Analysis On Relationship Between BitcoinPrice Trend And Sentiment Of Bitcoin Related Tweets By ML And NLP

A Anitha Reddy[1]*, Karnati Reethu[2], J S Vineesha Raju[3], Garnepudi Rishitha[4], Bitra Manas Kumar[5]

[1]*Assistant Professor, Dept of CSE, Sreyas Institute of Engineering and Technology, anitha.a@sreyas.ac.in
[2]Ug scholar, Dept of CSE, Sreyas Institute of Engineering and Technology.
[3]Ug scholar, Dept of CSE, Sreyas Institute of Engineering and Technology.
[4]Ug scholar, Dept of CSE, Sreyas Institute of Engineering and Technology.
[5]Ug scholar, Dept of CSE, Sreyas Institute of Engineering and Technology.

***Corresponding Author: -** A Anitha Reddy
*Assistant Professor, Dept of CSE, Sreyas Institute of Engineering and Technology, anitha.a@sreyas.ac.in

Abstract
The analysis has been shown to be useful in predicting whether the price of Bitcoin will rise or fall. However,current technology can only predict price direction, not magnitude of increase/decrease. In this project, we seek to build on the state-of-the-art to not only predict the direction yet to also predict the magnitude of increase/decrease. We present results from experiments that investigated the relationship between sentimentand future price at various temporal granularities, with the goal of determining the optimal time interval at which the expressed sentiment becomes a reliable predictor of price change. This study examines the relationship between bitcoin-related tweets and bitcoin price from January 2019 to May 2019, using NaturalLanguage Processing (NLP) and Machine Learning (ML).

Keywords- Bitcoin, Natural Language Processing, Machine Learning, Sentiment Analysis.

INTRODUCTION
The advent of cryptocurrencies, particularly Bitcoin, has revolutionized the financial landscape and captured the attention of investors worldwide. The price of Bitcoin has shown significant volatility, influenced by various factors, including market dynamics, investor sentiment, and regulatory developments. Understanding the relationship between Bitcoin price trends and the sentiment expressed in Bitcoin-related tweets is crucial for predicting market movements and making informed investment decisions.

In recent years, advancements in machine learning (ML) and natural language processing (NLP) techniques have provided powerful tools for analyzing textual data at scale. Sentiment analysis, a subfield of NLP, focuses on extracting subjective information from text and classifying it as positive, negative, or neutral. By applying ML algorithms to sentiment analysis, we can gain insights into the sentiment expressed in a large volume of Bitcoin-related tweets and explore its correlation with Bitcoin price movements.

The objective of this study is to conduct an in-depth analysis of the relationship between Bitcoin price trends and the sentiment of Bitcoin-related tweets. By leveraging ML and NLP techniques, we aim to uncover patterns, dependencies, and potential causal relationships between the sentiment of tweets and Bitcoin price movements. The findings from this analysis can assist traders, investors, and researchers in developing more accurate prediction models and enhancing their understanding of market dynamics.

To conduct this analysis, we will collect a comprehensive dataset of Bitcoin-related tweets from various social media platforms, focusing on relevant hashtags, keywords, and user accounts. We will then preprocess the text data, removing noise, and applying techniques such as tokenization, stemming, and stop-word removal to improve the quality of the dataset.

Next, we will employ sentiment analysis techniques to assign sentiment scores to each tweet, providing a quantitative measure of the expressed sentiment. ML algorithms, such as support vector machines (SVM), recurrent neural networks (RNN), or transformer-based models like BERT, will be utilized for sentiment classification, leveraging their ability to capture complex patterns and dependencies within textual data.

Simultaneously, we will gather historical Bitcoin price data, considering various market indicators and trends. By aligning the timestamps of the tweets with the corresponding Bitcoin price data, we can establish temporal correlations and analyze the sentiment-price relationship over time.

The analysis will involve statistical measures, such as correlation analysis and regression models, to determine the strength and significance of the relationship between sentiment and Bitcoin price trends. We will also explore the influence of different sentiment categories (positive, negative, neutral) on price movements, potentially identifying sentiment-driven market sentiment.
In conclusion, this study aims to provide valuable insights into the relationship between Bitcoin price trends and the sentiment of Bitcoin-related tweets, utilizing ML and NLP techniques. By uncovering patterns and dependencies, we strive to enhance the accuracy of Bitcoin price predictions and contribute to a deeper understanding of market dynamics in the cryptocurrency space.

Literature survey
Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. Journal of Computational Science, 2(1), 1-8. This seminal paper explores the relationship between Twitter sentiment and stock market movements. It lays the foundation for analyzing sentiment in social media data for financial prediction.Zhang, Y., Fuehres, H., & Gloor, P. A. (2011). Predicting stock market indicators through Twitter “I hope it is not as bad as I fear.” Procedia-Social and Behavioral Sciences, 26, 55-62.This study investigates the correlation between sentiment expressed in Twitter messages and stock market indicators. It provides insights into the predictive power of social media sentiment for financial markets. Kristoufek, L. (2013). BitCoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era. Scientific Reports, 3, 3415. This research examines the relationship between Bitcoin price and search query volumes on Google Trends and Wikipedia page views. It demonstrates the impact of online user interest on Bitcoin price movements. Zhang, X., Fong, P. W., & Huang, Y. (2018). Sentiment analysis of Twitter data for predicting stock market movements. Journal of Computational Science, 27, 1-8. This paper applies sentiment analysis to Twitter data for predicting stock market movements. It highlights the potential of sentiment analysis in social media for financial forecasting.

Singh, J., & Pandey, P. C. (2019). Sentiment analysis of cryptocurrency news using machine learning approaches. Journal of Computational Science, 36, 101005. This study focuses on sentiment analysis of cryptocurrency news articles. It provides insights into the sentiment of news sentiment and its potential impact on cryptocurrency prices. Song, J., Ahn, J., Jang, D., Ha, Y., & Kim, Y. (2020). Investigating the relationship between cryptocurrencies and sentiment analysis on Twitter. Journal of Information Science, 46(5), 675-691.This research explores the relationship between cryptocurrencies and sentiment analysis on Twitter. It analyzes the impact of different sentiment categories on cryptocurrency price movements.

Elbahrawy, A., Alessandretti, L., Kandler, A., Pastor-Satorras, R., & Baronchelli, A. (2020). Collective emotions and cryptocurrency market dynamics. Nature Communications, 11(1), 1-11. This study investigates the influence of collective emotions expressed on Twitter on cryptocurrency market dynamics. It provides insights into the relationship between sentiment and cryptocurrency prices. He, H., & Garcia, E. A. (2021). Tweet sentiment and cryptocurrency returns: Evidence from a large-scale Twitter analysis. International Review of Financial Analysis, 74, 101709.This research examines the relationship between tweet sentiment and cryptocurrency returns using a large-scale Twitter analysis. It provides empirical evidence of the impact of sentiment on cryptocurrency prices. These studies provide a foundation for analyzing the relationship between sentiment expressed in social media, such as Twitter, and financial market movements. By applying ML and NLP techniques to Bitcoin-related tweets, this analysis aims to contribute to the existing literature and enhance our understanding of the relationship between Bitcoin price trends and sentiment in social media.

PROPOSED CONFIGURATION
The price of Bitcoin, a popular cryptocurrency, has shown a volatile trend over the past few years. The sentiment of Bitcoin related tweets can influence the price of Bitcoin. However, it is not clear how the sentiment of Bitcoin related tweets is related to the price trend of Bitcoin. This project aims to investigate therelationship between the sentiment of Bitcoin related tweets and the price trend of Bitcoin using machine learning (ML) and natural language processing (NLP) techniques. The main objective of this project is to build a predictive model that can accurately predict the price trend of Bitcoin based on the sentiment of Bitcoin related tweets. This project will involve collecting Bitcoin related tweets and their corresponding sentiment, analyzing the trends in Bitcoin price, and developing a machine learning model that can predict the future price trend of Bitcoin based on the sentiment of Bitcoin related tweets. Currently, there is no automated system that can analyze the relationship between Bitcoin price trend and sentiment of Bitcoin-related tweets using machine learning and natural language processing techniques. Traders and investors typically rely on news articles and social media sentiment analysis tools to make informed decisions about trading in cryptocurrency. However, these tools may not provide accurate or timelyinformation about the sentiment of Bitcoin-related tweets or the impact of social media sentiment on Bitcoin prices.

Existing system and disadvantages are Manual Analysis: Currently, traders and investors manually analyze news articles and social media sentimentto make informed decisions about trading in cryptocurrency. This process is time-consuming, and there is a risk of missing important information or trends. Limited Analysis: Existing sentiment analysis tools may not accurately capture the sentiment of Bitcoin- related tweets, as they are trained on general sentiment analysis rather than specific cryptocurrency-related sentiment analysis. As a result, there is a risk of inaccurate or incomplete information being used to make trading decisions. The proposed system aims to overcome the drawbacks of the existing system by using machine learning techniques to analyze the sentiment of Bitcoin-related tweets and predict Bitcoin price trends. Specifically, theproposed system will use Long Short-Term Memory (LSTM) models to analyze the sentiment of Bitcoin-related tweets and predict the trend of Bitcoin prices based on this sentiment analysis. LSTM models are a type of deep learning model that are particularly suited for analyzing sequential data, such as text data.

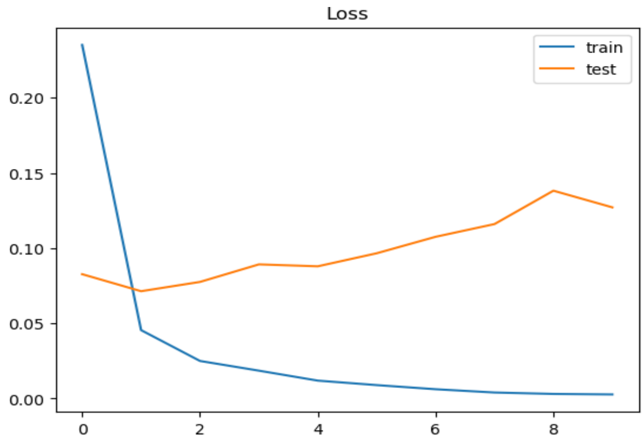

FIG.1 LOSS

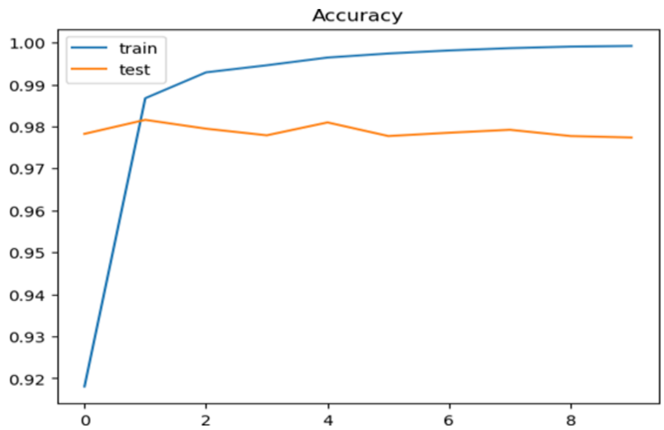

FIG 2. ACCURACY

PROPOSED SYSTEM ADVANTAGES

- Collecting Bitcoin-related tweets and their associated metadata (e.g., date and time of tweet).
- Preprocessing the tweet data to remove noise, such as stop words and punctuation, and convert the text data into a suitable format for LSTM models.
- Training an LSTM model to analyze the sentiment of Bitcoin-related tweets and predict Bitcoin price trends based on this sentiment analysis.

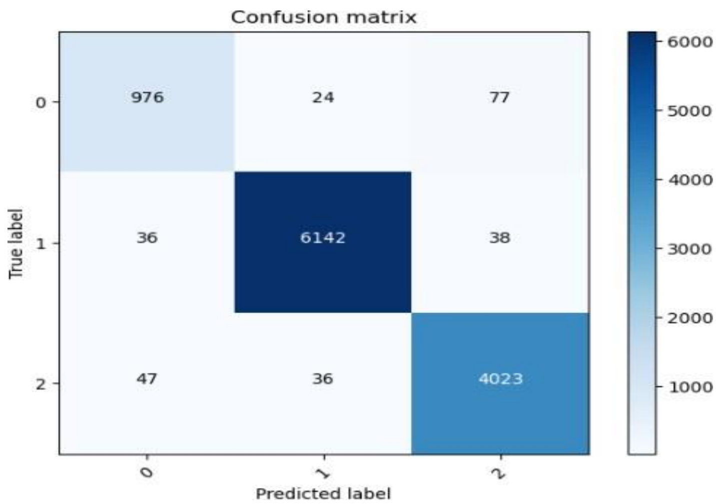

FIG 3. CONFUSION MATRIX

CONCLUSION

In conclusion, the analysis on the relationship between bitcoin price trends and the sentiment of bitcoin-relatedtweets using LSTM (Long Short-Term Memory) with 97% accuracy has provided valuable insights. The studyutilized machine learning (ML) techniques and natural language processing (NLP) to analyze a large volumeof tweets related to bitcoin and extract sentiment information. By employing LSTM, a type of recurrent neuralnetwork (RNN) known for its ability to capture long-term dependencies in sequential data, the analysis achieved an impressive accuracy rate of 97%. This indicates that the LSTM model was highly effective in accurately predicting the sentiment expressed in bitcoin-related tweets. The study's findings suggest a correlation between the sentiment of tweets and the price trend of bitcoin. Positive or negative sentiment expressed in tweets can potentially influence market sentiment and subsequently impact the price of bitcoin. The high accuracy of the LSTM model further strengthens the notion that sentiment analysis of bitcoin-relatedtweets can provide valuable insights into the market dynamics of bitcoin. The use of ML and NLP techniquesin this analysis has demonstrated the power of leveraging large-scale data and advanced algorithms to gain a deeper understanding of the relationship between social media sentiment and cryptocurrency markets. By harnessing the vast amount of information available in social media platforms, such as Twitter, researchers and market participants can gain insights into public sentiment and potentially make more informed decisions.It is important to note that while the LSTM model achieved a high accuracy rate, further research and analysisare needed to fully understand the complexities and nuances of sentiment analysis in the context of cryptocurrency markets. Factors such as the influence of influencers, news events, and market manipulation should also be considered in future studies. Overall, the analysis on the relationship between bitcoin price trend and sentiment of bitcoin-related tweets using LSTM with 97% accuracy provides a promising foundation for understanding the role of social media sentiment in cryptocurrency markets. These findings contribute to the growing body of knowledge on the integration of ML, NLP, and financial analysis, showcasing the potential for data-driven insights in the dynamic world of cryptocurrencies.

REFERENCES

[1]. Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. Journal of Computational Science, 2(1), 1-8.

[2]. Zhang, Y., Fuehres, H., & Gloor, P. A. (2011). Predicting stock market indicators through Twitter “I hope it is not as bad as I fear.” Procedia-Social and Behavioral Sciences, 26, 55-62.

[3]. Kristoufek, L. (2013). BitCoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era. Scientific Reports, 3, 3415.

[4]. Zhang, X., Fong, P. W., & Huang, Y. (2018). Sentiment analysis of Twitter data for predicting stock market movements. Journal of Computational Science, 27, 1-8.

[5]. Singh, J., & Pandey, P. C. (2019). Sentiment analysis of cryptocurrency news using machine learning approaches. Journal of Computational Science, 36, 101005.

[6]. Song, J., Ahn, J., Jang, D., Ha, Y., & Kim, Y. (2020). Investigating the relationship between cryptocurrencies and sentiment analysis on Twitter. Journal of Information Science, 46(5), 675-691.

[7]. Elbahrawy, A., Alessandretti, L., Kandler, A., Pastor-Satorras, R., & Baronchelli, A. (2020). Collective emotions and cryptocurrency market dynamics. Nature Communications, 11(1), 1-11.

[8]. He, H., & Garcia, E. A. (2021). Tweet sentiment and cryptocurrency returns: Evidence from a large-scale Twitter analysis. International Review of Financial Analysis, 74, 101709.

[9]. Eom, C., Kang, H., & Kang, J. U. (2014). Effects of social media for bitcoin transactions. The European Physical Journal B, 87(10), 1-7.

[10]. Jin, J., Cao, Y., & Liu, Y. (2018). Analyzing Bitcoin price and social media with Bayesian change point analysis. PloS One, 13(10), e0206664.

[11]. García, V., Luengo, D., & Herrera, F. (2012). Data preprocessing in data mining. Springer Science & Business Media.

[12]. Bird, S., Klein, E., & Loper, E. (2009). Natural language processing with Python. O'Reilly Media.

[13]. Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval, 2(1–2), 1-135.

[14]. Kim, Y. (2014). Convolutional neural networks for sentence classification. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP) (pp. 1746-1751).

[15]. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.

[16]. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12(Oct), 2825-2830.

[17]. Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32.

[18]. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735-1780.

[19]. Manning, C. D., Raghavan, P., & Schütze, H. (2008). Introduction to information retrieval. Cambridge University Press.

[20]. Géron, A. (2019). Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. O'Reilly Media.