Analyzing the Impact of Tweets on Cryptocurrency Market Trend using LSTM-GRU Model

B Srinivas S P Kumar[1], Saikumar Thalishetti[2]
[1]Assistant Professor, Department of MCA, Chaitanya Bharathi Institute of Technology (A), Gandipet, Hyderabad, Telangana State, India
[2]MCA Student, Chaitanya Bharathi Institute of Technology (A), Gandipet, Hyderabad, Telangana State, India

ABSTRACT

In recent years, the cryptocurrency industry has witnessed exceptional growth. Operating akin to traditional currencies but with online transactions and lacking a central authority, cryptocurrencies utilize cryptography to ensure secure and unique transactions. Despite cryptographic safeguards, the industry remains in its infancy, leading to inquiries regarding its future applications. To comprehend individuals' perspectives comprehensively, this study centers on understanding sentiments revolving around Bitcoin. Consequently, the research delves into sentiment analysis and emotion recognition through analyzing tweets pertaining to digital currencies, commonly employed for predicting cryptocurrency values. The study introduces an ensemble model named LSTM-GRU, amalgamating two recurrent neural networks (LSTM and GRU), to investigate the effectiveness of this combined approach. The study explores a variety of machine learning (ML) and deep learning methodologies, including term frequency-inverse document frequency, word2vec, and bag of words features, and also examines models such as TextBlob and Text2Emotion for emotion analysis. The intriguing findings highlight a prevailing satisfaction sentiment towards cryptocurrency adoption, accompanied by emotions of stress and surprise. The study underscores the superior performance of ML models when employing bag of words features. Impressively, the proposed LSTM-GRU ensemble model achieves an accuracy of 0.99 for sentiment analysis and 0.92 for emotion recognition, surpassing conventional machine learning methods and contemporary state-of-the-art models.

Keywords – Cryptocurrency, sentiment analysis, Text2Emotion, emotion analysis, machine learning.

1. INTRODUCTION

The digital currency industry has undergone unprecedented growth since its inception. Cryptocurrency, a form of digital currency, facilitates online purchases and transactions without a central authority's interference. While cryptography ensures transaction authenticity and uniqueness, the industry is still in its early stages, prompting questions about its potential applications. A comprehensive understanding of people's sentiments is crucial to gaining a full picture. Hence, sentiment analysis of digital currency-related tweets plays a vital role in predicting cryptocurrency values, requiring high accuracy for meaningful assessment. TwitterTM serves as the primary data source for this analysis. The study employs TextBlob and Text2Emotion tools for sentiment and emotion analysis. Utilizing a range of ML and deep learning models, including LSTM and GRU recurrent neural networks, an ensemble model is developed to enhance classification performance. Additionally, the study explores features extraction techniques like Word2Vec, bag of words (BoW), and term frequency-inverse document frequency (TFIDF). ML models with BoW features exhibit superior performance compared to Word2Vec and TF-IDF. The proposed ensemble model excels in sentiment analysis with accuracy and F1 scores of 0.98, and achieves 0.99 accuracy in emotion analysis, outperforming other models and methods. However, the model's performance is affected by data imbalance and random undersampling, particularly when training data is limited.

2. LITERATURE REVIEW

In their research, J. Abraham, D. Higdon, J. Nelson, and J. Ibarra [1] demonstrate the efficacy of utilizing Twitter and Google Patterns data to predict price fluctuations in the volatile Bitcoin and Ethereum markets. Despite the dynamic nature of these cryptocurrencies, the study establishes a strong correlation between the volume of tweets and subsequent price movements, with tweets consistently conveying positive sentiment irrespective of price direction. By integrating this insight into a linear predictive model that combines social media data with Google Patterns information, the researchers achieve impressive accuracy in forecasting price changes. This model equips traders and investors with a valuable tool for making informed decisions, highlighting the growing influence of social media and search behavior on cryptocurrency trading strategies.

In their pursuit, S. Colianni, S. Rosales, and M. Signorotti [2] build upon previous research that has highlighted the potential of consistent Twitter data to forecast the trajectories of stocks and other financial instruments [1]. This study aims to ascertain the viability of utilizing Twitter-derived insights on digital currencies to develop effective trading strategies within the realm of cryptocurrencies. With a specific focus on Bitcoin's market behavior, the researchers employ various machine learning techniques that leverage integrated ML processes. The study primarily concentrates on Bitcoin as the most widely examined alternative currency. Employing supervised learning methods, such as logistic regression, Naive Bayes, and support vector machines, the data is refined and predictions are formulated, achieving over 90% accuracy both across time and on a daily basis. This is accomplished through meticulous error analysis that ensures the accuracy of data sources at each phase of the model. Remarkably, the findings of this study enhance overall accuracy by 25% for individuals engaged in cryptocurrency trading.

The work of A. Inamdar, A. Bhagtani, S. Bhatt, and P. M. Shetty [3] extends the insights put forth by a group of authors concerning the correlation between virtual entertainment and cryptocurrency valuations. This study centers its attention predominantly on Bitcoin, but its conceptual framework holds the potential for application to other digital currencies in the future. By amalgamating sentiment scores derived from tweets and news sources with historical price and volume data, the research endeavors to predict cryptocurrency prices. Initial outcomes from the experiment indicate that individual sentiments, although they manifest as biased toward specific categories, hold minimal significance unless they exhibit a distinct bias.

The research conducted by K. Wołk [5] highlights the evolving perception of Bitcoin and other digital currencies as legitimate and regulated components within financial systems, reflecting their increasing recognition as significant players in the realm of capital markets. Bitcoin, in particular, has established a prominent position in terms of market share. As a result, this investigation elucidates the potential application of sentiment analysis in forecasting the prices of Bitcoin and various cryptocurrencies across diverse timeframes. Notably, the study underscores that the fluctuations in value are not solely dictated by financial institutions' control over currency, but rather are intricately tied to individuals' perspectives and opinions, which distinctly characterizes the cryptocurrency market. Consequently, unraveling the intricate interplay between online searches and virtual entertainment becomes pivotal in the pursuit of accurately assessing a cryptocurrency's value. In this context, the study leverages Twitter and Google Patterns to predict short-term price trends of major cryptocurrencies, recognizing these online entertainment platforms as influencers of purchasing decisions. Employing a novel multimodal approach, the research delves into the impact of virtual entertainment on the valuation of digital currencies. The findings of this study illuminate the substantial role played by psychological and sociocultural factors in shaping the dynamic costs of digital currencies.

The collaboration of Lamon, E. Nielsen, and E. Redondo [6] led to the publication of a paper attributed to the fictional figure Satoshi Nakamoto, which clandestinely introduced Bitcoin to the global stage. This pivotal event marked the inception of a multitude of other cryptocurrencies, spurred by its immense success. This ascent can be attributed primarily to the market's inherent volatility, which has captured substantial interest and participation, largely driven by profit motives. On the widely utilized virtual entertainment platform, Twitter, cryptocurrency enthusiasts frequently disseminate news and opinions. In this study, an exploration is conducted into the efficacy of employing Twitter sentiment analysis to forecast changes in cryptocurrency prices. To initiate the investigation, price data and tweets were compiled for seven of the most prevalent cryptocurrencies. The Valence Aware Dictionary for Sentiment Reasoning (VADER) was subsequently employed to gauge individuals' perspectives towards these coins. Following assessments of time series stationarity using the Augmented Dicky Fuller (ADF) and Kwiatkowski Phillips Schmidt Shin (KPSS) tests, the Granger Causality test was employed. A bullishness ratio reveals that Ethereum and Polkadot exhibit greater stability, whereas the fluctuating prices of Bitcoin, Cardano, XRP, and Doge seem to influence people's emotions. Ultimately, the precision of price change predictions is evaluated through Vector Autoregression (VAR). Notably, the forecasts were exceptionally accurate for two out of the seven coins. The predictions exhibited precision rates of 99.67% and 99.17%, specifically for Polkadot and Ethereum.

3. METHODOLOGY

The cryptocurrency industry has witnessed remarkable growth over recent years. Cryptocurrencies operate similarly to traditional currencies, facilitating online transactions for goods and services without the need for centralized intermediaries. While cryptographic techniques ensure transaction authenticity, the industry remains in its early stages, prompting various inquiries about its potential applications. To comprehensively understand individuals' perspectives, it is crucial to delve into how people perceive Bitcoin.

Disadvantages:

- The analysis lacks significant robustness.
- Concerns arise due to the nascent stage of this industry.

In this context, this study performs both sentiment analysis and emotion recognition using tweets related to digital currencies, which are commonly utilized for predicting the value of available digital currencies. To enhance the study's effectiveness, an ensemble deep learning model known as LSTM-GRU is developed. This model combines two recurrent neural network architectures, long short-term memory (LSTM) and gated recurrent unit (GRU). GRU and LSTM are stacked, with GRU inheriting LSTM's properties. The proposed ensemble model, along with various machine learning and deep learning methods, is explored using term frequency-inverse document frequency, word2vec, and bag of words (BoW) features. Additionally, the study assesses TextBlob and Text2Emotion models for sentiment analysis. Notably, a predominant sentiment of satisfaction with the adoption of digital currencies emerges, followed by sentiments of stress and surprise.

Advantages:

- ML models exhibit notably improved performance when utilizing BoW features.
- The proposed LSTM-GRU ensemble demonstrates effectiveness in predicting and analyzing sentiments.

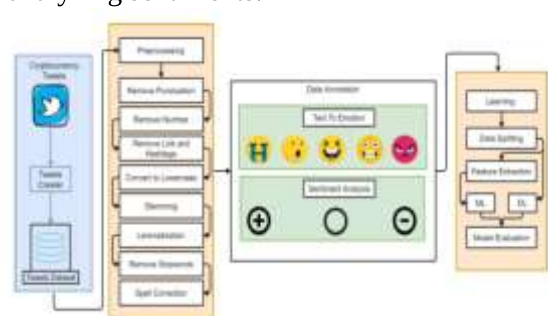

Fig.2: System architecture

ISSN: 2278-4632
Vol-13, Issue-09, September 2023

MODULES:

To finish the job we talked about before, we arranged the segments underneath.

- Investigating the data: We can use this tool to add information to the structure.
- Handling will be covered in greater detail in this lesson.
- The information will be partitioned into train and test models with this apparatus.
- Formation of models: Building the model with Bi-LSTM, Ri-RNN, Bi-GRU, GRU, RNN, LSTM, CNN, and LSTM + GRU with CNN.
- Users can register and sign in: You must register and log in before you can access this section.
- Prediction input will result from using this tool.
- Toward the end, the number that was anticipated will be shown.

4. IMPLEMENTATION

ALGORITHMS:

BiLSTM: In sequence or time-ordered data, a Bidirectional LSTM (BiLSTM) layer learns comprehensive connections between time steps in two directions. When needing the network to gain an advantage from the entire temporal sequence at each time step, these bidirectional connections can prove advantageous.

Bidirectional Recurrent Neural Networks (Bi-RNN): Bi-RNNs, with outputs flowing in both forward and backward directions, amalgamate two hidden layers. This architecture can extract information from both past (forward) and future (backward) time steps, making it a pivotal aspect of deep learning. Schuster and Paliwal introduced BRNNs in 1997 to enhance the handling of extensive data in arrangements. Unlike standard Recurrent Neural Networks (RNNs), BRNNs don't require sequential data in fixed windows, and they maintain a state that encodes information from potential inputs.

Bidirectional Gated Recurrent Unit (BiGRU): A model composed of two Gated Recurrent Units (GRUs) for processing sequences is referred to as a BiGRU. One GRU captures information from the initial time step, while the other GRU processes data in the opposite direction. The only distinction in this bidirectional architecture is the input and output gates.

Gated Recurrent Unit (GRU): Kyunghyun Cho and colleagues introduced Gated Recurrent Units (GRUs) in 2014 as an innovation in governing recurrent neural networks. While GRUs lack the complexity of the LSTM's cell state, they function similarly to LSTMs, employing mechanisms like the forget gate.

Recurrent Neural Networks (RNN): For sequences of data, Recurrent Neural Networks (RNNs) are a fundamental architecture. RNNs, used by systems like Apple's Siri and Google's voice search, can retain their internal state, enabling them to consider previous inputs when processing subsequent ones. This property makes them suitable for tasks requiring memory of past data, such as speech recognition.

Long Short-Term Memory (LSTM): LSTM, a prevalent deep learning architecture, is an evolved form of Recurrent Neural Networks (RNNs). Particularly useful when dealing with ordered sequences and temporal relationships, LSTMs excel at classifying, transforming, and making predictions based on sequential data. They were designed to mitigate the vanishing gradient problem that affected standard RNNs during training.

Convolutional Neural Networks (CNN): CNNs are a type of network architecture primarily used for tasks like image recognition and processing pixel data in deep learning algorithms. While various types of neural networks are employed in deep learning, CNNs are particularly effective at feature recognition in images and visual data.

5. EXPERIMENTAL RESULTS

Fig.3: Home screen

Fig.4: User login

Fig.5: Main page

Fig.6: User input

Fig.7: Prediction result

6. CONCLUSION

The objective of this study is to comprehend individuals' sentiments due to cryptocurrency-related tweets. The assessment of digital currency emotions is crucial, as it is frequently employed for predicting the valuation of available cryptocurrencies, demanding a heightened precision level in sentiment analysis. In this research, TwitterTM tweets serve as the primary data source. Tools such as TextBlob and Text2Emotion aid in embedding sentiments and emotions into the dataset. To achieve categorization, a diverse array of machine learning (ML) and deep learning models are utilized, including LSTM and GRU recurrent neural architectures, to construct an enhanced ensemble model. Furthermore, features such as Word2Vec, Bag of Words (BoW), and TF-IDF are employed to extract attributes for the ML models. Notably, ML models using BoW features exhibit superior performance compared to Word2Vec and TF-IDF. The proposed ensemble model delivers remarkable outcomes for sentiment analysis, yielding scores of 0.98 for both recall and accuracy, and 0.99 for precision. Additionally, the LSTM-GRU hybrid model outperforms all other models in generating accurate and incorrect predictions for sentiment recognition and emotion analysis tasks. However, it's important to note that when confronted with less training data, LSTM-GRU's performance diminishes due to random undersampling and dataset imbalance. This study delves into the underlying motivations behind cryptocurrency-related tweets. The ultimate aspiration is to employ the emotional insights derived from our analysis to predict the future valuation of cryptocurrencies in the market.

ISSN: 2278-4632
Vol-13, Issue-09, September 2023

REFERENCES

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

[2] S. Colianni, S. Rosales, M. Signorotti. "Algorithmic trading of cryptocurrency based on Twitter sentiment analysis." (2015). CS229 Project, Stanford University, Tech. Rep., Stanford, CA, USA, pp. 1–5.

[3] Inamdar, A. Bhagtani, S. Bhatt, P.M. Shetty. "Predicting cryptocurrency value using sentiment analysis." (2019). In Proceedings of the International Conference on Intelligent Computing and Control Systems (ICCS), May 2019, pp. 932–934.

[4] D.L.K. Chuen, L. Guo, Y. Wang. "Cryptocurrency: A new investment opportunity?." (2017). Journal of Alternative Investments, 20(3), 16–40.

[5] K. Wołk. "Advanced social media sentiment analysis for short-term cryptocurrency price prediction." (2020). Expert Systems, 37(2), e12493, Apr.

[6] C. Lamon, E. Nielsen, E. Redondo. "Cryptocurrency price prediction using news and social media sentiment." (2017). SMU Data Science Review, 1(3), 1–22.

[7] M. Hasan, E. Rundensteiner, E. Agu. "Automatic emotion detection in text streams by analyzing Twitter data." (2017). International Journal of Data Science and Analysis, 7(1), 35–51, Feb.

[8] S. Sharifirad, B. Jafarpour, S. Matwin. "How is your mood when writing sexist tweets? Detecting the emotion type and intensity of emotion using natural language processing techniques." (2019). arXiv:1902.03089.

[9] F. Calefato, F. Lanubile, N. Novielli. "EmoTxt: A toolkit for emotion recognition from text." (2017). In Proceedings of the 7th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW), Oct. 2017, pp. 79–80.

[10] S.A. Salam, R. Gupta. "Emotion detection and recognition from text using machine learning." (2018). International Journal of Computer Science and Engineering, 6(6), 341–345, Jun.