its role in cryptocurrency markets is unclear. This research explores the predictive power of public sentiment on Bitcoin trading volume. We develop a novel sentiment analysis pipeline for processing Bitcoin-related tweets and achieve state-of-the-art accuracy on a benchmark dataset. Our pipeline also leverages information gain theory to incorporate the impact of textual and non-textual features. We use such features to discern a non-linear relationship between public sentiment and Bitcoin trading volume and discover the optimal predictive horizon for Bitcoin. This research provides a useful module and a foundation for future studies and understanding of Bitcoin market dynamics, and its interaction with social media buzzing.

Keywords: Sentiment Analysis, Bitcoin Volume, SenticNet, FinBERT, Information Gain

Introduction

A cryptocurrency is a form of virtual or digital currency that is protected by cryptography, making it technically impossible to forge. Most cryptocurrencies are decentralized networks based on blockchain technology: a distributed ledger enforced by a disparate network of computers (Hassan et al. 2020). Bitcoin is the most popular cryptocurrency today. It is powered by blockchain that allows peer-to-peer transactions enforced by highly complex cryptographic algorithms. Several studies have engaged in predicting trading volumes in stock markets, to visualize their impact on the price movements of their respective stocks. Sabri (2008) concluded that the trading volume of stocks at an instant, is an important indicator for stock price prediction.

Since currency volume (Bouri et al. 2019) is not much explored yet an equally important indicator for cryptocurrencies, we extrapolate the above inference and focus on the prediction of Bitcoin trading volume, which subsequently shall become important for estimating other parameters and aspects of that currency. We choose Bitcoin as behavior of several crypto-currencies are mainly driven by Bitcoin (S Kumar et. al 2019). The volume–price dependency might vary depending on the market investigated and the assets considered. Based on (Rogalski, 1978), we assume that trade volumes are strong positive indicators of price change and conclude that our research ultimately aids in crypto price estimation. Moreover, since cryptocurrency is not controlled by any central power, it makes opinion of the public and a wide range of market participants play an even more important role.

It has also been observed that the general sentiment about a particular cryptocurrency drives it towards a net gain or loss, and a high correlation can be extracted from the same (Huynh 2021). Hence, it becomes imperative to tap the potential of the social media sentiment for predictive analysis. Our study follows the call to promote more predictive analytics in the IS community (Shmueli and Koppius 2011) and aims to establish a connection between two important issues, i.e., the role of sentiment analysis in the realm of cryptocurrency (Gao 2021), and asset trading volume as a feature in predicting many variables, including price associated with the same. Although the impact of social media sentiment on conventional stock markets has been studied (Ho et al. 2017), it remains unclear whether the findings can be directly migrated to crypto markets. We bridge these two aspects by determining the impact of social media sentiment on Bitcoin volume, using statistics and natural language processing (Xing et al. 2018). To the best of our knowledge, we are the first to explore this relationship and develop a viable solution for describing and determining such a correlation. This paper primarily aims to answer these research questions:

RQ1: Does Public Sentiment affect the trading volume of a cryptocurrency, particularly Bitcoin?

RQ2: Is there a time span within which the impact of public emotion is reflected on Bitcoin volume?

The findings of these research statements can have a meaningful impact in the financial industry where cryptocurrency has a sizeable market capitalization (over a trillion dollars) and is traded as an asset. Our research provides additional information to pipelines for effectively predicting cryptocurrency movements, and thereby influence investment decisions in the segment. Our primary contributions to the research focus are three-fold. Frist, we propose a novel pipeline for attributing the social media sentiment to the volume of Bitcoin. Second, we implement transfer learning on FinBERT for sentiment classification and achieve accuracy and F1 scores that outperform existing approaches on the Sentiment140 dataset. Finally, we formulate a statistical approach to incorporate both the textual and non-textual features of tweets which contribute to a holistic sentiment feature engineering.

Related Work

There has been a lot of research going on around cryptocurrency, particularly Bitcoin. Typically, many studies have been conducted around social media sentiment as cryptocurrency, at the end of the day, responds to the demand-supply law. (S Kumar et. al 2019) and (Bouri et. al. 2020) deduce that several cryptocurrencies imitate the behaviour of Bitcoin. Valencia et al. (2019) conducted a comparative analysis on different machine learning models for predicting the price of Bitcoin using twitter sentiment analysis. This research shows that neural networks achieve the highest accuracy in predicting the future prices. Abraham et al. (2018) incorporated total tweet volume and Google trends data along with tweet sentiment for their research. The former applied SentiStrength tool and determined the impact on Bitcoin prices whereas the latter study used VADER compound score to predict future prices. Mohapatra et al. (2018) created an adaptive learning model which performs predictive analysis of new sentiments and prices, adapting the weights as a coping mechanism. (Karasu 2018) feeds time-series data to ML algorithms for Bitcoin price estimation. Other studies also discussed the high correlation between Bitcoin prices and twitter sentiments. It also considers the historical Bitcoin price while training various models.

Researchers have also investigated into other aspects of cryptocurrency transactions such as volume traded at various exchanges and its value volatility. Balcilar et al. (2017) conducted causality-in-quantile tests to analyze the causal relationship between trading volume and Bitcoins returns and volatility over their conditional distributions. This study reveals that Bitcoin volume can predict market returns except for bull and bear regimes. Szetala et al. (2021) conducted a study to verify the existence of short and long-term relationships between the strength of a trend and the volume in bullish and bearish cryptocurrency markets. Nasir et al. (2019) primarily employed various established empirical methodologies to forecast the Bitcoin volume predictability and returns using Google search indexing and values. The research suggests that the Google search frequency equals a positive return and a surge in Bitcoin trading volume. These research works establish a link between Bitcoin volume and Bitcoin price movement and thereby, serve in abetting our research hypothesis. Various pipelines exist that work in an ensemble to have a better performance in predicting cryptocurrency market volatility (Walther et al. 2019). With this volumetric addition to the pipeline, we will have more information to build a predictive model which better than the current state of the art by (Ma et al. 2020).

Methodology

This section elaborates on the pipeline shown in Figure 1., which is used in our approach.
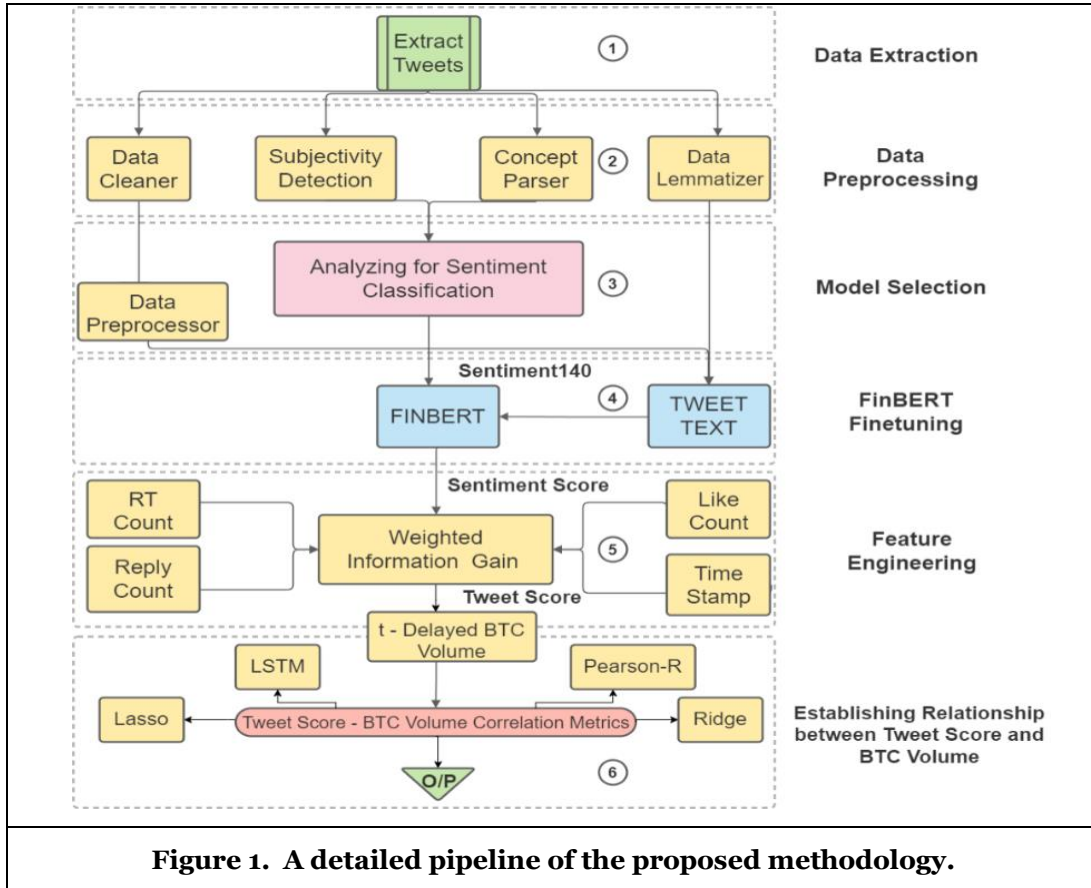
Figure 1. A detailed pipeline of the proposed methodology.

Data Extraction and Preprocessing

Approximately 1400 tweets were extracted every hour[1] after filtering retweets and replies to eliminate redundancy. The whole dataset, therefore, consists of around 4 million tweets spanning over 120 days ranging from 2021-05-28 to 2021-09-25, maximizing data and maintaining a proper hourly intake ratio. The data pull is based on specific keywords like @Bitcoin, #bitcoin, #BTC, #Bitcoin, #bitcoins, #bitcoinnews, #bitcointrading, #bitcoininvestment, #bitcointrader, making the extraction process specific to Bitcoin. The next major part of the analysis involves preprocessing the datasets for finer and more focused analysis. Before pushing into the pipeline, an intensive filter against crypto based spam keywords like "giveaway" or "top deals" was applied, in a parallel ensemble with the SenticNet Subjectivity detection to mitigate the presence of irrelevant and junk data. The data contains unlabeled tweets, for example "I think the next big news for #BTC will be @user announcing he has come up with a solution for #Bitcoin power usage." The data processing pipeline firstly eliminates any retweet duplication via comparing against a dynamic weighted metric threshold. Next, every tweet was passed into a batch preprocessing algorithm where all emoticons were removed as we empirically concluded that most emoticons in our dataset conveyed meaningless information which added bias to the model. Furthermore, every instance of a URL, hashtags, and user tags was replaced with a constant keyword and consecutive repetitions of words were eliminated. Lemmatization was applied to the tweets. Afterward, they were passed through the SenticNet subjectivity detection module to filter objectiveness.

---

[1] We use Twython, a wrapper for Twitter API, to extract tweets that contain Bitcoin-related hashtags.

SenticNet Implementation

SenticNet (Cambria et al. 2022) is a commonsense knowledgebase for sentiment analysis that makes the conceptual information conveyed by natural language more easily accessible to machines. This is accomplished by replacing the bag-of-words model with a new model that sees the text as a bag of concepts. By doing so, SenticNet goes beyond counting word co-occurrence frequencies and leverages an ensemble of symbolic and sub-symbolic AI techniques by integrating top-down and bottom-up learning to gain a deeper understanding of natural language. Moreover, SenticNet is rooted in a theory of emotion namely Hourglass of Emotions (Susanto et al. 2020), which is correlated to arousal-valence that is often used as a measure for web word-of-mouth We have used the SenticNet model for primarily two subtasks in our pipeline, namely (i) Concept parsing and (ii) subjectivity detection.

The SenticNet Parser is a tool that extracts the commonsense concepts from a sentence without spending excessive time on phase structure analysis. The algorithm first breaks text into clauses and then deconstructs clauses into small bags of concepts (SBoC) using a 2-step normalization process.

The SenticNet subjectivity detection is a tool to filter out factual or objective, information from free text, i.e., text which has no sentiment associated with it. The tool extends Extreme Learning Machines to a model consisting of Bayesian networks and fuzzy RNNs to detect subjectivity in a sentence. Bayesian networks are used to capture dependencies in large-dimensional data by building a network of connections among the hidden Extreme Learning Machine neurons. The fuzzy RNN models the temporal features in the classifier by inheriting the entire structure derived from the previous step. SenticNet's novelty lies in the fact that it engenders a decision-making process far superior to mere probabilistic guesses adopted by most NLP approaches and instead relies on implicit and semantic meanings associated with commonsense concepts.

SenticNet has been instrumental in providing realistic and unbiased insights on text analysis for the next steps in the pipeline and enhancing the quality of our dataset, thereby breaking the paradigm of relying on task-specific algorithms and creating a huge impact on the overall performance of our proposed approach.

Model Selection

As part of our research, we attempted to determine the correlation between financial and crypto datasets and noticed that there was minimal coherency with respect to word tokens present in the two corpora. SenticNet Parser was applied on text data to extract key tokens using priority queues and the commonality between the Financial Phrase Bank dataset, and the extracted twitter dataset was observed. SenticNet Parser was applied on text data to extract key tokens using priority queues and the commonality between the Financial Phrase Bank dataset, and the extracted twitter dataset was observed. The results showed that out of the top 300 most relevant words in the corpus, only 0.5% matched with the real-time extracted tweets. Figure 2 shows a smaller version of the results obtained from the above comparison. It is evident that amongst the top 40-50 most occurring tokens, 'new' was the only common token amongst both corpora. Other statistical analyses like maximum likelihood estimations too resulted in low similarity scores.

Figure 2. Corpora Token Comparison for Model Selection
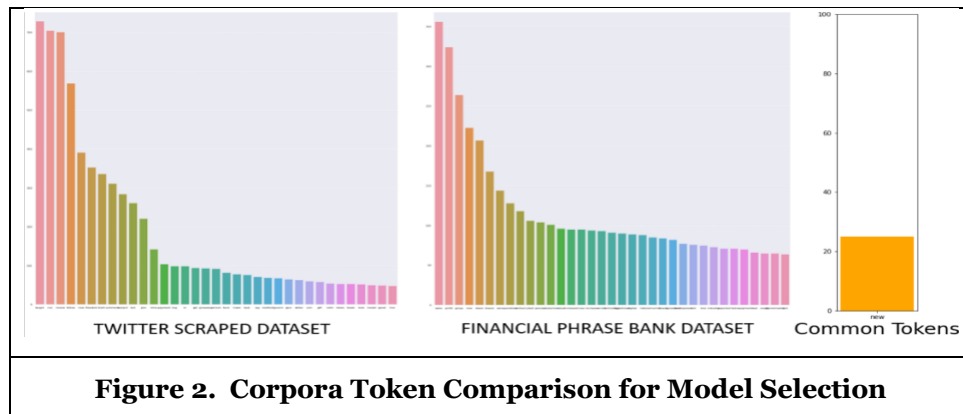
This lets us hypothesize that the Bitcoin-based Twitter data is not purely a financial textual matter. We performed the same tests every day for a 1-month period on the Sentiment140 dataset and the results obtained had extremely high correlations with the actual tweets establishing consistency. We experimented with various models for sentiment classification of tweets and finally used FinBERT for our pipeline as given in Section "Experimental Results". These experiments were instrumental in model selection for polarity detection of tweets.

FinBERT fine-tuning

The accuracy of financial sentiment analysis (Xing et al. 2020) is important for downstream applications. In the financial world, the professional vocabulary and expressions are different with respect to the daily colloquial language that general domain sentiment analysis models deal with. Hence, to harness the power of financial text mining along with state-of-the-art language model capabilities, we use the FinBERT (Liu et al. 2020) architecture of tweet sentiment classification. FinBERT is trained through multi-task learning on a large financial corpus for transferring knowledge from financial world data. It is trained for solving six self-supervised tasks which enable it to extract more knowledge and capture underlying semantic information.

As inferred from model selection, the Twitter data which is used for our study has a small correlation to Financial Phrase bank (a very finance-specific corpus). In order to tune our model to a more general conversational twitter corpus, we decided to apply transfer learning and fine-tune our model on Stanford's Sentiment140 dataset. This approach is novel as it enables the FinBERT to accommodate twitter expression and conversational jargon while maintaining the ability to capture financial details in the tweet if any.

Feature Engineering

Non-textual features of a tweet such as the likes, retweets, and similar features reinforce the sentiment of the tweet. Intuitively speaking, a positive tweet shared a greater number of times has a much larger outreach and should have a larger impact on the market sentiment than the other low-profile tweets. To take this into account, we incorporated retweet count, like count, and reply count along with sentiment analysis in our approach. In order to achieve this, we designed an algorithm that rules out hardcoding and uses statistical probabilities to empirically calculate feature weights. We use a multi-step approach for the same. Since the non-textual features reinforce the sentiment, we use sentiment scores as the target feature and the non-textual features as the descriptive features.

Firstly, The Shannon entropy is calculated for sentiment scores obtained from FinBERT across the entire dataset. We then calculate the remaining entropy of individual descriptive features with respect to the target feature. The information gains *IG* for every descriptive feature *f* is calculated as shown in (1). *H(T)* signifies the entropy of that feature, *H(T|f)* is the conditional entropy of target feature *T* with respect to *f,* and *n(x)* is the tweet count for *x, n_{total}* is the total tweet count and *m* is the number of features. We calculate the weight of every feature by using the formula denoted by (2). In (2), the *m^{th}* root of the *Information Gain* ratio is taken where *n* is the number of features which is 3 in our case. We note that the weights are static throughout our dataset. A one-to-one mapping was needed for the Twitter sentiment and the Bitcoin volume traded per hour. Hence, the tweet score for every hour *i* is calculated by summation of the tweet score of all tweets as in (3). Individual tweet score is calculated by multiplying the absolute value of the sentiment score by the weighted average of non-textual features using weights obtained from (2). The final value is obtained which encompasses textual and non-textual features of all tweets within an hour. In (3), $W^{feature}$ denotes the weight of the respective feature.

$$\text{Information\_Gain}(T, f) = H(T) - \sum_{i=1}^{m} H\left({}^{T}/_{f}\right) - \left(n\left({}^{T}/_{f}\right)/n_{total}\right) \qquad \textbf{(1)}$$

$$\text{weight}(f) = (IG_f/IG_{mean})^{1/m} \qquad \textbf{(2)}$$

$$TweetScore_i = \sum_{j=1}^{n} \left|\left(SentimentScore_j\right)\right| * \left(W_j^{likes} * n_j^{likes} + W_j^{retweet} * n_j^{retweet} + W_j^{reply} * n_j^{reply}\right) \qquad \textbf{(3)}$$

Experimental Results

Model	Accuracy (%)
VADER	12.72
SenticNet Polarity Classifier	8.80
Table 1. Comparison of models on StockTwit dataset	

We use tweets extracted using Twitter Academic API. While experimenting with models for sentiment analysis, we conducted a comparative analysis between VADER and SenticNet Polarity Classifier to study their classification capabilities on free text containing financial details. For this purpose, we used a small StockTwit corpus for testing. The results as shown in Table 1 indicated that both models failed to capture the sentiment associated with free texts having a combination of general and financial settings. Hence, we decided to fine-tune and apply FinBERT in our proposed approach. We use the Sentiment140 dataset for its robustness to fine-tune FinBERT. Since a small corpus is sufficient for transfer learning, we reduced the size of the dataset from 1.6 million to 30000.

Activation	Accuracy (%)	F1 Score	Loss
Softmax	82.14	0.7516	0.62
Sigmoid	**83.52**	**0.8315**	**0.42**
Table 2. Evaluation metrics of FinBERT			

The FinBERT architecture employed in our approach has 12 hidden layers and 12 attention nodes. The "bert base uncase" tokenizer was used to convert the text into tokens for feeding into the BERT model. The model was trained for 20 epochs as saturation was observed at around 17 epochs for every configuration. We experimented with two combinations for the final layer of BERT during transfer learning, Categorical Cross-Entropy Loss + Softmax activation and Binary Cross-Entropy Loss + Sigmoid activation. As evident from Table 2., BCE loss with sigmoid activation yields much higher accuracy, hence we use this configuration throughout the rest of the approach. The output of the model has probabilities of both "positive" and "negative" labels, the final score from which is computed by taking their difference (positive-negative). The FinBERT model is trained on two NVIDIA GeForce RTX 3090 GPUs.

| Model	Accuracy (%)	F1 Score
Our Approach	**83.52**	**0.8315**
Word2vec + RNN	82.81	0.8296
Word2vec + CNN	80.06	0.8006
SVM	75.61	0.7649
Naive Bayes	75.19	0.7505
Majority Voting	74.80	0.7671
Logistic Regression	74.15	0.7521
Random Forest	71.76	0.7470
Attention-LSTM	69.29	0.6948
FCDNN	56.67	0.5654
Table 3. Comparison of performance on Sentiment140 dataset		

Table 3 depicts the accuracy and F1 scores of our model against other machine learning approach benchmarked on the Sentiment140 dataset. Our proposed model outperforms all other methods with an accuracy of 83.52 and F1 score of 83.15. From the above Table 3, we see that traditional machine learning algorithms like SVM and Naive Bayes perform moderately with accuracies of around 75%. On the other hand, we notice that the word embedding technique coupled with neural networks (Dang, 2020) enhances the performance of the classifier achieving scores over 80%. It can also be observed that attention-based LSTM and FCDNN (Zhang et al. 2019), two compound deep learning models perform extremely poorly on the dataset.

After sentiment analysis on the dataset, the Bitcoin volumes are analyzed with the overall tweet scores as in the subsection "Feature Engineering" for every hour to determine a correlation between the two metrics. It is intuitive that if sentiment analysis has a role to play, it would not instantaneously reflect upon the market depth. There would be a finite gap between the creation of tweets and their impact on the Bitcoin market. We incorporate this crucial observation in our research study by estimating the correlation at a regular interval of 2 hours. A t-delay of T hours signifies that the difference between the timestamp of volume and tweet is T hours.

T-Delay	Pearson	P-value	Lasso	LSTM Loss
0	0.1584	$3.9 * 10^{-12}$	0.021	0.009
2	0.1161	$4.3 * 10^{-10}$	0.021	0.008
4	0.0919	$7.8 * 10^{-9}$	0.009	0.001
6	0.0902	$8.1 * 10^{-9}$	0.006	0.006
8	0.0828	$8.9 * 10^{-9}$	0.002	0.012
Table 4. Tweet Score - Bitcoin Volume correlation				

We use multiple metrics to evaluate the dependency of Bitcoin volume on social media sentiment. Pearson's R along with the Lasso regression model was applied to determine the presence of a linear relationship between the two input parameters. We choose Lasso over linear regression as it offers efficient regularization and reduces the complexity of the model while retaining all the variables. We also compare the Pearson's R values at different intervals to further affirm the inference obtained for linear dependency. A Deep LSTM architecture is employed to detect a nonlinear relationship. The trend in the loss obtained during validation explains the dependency as we change T-delay. Moreover, a very small loss is also indicative of a more complex relationship that is captured by RNN. In order to validate the significance of the results, we also calculate the p-values for every T-delay. A small p-value signifies the minimal influence of the null hypothesis on the results.
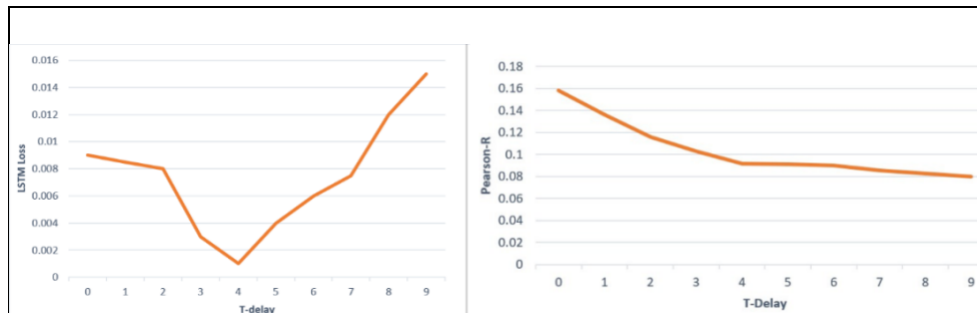
Figure 3. The bell-curve like relation between t-delay and LSTM Loss

The results obtained in Table 4. show that there is no significant linear relationship between social media sentiment and Bitcoin volume. The Lasso regression scores at all intervals are too small to convey linearity between the two inputs. Even though The Pearson-R scores show a downward trend as in Figure 3, the maximum value is 0.1584, which is not sufficient for a concrete relationship. Since the P-values are small, hence the Pearson Correlation calculated is statistically significant. Hence, we conclude that there is no linear relationship between Twitter sentiment and Bitcoin volume.

Table 4. also depicts a bell-curve-like relationship in the LSTM Loss between the two inputs by varying the hour gap. It can be observed in Figure 3 that the curve achieves minima around the 3–4-hour gap range, thereby affirming the impact of a tweet is most visible on the Bitcoin volume after 3-4 hours. The small value of LSTM losses also indicates a form of non-linear relationship which cannot be ascertained by simple correlation models. Lastly, the extremely minute p-values obtained in the experiment highlight that the null hypothesis is absent, and the results are definitive. Beyond 3 -4 hour horizon, we find that the sentiment and volumes do not have a substantial correlation and hence no strong conclusion can be derived for the same at this stage. On testing for T-delay over 12 hours, we found that the p-values were greater than 0.5. The p-value for a T-delay of 36 hours was 0.8655 and the Pearson-R score was close to 0. The high p-value stipulated the dominance of the null hypothesis and the statistical insignificance of the results. Hence, we limited the scope of the T-delay in our study to 8 hours.

Conclusion and Future Works

This study empirically determines the impact of social media sentiment on the volume of Bitcoin traded. We adopted the novel FinBERT architecture and fine-tuned it to obtain a high sentiment classification accuracy of 83.52% on the sentiment140 dataset. Information gain was used to incorporate non-textual features and detect correlation using various metrics. We discover that there is no strong linear relationship between Bitcoin trading volume and lagged social media sentiments. However, experiments with LSTM losses indicate a more complex, non-linear relationship between the two features, and the maximum dependency is visible at an predictive horizon of 3-4 hours. Our study is an important step toward understanding the volatility and the future price of the cryptocurrency market and its interaction with social media content using textual analysis.

There are certain limitations to our approach. An extremely important direction of future work is to employ the pipeline on other cryptocurrencies such as Ethereum, Cardano, Solana; not just in Finance but also on data from sectors like Accounting in order to research extensively and generalize the observations found in this paper. The entire study is based on the dataset extracted using the Twitter API. Hence the type of correlation that has been determined may be localized. The dataset contains all types of tweets having hashtags as mentioned in previous sections. Hence certain tweets containing promotional or irrelevant content may have distorted the tweet score for every hour. Another limitation that can be observed is that the weights obtained by information gain are subjective to the current dataset. Further works also include using the granularity offered by SenticNet with respect to Hourglass of emotions and combine with other valences to gather insights on impact of specific emotions. Due to restrictions on resources, we were not able to do modeling on an extensive dataset that would yield constant weights applicable for every dataset. Another interesting direction of future research is to deduce the cause for the lack of correlation after a 3-4 hour time horizon. We aim to improve on some of these drawbacks by designing an algorithm to filter tweets that do not contribute to the study. We are also working towards developing a robust evaluation metric for accurately detecting the non-linear relation in this study. We also plan on incorporating Google trends and total Bitcoin tweet volume as key features for re-evaluating the hypotheses.

References

Abraham, J., Higdon, D., Nelson, J., and Ibarra, J. 2018. "Cryptocurrency price prediction using tweet volumes and sentiment analysis." *SMU Data Science Review* (1:3), pp. 1-22.

Balcilar, M., Bouri, E., Gupta, R., and Roubaud, D. 2017. "Can volume predict Bitcoin returns and volatility? A quantiles-based approach," *Economic Modelling* (64), pp. 74–81.

Bouri, E., Lau, C. K., Lucey, B., and Roubaud, D. 2019. "Trading volume and the predictability of return and volatility in the cryptocurrency market," Finance Research Letters (29), pp. 340–346 (doi: 10.1016/j.frl.2018.08.015).

Bouri, E., Roubaud, D., and Shahzad, S. 2020. "Do Bitcoin and other cryptocurrencies jump together?", The Quarterly Review of Economics and Finance (76), pp. 396-409(doi: 10.1016/j.qref.2019.09.003).

Cambria, E., Liu, Q., Decherchi, S., Xing, F. and Kowk, K. 2022. "SenticNet 7: A commonsense-based neurosymbolic AI framework for explainable sentiment analysis," in *Language Resources and Evaluation Conference*, Marseille, France, pp. 3829-3839.

Gao, X., Huang, W., and Wang, H. 2021. "Financial Twitter Sentiment on Bitcoin Return and High-Frequency Volatility", Virtual Economics (4:1), pp. 7-18(doi: 10.34021/ve.2021.04.01(1).

Hassan, A., Mas' ud, M. Z., Shah, W. M., Abdul-Latip, S. F., Ahmad, R., Ariffin, A., and Yunos, Z. 2020. "A systematic literature review on the security and privacy of the blockchain and cryptocurrency," *OIC-CERT Journal of Cyber Security* (2:1), pp. 1–17.

Ho, C., Damien, P., Gu, B., and Konana, P. 2017. "The time-varying nature of social media sentiments in modeling stock returns", Decision Support Systems (101), pp. 69-81(doi: 10.1016/j.dss.2017.06.001).

Huynh, T. 2021. "Does Bitcoin React to Trump's Tweets?", Journal of Behavioral and Experimental Finance (31), p. 100546(doi: 10.1016/j.jbef.2021.100546).

Karasu, Seçkin 2018. "Prediction of Bitcoin prices with machine learning methods using time series data." 2018 26th signal processing and communications applications conference (SIU). IEEE.

Liu, Z., Huang, D., Huang, K., Li, Z. and Zhao, J. 2020. "FinBERT: A Pre-trained Financial Language Representation Model for Financial Text Mining." In: IJCAI, pp. 4513–4519.

Ma, F., Liang, C., Ma, Y., and Wahab, M. 2020. "Cryptocurrency volatility forecasting: A Markov regime-switching MIDAS approach", Journal of Forecasting (39:8), pp. 1277-1290(doi: 10.1002/for.2691).

Mohapatra, S., Ahmed, N., and Alencar, P. 2019. "KryptoOracle: A Real-Time Cryptocurrency Price Prediction Platform Using Twitter Sentiments." In: *IEEE International Conference on Big Data*, pp. 5544–5551. (10.48550/ARXIV.2003.04967)

Nasir, M., Huynh, T., Nguyen, S., and Duong, D. 2019. "Forecasting cryptocurrency returns and volume using search engines", Financial Innovation (5:1)(doi: 10.1186/s40854-018-0119-8).

Rogalski, R. 1978. "The Dependence of Prices and Volume", The Review of Economics and Statistics (60:2), p. 268(doi: 10.2307/1924980).

Sabri, N. 2008. "The Impact of Trading Volume on Stock Price Volatility in the Arab Economy", SSRN Electronic Journal (14:3), pp. 285–298 (doi: 10.2139/ssrn.1097624).

S Kumar, A., and Ajaz, T. 2019. "Co-movement in crypto-currency markets: evidences from wavelet analysis", Financial Innovation (5:1)(doi: 10.1186/s40854-019-0143-3).

Shmueli, G., and Koppius, O. 2010. "Predictive Analytics in Information Systems Research", SSRN Electronic Journal (35:3), pp. 553–572 (doi: 10.2139/ssrn.1606674).

Susanto, Y., Livingstone, A., Ng, B., and Cambria, E. 2020. "The Hourglass Model Revisited", IEEE Intelligent Systems (35:5), pp. 96-102 (doi: 10.1109/mis.2020.2992799).

Szetela, B., Mentel, G., Bilan, Y., and Mentel, U. 2021. "The relationship between trend and volume on the bitcoin market", Eurasian Economic Review (11:1), pp. 25-42(doi: 10.1007/s40822-021-00166-5).

Valencia, F., Gómez-Espinosa, A., and Valdés-Aguirre, B. 2019. "Price Movement Prediction of Cryptocurrencies Using Sentiment Analysis and Machine Learning", Entropy (21:6), p. 589(doi: 10.3390/e21060589).

Walther, T., Klein, T., and Bouri, E. 2019. "Exogenous drivers of Bitcoin and Cryptocurrency volatility – A mixed data sampling approach to forecasting", Journal of International Financial Markets, Institutions and Money (63), p. 101133(doi: 10.1016/j.intfin.2019.101133).

Xing, F., Cambria, E., and Welsch, R. 2017. "Natural language based financial forecasting: a survey", Artificial Intelligence Review (50:1), pp. 49-73(doi: 10.1007/s10462-017-9588-9).

Xing, F. Z., Malandri, L, Zhang, Y., and Cambria, E. 2020. "Financial Sentiment Analysis: No Silver Bullet and An Investigation into the Common Mistakes," in *International Conference on Computational Linguistics*, Barcelona, Spain, pp. 978-987 (10.18653/v1/2020.coling-main.85).

Zhang, Y., Song, D., Zhang, P., Li, X., and Wang, P. 2019. "A quantum-inspired sentiment representation model for twitter sentiment analysis", Applied Intelligence (49:8), pp. 3093-3108(doi: 10.1007/s10489-019-01441-4).