International journal of engineering science and management (IJESM).
ISSN: XXXX-XXXX.
https://doi.org/XX.XXXXX/XXXXXXXX/IJESM-XXXXXXXXX

Volume: 24, Issue 1, 43-53, October 2024.

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

Dr Kavitha Lakshmi[1], Dr. K Swathi[2]
[1]Dept of CSE, Professor, Jawaharlal Nehru Technological University Hyderabad,
Email Id: kavithalkshm067@gmail.com
[2]Dept of CSE, NRI Institute of Technology, Vijayawada, India,
Email Id: swathik1985phd@gmail.com

Received: 6 September 2024 Revised: 24 September 2024 Accepted: 2 October 2024 Published: 4 October 2024

ABSTRACT:
In recent years, the cryptocurrency industry has experienced remarkable growth. Unlike traditional currencies, cryptocurrencies operate online without the need for a central authority, relying on cryptography to ensure secure and unique transactions. Despite the cryptographic safeguards in place, the cryptocurrency industry is still in its nascent stage, leading to questions about its potential applications. To gain a comprehensive understanding of public sentiment, this study focuses on analyzing sentiments related to Bitcoin. To achieve this, the research employs sentiment analysis and emotion recognition techniques by analyzing tweets related to digital currencies, a common approach used for predicting cryptocurrency values. The study introduces an ensemble model called LSTM-GRU, which evaluates the efficacy of a mixed LSTM and GRU recurrent neural network. Several ML and DL approaches, such as term frequency-inverse document frequency, word2vec, and the bag of words features, are investigated. Also included are emotion analysis models like Text Blob and Text2Emotion. Intriguingly, the results show that among other feelings, satisfaction over cryptocurrency acceptance ranks high. Using bag of words characteristics results in better performance for ML models, according to the study. Unbelievably, the proposed LSTM-GRU ensemble model outperforms state-of-the-art methods in both sentiment analysis (0.99) and emotion recognition (0.92). conventional machine learning methods and contemporary state-of-the-art models.

KEYWORD'S – Cryptocurrency, sentiment analysis, Text2Emotion, emotion analysis, machine learning.

1. INTRODUCTION:

The digital currency industry has experienced remarkable growth since its inception. Cryptocurrency, the ability to make purchases online without the involvement of a central authority is made possible by the use of a digital currency. Even if encryption guarantees that each and every transaction is genuine and one-of-a-kind, the industry is still in its early stages, raising questions about its potential applications. To gain a comprehensive understanding of public sentiment, sentiment analysis of digital currency-related tweets is crucial, as it plays a significant role in predicting cryptocurrency values, requiring high accuracy for meaningful assessment. Twitter TM serves as the primary data source for this analysis. The study employs tools such as Text Blob and Text2Emotion for sentiment and emotion analysis. To improve classification accuracy, we create an ensemble model that combines multiple ML and Long short-term memory (LSTM) and generative adversarial networks (GRU) are two types of deep learning models. What does Word2Voice mean? To put it another way, we use the opposite of the frequency to extract. BoW-featured machine learning models demonstrate superior performance when compared to Word2Vec and TF-IDF. The proposed ensemble model excels in sentiment analysis, achieving accuracy and F1 scores of 0.98, and it attains an accuracy of 0.99 in emotion analysis, outperforming other models and methods. It is important to remember, meanwhile, that when training data is few, data imbalance and random under sampling might negatively impact the model's performance.

Numerous methodologies for sentiment analysis have been proposed; however, several unresolved issues warrant further investigation. One such challenge is the complexity of sentence structures, which poses difficulties in sentiment annotation. Achieving highly accurate annotations often necessitates the use of concise sentences. It is important to recognize that a one-size-fits-all approach cannot be universally applied across the entire corpus. Results obtained in a different domain may not always align with the goals of sentiment analysis. Additionally, the significance of specific feature extraction methods cannot be discounted entirely. In light of these considerations, Use of supervised machine learning algorithms for attitude prediction is the primary focus of this investigation and emotions concerning the Bitcoin market. Given the widespread use of Twitter TM as a platform for sharing ideas and opinions on specific topics, this study utilizes a database of tweets. These are the key contributions that this study brings to the field.

- A highly accurate ensemble model is suggested for sentiment analyses. Due to their synergistic benefits, gated recurrent units (GRUs) and long short-term memory (LSTM) are now used together.
- Both sentiment & emotion analyses are carried out. The Text2Emotion model is utilized to annotate emotions, whereas Text Blob is used to annotate sentiments data. Sentences can be positive, negative, or neutral, and emotions can be divided into joyful, sad, surprised, furious, or fearful categories.
- There are three types of feature engineering: bag of words (BoW), word2vec, and term frequency-inverse document frequency (TF-IDF)—are examined for appropriateness and performance. This study uses several well-known machine learning models, such as decision trees (DT), logistic regression (LR), support vector machines (SVM), extra tree classifiers (ETC), Gaussian Naive Bayes (GNB), and k closest neighbour (KNN). Furthermore, examined is the operation of the GRU and LSTM models.
- Utilizing the Tweepy library, the suggested method starts with gathering data from Twitter^{TM}. Tweets are removed in this way, and a Twitter^{TM} developer account is made. To get tweets, people use terms like "#BTC", "#cryptocurrency", "#cryptomarket", and "#cryptocurrency". This method gathers a total of 40,000 tweets. In 2021, the gathering of information will start in July and go through August.

2. LITERATURE REVIEW:

In the research demonstrated the effectiveness of utilizing Twitter and Google Patterns data for predicting price fluctuations in the volatile Bitcoin and Ethereum markets. Despite the dynamic nature of these cryptocurrencies, the study establishes a robust correlation between the volume of tweets and subsequent price movements. Tweets consistently convey positive sentiment regardless of price direction. By incorporating this insight into a linear predictive model that integrates 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, underscoring the growing impact of social media and search behavior on cryptocurrency trading strategies [1].

Build upon prior research that has highlighted the potential of utilizing Twitter data for forecasting the trajectories of stocks and other financial instruments. This study seeks to determine the feasibility of using 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 studied alternative currency. Through the application of 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 across different time frames and on a daily basis. This accomplishment is the result of a meticulous error analysis process that ensures data source accuracy at each stage of the model. Remarkably, the findings of this study enhance overall accuracy by 25% for individuals engaged in cryptocurrency trading [2].

The work of builds upon the insights introduced by a previous group of authors who explored the correlation between virtual entertainment and cryptocurrency valuations [3].

This study primarily focuses on Bitcoin, but its conceptual framework has the potential for application to other digital currencies in the future. By integrating sentiment scores derived from tweets and news sources with historical price and volume data, this research aims to predict cryptocurrency prices. Initial findings from the experiment suggest that individual sentiments, even when exhibiting bias toward specific categories, carry minimal significance unless they display a distinct bias [4].

The research conducted the highlights of evolving perception of Bitcoin and other digital currencies as legitimate and regulated components within financial systems, underscoring their growing recognition as significant players in capital markets. Bitcoin, in particular, has secured a prominent position in terms of market share. Consequently, this investigation elucidates the potential application of sentiment analysis in forecasting the prices of Bitcoin and various cryptocurrencies across diverse timeframes. Importantly, the study emphasizes that fluctuations in value are not solely influenced by financial institutions' control over currency but are intricately linked to individuals' perspectives and opinions, a characteristic feature of the cryptocurrency market. As such, unraveling the complex interplay between online searches and virtual entertainment becomes crucial in 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 shed light on the significant role played by psychological and sociocultural factors in shaping the dynamic costs of digital currencies [5].

The collaboration resulted in the publication of a paper attributed to the fictional figure Satoshi Nakamoto, clandestinely introducing Bitcoin to the global stage. This pivotal event marked the beginning of numerous other cryptocurrencies, spurred by its immense success. This surge can be primarily attributed to the inherent volatility of the market, which has garnered substantial interest and participation, largely driven by profit motives. On the widely-used virtual entertainment platform, Twitter, cryptocurrency enthusiasts frequently disseminate news and opinions. In this study, an exploration is conducted into the effectiveness 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 Polka dot exhibit greater stability, whereas the fluctuating prices of Bitcoin, Cardano, XRP, and Doge appear 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, with precision rates of 99.67% and 99.17% specifically for Polka dot and Ethereum [6].

3. METHODOLOGY:

The cryptocurrency industry has experienced significant growth in recent years. Cryptocurrency's function much like traditional currencies, enabling online transactions for goods and services without the necessity of centralized intermediaries [7].

Despite cryptographic techniques ensuring transaction authenticity, the industry is still in its nascent stages, giving rise to various questions regarding its potential applications. To gain a comprehensive understanding of people's viewpoints, it is essential to explore how individuals perceive Bitcoin, as illustrated in Figure 1, which displays the proposed system architecture.

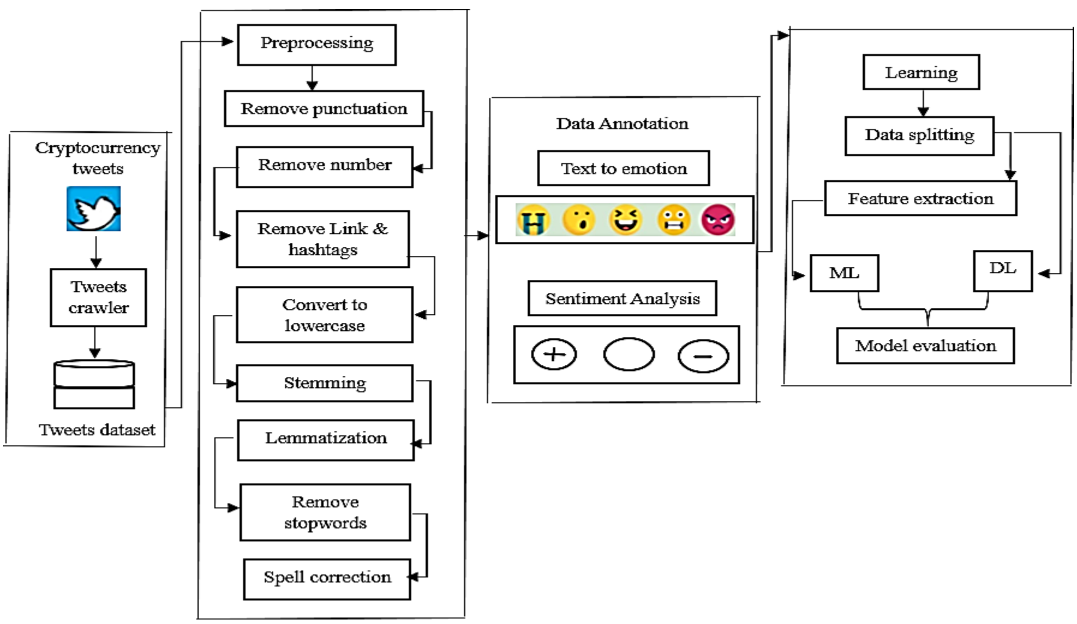

Figure.1: System architecture

Sample Text from the Dataset is shown in Table 1.

Table1: selected Text from the Collection.

User	Location	Text
Aureum Victoria	sachsen	Nice!My #trading Bot 5 just sold $NEO/$USDT with 4.73% profit on #binance! Bi-nance -20% on thr Fee:ˆæ| http://t.co/hhJIKeZD1b
Tron_age	USA	Striatus is the first strengthened 3D-printed concrete arched bridge in the world. #blockchain #dlike #cryptocurrency http://t.co/10i2IUAbG4

Tweets include characters, numbers, tags, usernames, & links that aren't necessary for training machine learning algorithms. Preprocessing methods like stemming, lemmatization, spell checking, stopword removal, etc. [8]. are used to extract this useless data from tweets are shown in Table 2.

Table 2: Text Blob and Text2Emotion Sentiment Scoring.

Text	TextBlob Text2Emotion			
	Polarity value	Sentiment	Highest value	Emotion
Nice trad bot just sold neo usdt profit on bo-nance fee	0.6	Positive	0.5	Happy
Striatu world first unreinforced print concrete archi bridge blockchain cryptocurrency	0.2	Poitive	0.5	Surprise

Text Blob rates the first tweet as extremely positive with a polarity score of 0.6, whereas the second tweet has a polarity value of 0.2, indicating that it is positive but not strongly [9].

Regarding emotion detection, based on its predictions, Text2Emotion thinks that the first tweet might get the highest score of 0.5 for the happy emotion as well as the second tweet will get the highest score of 0.5 for stress. There were a lot of neutral, positive, and negative tweets in the dataset. Figure 2a shows the ratio of those tweets, and Figure 2b shows the ratio of each mood. [10].

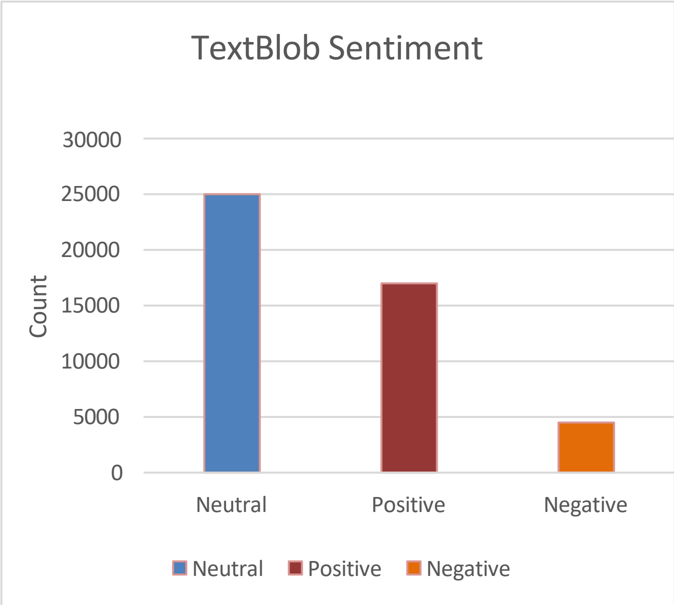

Figure 2(a): The ratio of neutral, positive, and negative tweets;

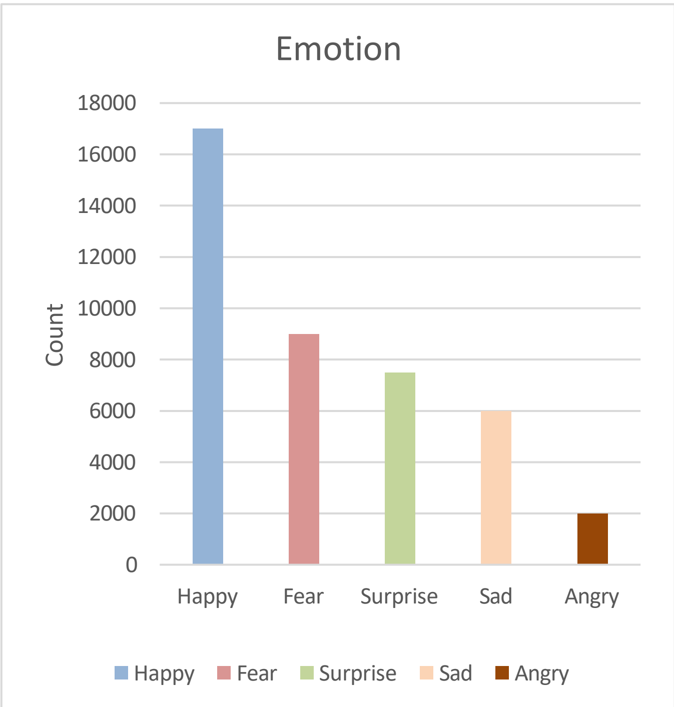

Figure 2(b): Emotion Found in the dataset.

Disadvantages:

- The analysis lacks significant robustness.
- Concerns arise due to the nascent stage of this industry.
- As a result, this study uses tweets about digital currencies for both sentiment analysis and emotion recognition, which are popular ways to guess how much different digital currencies will be worth. Additionally, an ensemble deep learning model called LSTM-GRU is created to improve the study's impact. A long short-term memory (LSTM) and a gated recurrent unit (GRU) recurrent neural network design are combined in this model. Although GRU and LSTM are stacked, GRU takes on the traits of LSTM. The suggested ensemble model is tested with different machine learning and deep learning methods using word2vec, BoW, and term frequency-inverse document frequency features. Additionally, the study assesses Text Blob 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.

4. METHODOLOGY:

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: Using Bi-LSTM, Ri-RNN, Bi-GRU, GRU, RNN, LSTM, CNN, and LSTM + GRU with CNN to create a framework.
- 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.

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. RESULTS:

Figure.3: Home screen

The above screen Fig 3 shows the home page for our proposed method and it can contain the details of our tweets.

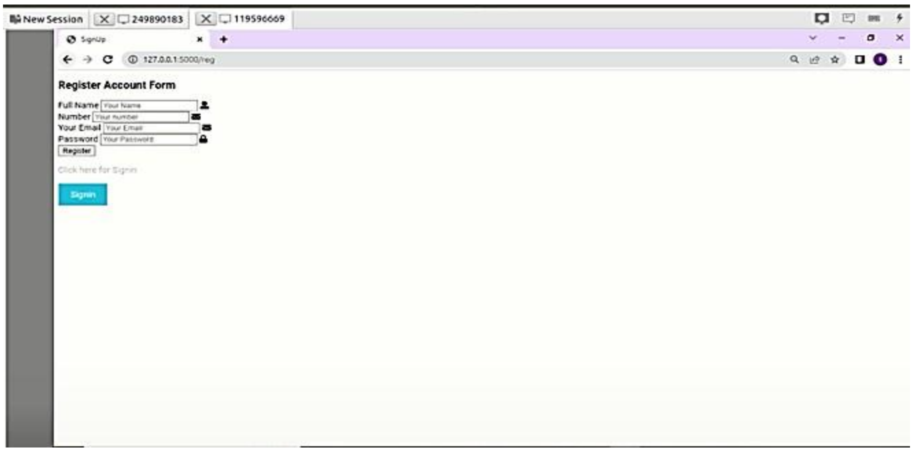

Figure.4: User login

The screen over Fig 4 shows the user login page to login into our approach and the screen Below Fig 5 shows the main page to start our proposed method.

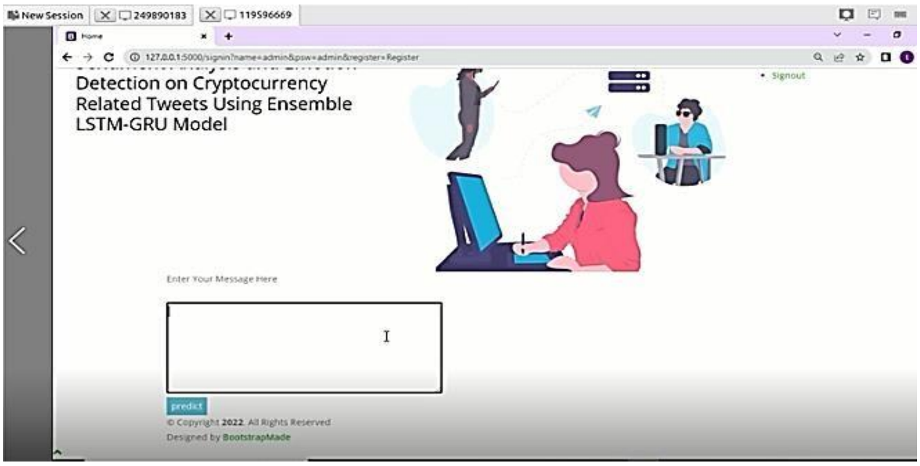

Figure 5: Main page

Figure.6: User input

The above screen Fig 6 displays the screen the user input page means the user needs to enter the basic details of the work and then the Fig 7 display the prediction results by considering user inputs.

Figure.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, Twitter TM tweets serve as the primary data source. Tools such as Text Blob and Text2Emotion aid in embedding sentiments and emotions into the dataset. To achieve categorization, A lot of different machine learning (ML) as well as deep learning methods are used, such as LSTM and GRU recurrent neural designs, 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 works better than all others in making correct as well as incorrect predictions for tasks involving emotion analysis and sentiment analysis. However, it's important to note that when confronted with less training data, LSTM-GRU's performance diminishes due to random under sampling 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.

7. 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.