

Twitter Sentiment Mining for Cryptocurrencies with Deep Learning Approach

Jimit Garala¹ and Siddhivinayak Kulkarni²

¹⁻²Department of Computer Engineering and Technology, MIT-World Peace University, Pune, India.
Email: jimit.garala@mitwpu.edu.in, siddhivinayak.kulkarni@mitwpu.edu.in

Abstract— Due to their volatility in price, the cryptocurrencies are used widely in different financial and business undertakings. Advancement in the processing of natural language and giant data analytics has caused the inception of automated systems that assess sentiments in online communities, which have now been crucial in facilitating publication of opinions by individuals in this forum. Twitter, one of the most popular social media there are, has several tweets about cryptocurrencies. With this information, we can predict the shift of bitcoin prices by utilizing deep learning (DL). Studying the sentiment of Twitter, the researchers aim to come to know more and analyze factors that are driving volatile price of cryptocurrencies. The central issue in this article is the lack of the standard model that may be referred to for the analysis of textual emotions, although textual emotions are among the reasons behind the increase and decrease in the price of cryptocurrencies. The aim of this article is to classify the sentiments for an expression into three classes: neutral, negative, and positive. The techniques that have been used include Word embedding Fast-Text model, Long-short-term memory networks (LSTMs), re-current neural networks (RNNs), gated recurrent units (GRUs), one-dimensional convolutional neural networks (CONV1D), a bi-LSTM+CONV1D, Recurrent-Convolutional Neural Networks (Three different datasets from Kaggle which contain different tweets on cryptocurrencies have been used in this research. The Long Short-Term Memory (LSTM) generated the best results in comparison with other similar techniques and outperformed other approaches in determining the sentiment conveyed within the text of the cryptocurrencies.

Keywords— Natural language processing, sentiment analysis, deep learning, algorithms, cryptocurrencies, and preprocessing.

I. INTRODUCTION

The adoption of the digital currencies has increased neck breaking pace over last few decades among consumers and corporations. It is because of it that crypto-currencies are making society's life and becoming an integral part of it. Given that the digital era has changed people's perception, there are many companies that have used the social media to promote their products or capture customer feedback. Such include social media sites such as Instagram, Google Plus, LinkedIn, Twitter, and WhatsApp. These websites are used by people where they get to express their thoughts and opinions.

Anyone can send and receive short text messages using the service provided by the popular social networking site Twitter. Even if they are strangers, users can access each other's posts on Twitter.

Twitter is a great the origin of emotional intelligence and offers real-time updates regarding cryptocurrency details because investors frequently use it to convey their emotions. Deep learning (DL) technology can be used to identify feelings in data.

It is exactly this absence of a good, accurate and precise model to assess Emotional material on social media, which inspired this paper. This is important for knowing how the value of cryptocurrencies changes, as this has a large effect on the potential for profit. It is important for those who buy these protected currencies to have a solid plan for dealing with this. Deep learning was employed as it has shown successful results in many research fields.

Readers of this post will learn to understand the site of person-environment relationships through definitions, examples, and recognition of the DL and its applications in evaluating the emotional expressions of individuals in social media sites. As this analysis is so important when considering the fluctuation prices of bitcoins, this is astonishing. According to the results from the experiment, Deep Learning method outperformed Traditional machine learning techniques. In addition, these discoveries can assist traders to choose the Deep learning model for the Sentiment analysis on Twitter and comments in social media and categories them as negative, positive or neutral.

Sentiment analysis (SA) is an approach to deducing polarity of people's subjective opinions from plain texts in the natural language. SA includes placing textual words in such categories as "positive" or "negative" or "neutral". Each emotion is given a polarity score; it is a floating-point value -1 and +1. It is possible to establish the sentiments underlying the statement. The polarity values which are greater than 0 assume the good emotions, 0 denotes neutrality, while <0 corresponds to the bad emotions. Therefore, in order to increase the accuracy of sentiment analysis we suggest applying deep learning.

Long-Short-Term Memory Networks (LSTMs), Recurrent Neural Networks (RNNs), Gated Recurrent Units (GRUs), One-Dimensional Convolutional Neural Networks (CONV1D), Bi-LSTM + CONV1D, Recurrent Neural Networks (RNNs), Gated Recurrent Units (GRUs), a Bi-LSTM + CONV1D, Recurrent-Convolutional Neural Networks (RCNN), Self-Explaining Neural Networks (SENNs), Capsule Networks (The Fast Text approach extracts semantic word vectors from the lexical words in each tweet's input word. to classify and train the tweet's emotion labels into three categories: negative, positive, and neutral. The experiment demonstrates that the LSTM model produces better predictions for sentiment classification.

The remaining seven sections of the article are as follows: Section 2 contains a review of the literature. Section 3 shows Datasets Information. Section 4 introduces our proposed model. Section 5 includes experimental results and analysis, and finally, Section 6 presents conclusions and future work.

II. LITERATURE REVIEW

Gupta et al. [1] All of this research indicate how social media information can be utilized for financial analysis whereby the door has been opened for more advanced methods of analysis that involves machine learning and natural language processing to predict the trend of the market from public opinion. Handan et al. [2] The literature study points out the fact that sentiment analysis is becoming more and more popular, particularly in the sphere of financial markets. Machine Learning (ML) and Deep Learning (DL) methods were used for sentiment classification in English, Arabic, Turkish and others.

N. Jackson et al. [3] These studies collectively demonstrate the importance of sentiment research in studying the dynamics of markets as well as the ways used to classify and analyses social media opinions especially those related to stock trading. G. Popoola et al. [4] With the use of the feature extraction technique Term Frequency – Inverse Document Frequency (TF-IDF) that was used in numerous studies — one discovers the most relevant terms in a given dataset. This approach has been used with a number of machine learning methods including support vector machines (SVM), and deep learning methods for improved classification.

S. Ao et al. [5] The research examines the increasing importance of emotion research in the financial industry with a special reference to how social media (Twitter in particular) enables people to post their opinion about stock market. M. Kulakowski et al. [6] Sentiment analysis, or SA, has grown rather common in the financial industry. This is especially true in the trading of cryptocurrencies since social media significantly affects market mood.

P. V. Mohini et al. [7] In the literature, it is possible that big data and Deep learning could revolutionize a variety of sectors like finance. One very relevant field through which these technologies could gain an insight into consumer behavior is sentiment analysis which classifies text data as neutral, negative or positive. J.D. Lee et al. [8] Identifying Social Media Sentiment and Its Correlation with Bitcoin Price Fluctuations discusses how social

media platforms, especially Twitter, are increasingly being used to forecast bitcoin price changes. Earlier research has shown that the Bitcoin price shifts has some correlation with social media sentiment.

Nair et al. [9] Literature Review In the literature review included in this article, a variety of works on sentiment analysis (SA) models applied to Bitcoin (BTC) are analyzed founded on methods for deep learning (DL) and machine learning (ML). Haritha G B et al. [10] In the paper, the increasing interest in the cryptocurrencies and cryptocurrency (Bitcoin), its volatility and the emergence of various prediction methods thereof are presented. Although there are other studies that have utilized the techniques such as deep reinforcement learning or the Bayesian regression to produce meaningful returns, the traditional approach was based on the past pricing information.

L. M. Zagi et al. [11] Literature review provides the coverage of the study on social network analysis (SNA) and its application to determine the major participants of the information-dissemination process on social networking sites and especially on Twitter. A. Mahmood et al. [12] The value of the sentiment analysis (SA) in various domains, especially politics and business has been indicated in the literature review. It explains how Twitter is a medium for public opinion where scholars can research for opinions associated with elections and political matters. As businesses utilize social media information for customer satisfaction and marketing strategy.

M. Mittal et al. [13] According to the literature review, the importance of sentiment analysis in social media comes to the fore, in particular, on sites, such as Twitter. It emphasizes the importance of the sentiment composition rules in polarity detection and mentions several methods such as rules, and case-based reasoning. Gaurav Shetty et al. [14] The study is based on the detection of spam accounts in Twitter by using machine learning classification techniques and sentiment analysis. It examines several ways how one could identify spam accounts currently and the usefulness of various classification algorithms such as Ada-Boost, Random Forests, XGBoost, and Decision Trees.

C. M. De Los Santos et al. [15] The literature review reveals the increasing trend in the availability of sentiment analysis especially in twitter and other social networking sites. The attitudes are analyzed with various approaches used by the researchers; such as Support Vector Machine (SVM) and K-Nearest Neighbours (KNN).

L. Lukas et al. [16] The literature review studies the development of the sentiment analysis and its importance in the perception of public opinion within the social media. Minimizing its attention on Naïve Bayes and other machine learning based techniques, Bayesian networks and Support Vector Machines.

D. A. Damaratih et al. [17] Literature review talks about the importance of sentiment analysis, specifically as applied to such social media outlets as Twitter. It points to increasing use of Twitter as a source for news and the public opinion, especially during the COVID-19 pandemic. A. S. Bhayyar et al. [18] It is from the literature study that a value of sentiment analysis in evaluating public views on Twitter and other social media sites is brought forward. The precision with which machine learning systems classify the emotions expressed in tweets was demonstrated in various researches.

M. Kavitha et al. [19] Movie comments are evaluated using Maximum Entropy, Naive Bayes (NB), and Support Vector Machines (SVM). SVM achieved an accuracy rate of 82.9%. When analyzing tweets during the Egyptian presidential election, TF-IDF was used to compare SVM and NB classifiers. plus, an emphasis on political tweets. H. Thakkar et al. [20] The tweets were categorized into three, by the sentiment analysis. subjective, negative and positive, with regard to subjectivity and polarity. When compared to the conventional machine learning models, the optimized BERT model was found to be much better exhibiting F1 score of 0.99.

III. DATASET

A. Cryptocurrency Tweets Dataset (10,000 Tweets)

Ten thousand cryptocurrency-related tweets were gathered from Kaggle and included in this dataset. In the context of bitcoin talks, it is designed for use in sentiment analysis, trend detection, and natural language processing (NLP) applications. The tweets provide information about sentiment and public opinion in the cryptocurrency field by mentioning a variety of well-known cryptocurrencies, including Dogecoin, Ethereum, Bitcoin, and others. Tweet id, username, timestamp, text, retweets, likes, hashtags, mentions are the features of the dataset.

B. Merged Twitter Data (107325 Records)

This dataset offers a thorough understanding of Twitter conversations about cryptocurrencies by combining tweets about different cryptocurrencies from several sources. It is appropriate for research and machine learning projects that include social media mining, sentiment analysis, market trend analysis, and natural language processing (NLP). A rich and varied collection of tweet content, including user information, timestamps, and

engagement indicators like likes and retweets, has been produced by cleaning and combining the data. Tweet id, timestamp, text, response, re-tweets, likes, hashtags, mentions, location, followers, following are the features of the dataset.

C. Bitcoin-Tweets (174,439 Tweets)

We collected Bitcoin tweets from the Kaggle website between February 10, 2021, and January 28, 2022. The dataset can be accessed at <https://www.kaggle.com/datasets/kaushiksuresh147/bitcoin-tweets>. This dataset is based on over 174,439 tweets and was created by testing a few algorithms on the entire collection. The purpose of the study was to determine public sentiment regarding Bitcoin; thus the focus was on conducting sentiment analysis (SA) of the tweets. The dataset encompasses a multitude of columns including 'user', 'userlocation', and 'userid's follower count', which is the total number of followers, in addition to the user's name, their location, and other attributes besides the tweet content and the user's rating. Within these, the research primarily discusses the text of the tweet. The dataset did not come with labels stating whether the sentiment was positive, negative, or neutral, which is why this study applied VADER classifier for polarity classification."

IV. PROPOSED METHODOLOGY

The Fig.1 The process of the sentiment analysis of twitter data is depicted above. Data preprocessing techniques of which include punctuation removal, filtering, lowercasing, tokenization, removal of stop words and stemming, are performed once data are collected from a Twitter dataset. VADER tool is then used to get sentiment score of the pre-processed text. FastText is used for feature extraction. Eighty percent of the dataset is used for training, whilst the rest, that is, twenty percent, is used for testing. Accuracy, recall, F1-score and precision are the metrics that are used to both train models and test them to estimate performance. After evaluation, the flow ends thus providing complete pipeline for sentiment analysis on Twitter based on natural language processing methods.

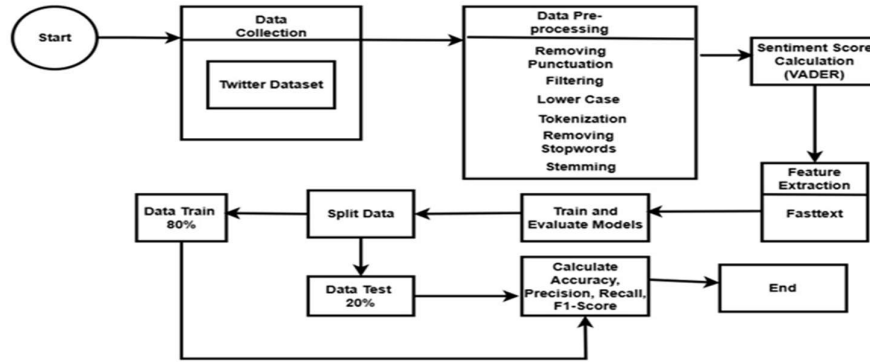


Figure 1 Twitter data sentiment analysis architecture

For producing the desired results, we used a number of Deep Learning models. The models used for the proposed system are RNN, GRU, LSTM, Conv1D, Bi-LSTM+Conv1D, RCNN, SENN, CAPSNET, and HAN detailed descriptions of these models are specified below.

A. Recurrent Neural Network (RNN)

Embedding Layer produces vector presentations of given words which are dense. Embedding matrices of pre-training embeddings (i.e. Word2Vec, GloVe) saved in the embedding_matrix. When trainable=False the semantic meaning of the embeddings is maintained. The subsequent layers can then attend to sequential dependences having made output the sequence through the initial SimpleRNN(300, re-turn_sequences=True) layer that works with input sequence. This sequence is then passed through another SimpleRNN (64) whereby the dimension is again reduced still maintaining the context information. Dense (32, activation='relu'): Reads high level sentiment features. Dense (1, activation='sigmoid'): gives either a positive or negative binary sentiment prediction.

B. Gated Recurrent Unit (GRU)

Embedding Layer forms short bases of the vector description of the utilized words. enhances the knowledge on words in utilization of the assistance of an embedding matrix, which it has been trained (e.g. the Word2Vec, the

GloVe). The `Trainable=False` will not modify the embeddings when Training it because it will ensure that they remain constant. GRU (200, `return_sequences=True`): Analyzes the tweet word by word according to the provided preceding circumstances. The outputs that are sequential are put in sequence to the next GRU layer on `return_No` transactions made in any country by the individual Caribbean countries may be considered independent of others. Dense (32, `activation='relu'`): Obtains more high level sentiment features. Dense (1, `activation='sigmoid'`): This code produces a binary classifier of sentiments is positive or negative.

C. Long Short-Term Memory (LSTM)

Embedding Layer helps create the compact representations of words in the form of vectors. The applications of initialized word embeddings (Word2Vec, GloVe, etc.) that are loaded in `embedding_matrix`. `trainable=False` offer a way to fix the embedding to be non-changeable based on the information that was acquired in advance about the meaning of the words. LSTM (300, `return_sequences=True`): Reads the input sequence and returns a sequence output to be further passed to layers of LSTM in order to be processed in information. LSTM (100): Extrudes last temporal dependencies and yields source of hidden state summarizing the input. Dense (32, `activation='relu'`): a high-level feature extractor of sentiments. Dense(1, `activation='sigmoid'`): Delivers a binary classification is positive or Negative sentiment.

D. 1D Convolutional Layer (Conv1D)

Embarks words and makes out of them thick numerical vectors. Uses pre trained embeddings (Word2Vec, GloVe) so that it can use semantic meaning. Word-embedding scan with filters (filters) or kernels. Such patterns as price increase, market down or bullish trend are identified by the filter. The amount of words that get expanded in one convolution is subject to kernel size. Helps to make the significant selections of features out of the output of the dimensionality reduction and convolutional. The most valuable information is preserved on shortening the length of the sequence by half the aid of MaxPooling1D (`pool_size=2`). The most powerful feature chosen on the sequence, GlobalMaxPooling1D() maximizes the efficiency of the model. As it is being trained, it avoids overfitting by randomly shutting off the neurons. High-level sentiment features are learnt by Dense (32, `activation='relu'`). Due to an activation functional Dense (1, `activation='sigmoid'`) a binary classification output (positive or negative sentiment) is received.

E. Bi-LSTM+Conv1D

Trains words with an embedding like word2vec and glove which represent the words as dense vectors. `Trainable=False` is used so that the embeddings cannot be trained to keep those meanings of words that were learned previously. Make use of n-gram properties of text through convolutional 1D filters. Tracks the local trend, such as the fact that the market is crashing, or the Ethereum crashing. Sentiment patterns with 5 words window pattern but a value of 5 is used rather than kernel size. Examples of the most relevant properties identified by Conv1D as it undergoes dimensionality reduction. Bi-LSTM on the other hand, can read the text left to the right direction like normal LSTM, but can also read the text backward. During training, the neurons are random turned off and on to avoid overfitting. Dense(32, `activation=relu`) attains the high level representation of sentiment. Dense(1, `activation= sigmoid`), does a binary classification of either positive or negative sentiment, the output of Dense is used.

F. Region-Based Convolutional Neural Networks (RCNN)

Encodes words to sparse numerical vectors via pre-trained word embeddings (such as Word2Vec, and GloVe). Maintains semantics meaning of terms in tweets. Just uses 1D convolution filters to tell me some valuable n-grams, e.g. "Bitcoin drop", "Ethereum rising". Takes a kernel of 5 which implies that it considers five words at a given time to identify sentiment related phrases. It makes the shape of output the same with `Padding='same'`. Retains the greatest sentiment features with a smaller dimensionality. Improves the computational efficiency by opting the best features. Sequential processes such data to trace dependencies. The preprocessed SimpleRNN(300 units, `return_sequences=True`) enables the second RNN to make use of sequential features. Another SimpleRNN (64 units) produces the last representation of the high-level sentiment. It does not overfit by disabling neurons randomly when training. The high-level features on the sentiment are learned by dense (32, `activation='relu'`). Dense (1, `activation='sigmoid'`) delivers a binary output (pos or neg) of a classification task.

G. Self-Explaining Neural Networks (SENN)

Based on pre-trained embeddings (e.g. Word2Vec, GloVe) in which words are represented as dense numeric vectors. Extracts semantic meaning of words used in the tweets associated with cryptocurrencies. Dense (64, `activation='relu'`) extracts interesting sentiment features of the text. Analyzes the essential phrases like BTC

crash, bullish momentum, or Ethereum pumping. The extractable features are given importance value by another Dense (64, activation='softmax') layer. Ensures that significant terms that concern the feelings also constitute a greater percentage in the final categorisation. The above example, the sentence cannot simply be written as "Bitcoin is recovering after a market dip"; rather the word recovering and after will be of more weightage. The Multiply() layer multiplies the features that have been extracted and their coefficients. Brings to readership attention critical features that have a potent impact on the sentiment. Constructs the 1D vector of the concatenation feature representation. A Binary sentiment classification is carried out using a Dense (1, activation='sigmoid') layer. Decides on the positive or negative sentiment of the cryptocurrency-related tweet.

H. Capsule Network (Caps Net)

Embeds (Word2Vec, GloVe, FastText) take words and encodes them into dense, numerical vectors using pre-trained embeddings. Screens semantic meaning of semantics of cryptocurrency terms like the term bullish, the term HODL, the term FOMO and so on. Based on the text, Convolutional Layer (Conv1D) extracts important characteristics. Captures word dependencies in the target language by feeding it through 128 kernel 9 size filters. Both the presence and pose (orientation) of a feature is encoded in each capsule. The Capsule Layer consists of ten sixteen dimensional capsules which capture strong representatives of sentiment. Capsule outputs are flattened into 1D vector to enable easier classification. Dense (1, activation='sigmoid') is used to carry out binary sentiment classification. Finds whether the tweet emotion of cryptocurrencies is good or negative.

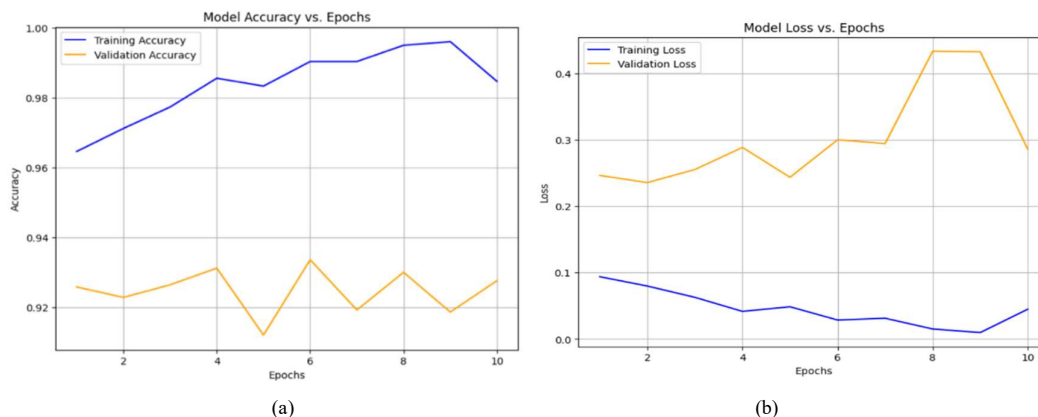
I. Hierarchical Attention Network (HAN)

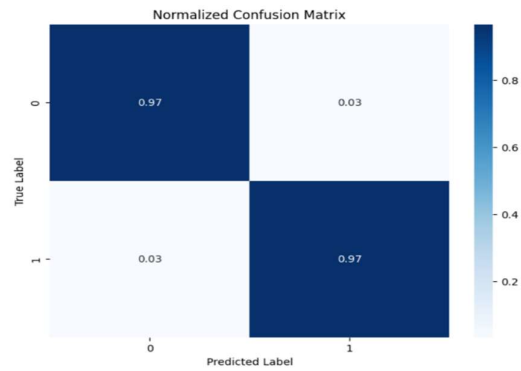
The input layer will accept a list of words as is typically expressed in a tweet discussing cryptocurrencies. Embedding Layer represents the pre-trained embeddings (Word2Vec, GloVe, FastText) in order to convert words to vectors of numeric values. A bidirectional GRU (100 units) maps each word in a forward, and back direction, and stores contextual dependencies. This assists HAN in knowing common terms like Bitcoin is pumping vs. Bitcoin is pump-ing, which may be given different meaning. Not every word is equally a sentimental word. The attention layer will give more weight to key words and neglect others. E.g. In Bitcoin is mooning! Super profits in the future!", there are more attention scores on the word mooning and gains. The sentiment is defined as positive or negative by a Dense(1, activation='sigmoid') layer. Loss Function: Binary Cross-Entropy (Binary Cross-Entropy because in this it is binary classification).

V. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, some graphs and results are shown below. The graphs and confusion matrices are being showed for the models which have performed best among all for all the three datasets. The graphs contain model accuracy vs epochs and model loss vs epoch showing how the model had behaved from time to time. Also, below there a table is included which shows the top 3 models of all 9 models which have performed better than others.

In Fig. 2., all results of LSTM are being showed on the dataset Crypto_10k_Tweets. LSTM is chosen because it is the model which have given best results for this dataset. In Fig. 2. (a), it shows a graph of model accuracy vs epochs means how the accuracy of model differs from time to time while training and validating. In Fig. 2. (b), it shows a graph of model loss vs epochs means how much loss have been made from time to time. And in Fig. 2. (c), it shows the confusion matrix of LSTM model on this dataset.

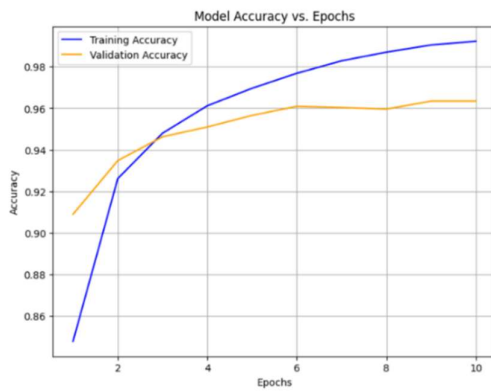




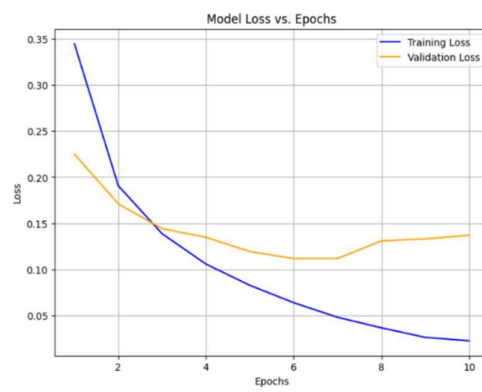
(c)

Figure 2 Results of LSTM on Dataset Crypto_10k_Tweets

In Fig. 3., all results of LSTM are being showed on the dataset Merged_Twitter_Data. LSTM is chosen because it is the model which have given best results for this dataset. In Fig. 3. (a), it shows a graph of model accuracy vs epochs means how the accuracy of model differs from time to time while training and validating. In Fig. 3. (b), it shows a graph of model loss vs epochs means how much loss have been made from time to time. And in Fig. 3. (c), it shows the confusion matrix of LSTM model on this dataset.



(a)



(b)

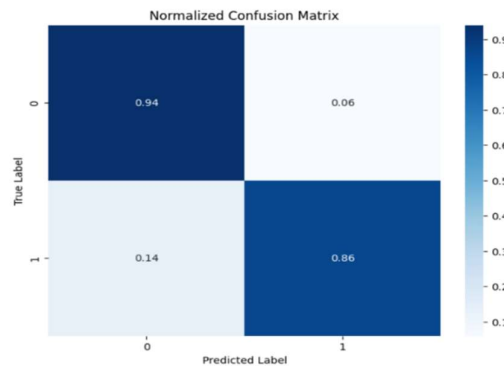


Figure 3 Results of LSTM on Dataset Merged_Twitter_Data

In Fig. 4., all results of GRU are being showed on the dataset Bitcoin_Tweets_Dataset_2. GRU is chosen because it is the model which have given best results for this dataset. In Fig. 4. (a), it shows a graph of model accuracy vs epochs means how the accuracy of model differs from time to time while training and validating. In Fig. 4. (b), it shows a graph of model loss vs epochs means how much loss have been made from time to time. And in Fig. 4. (c), it shows the confusion matrix of GRU model on this dataset.

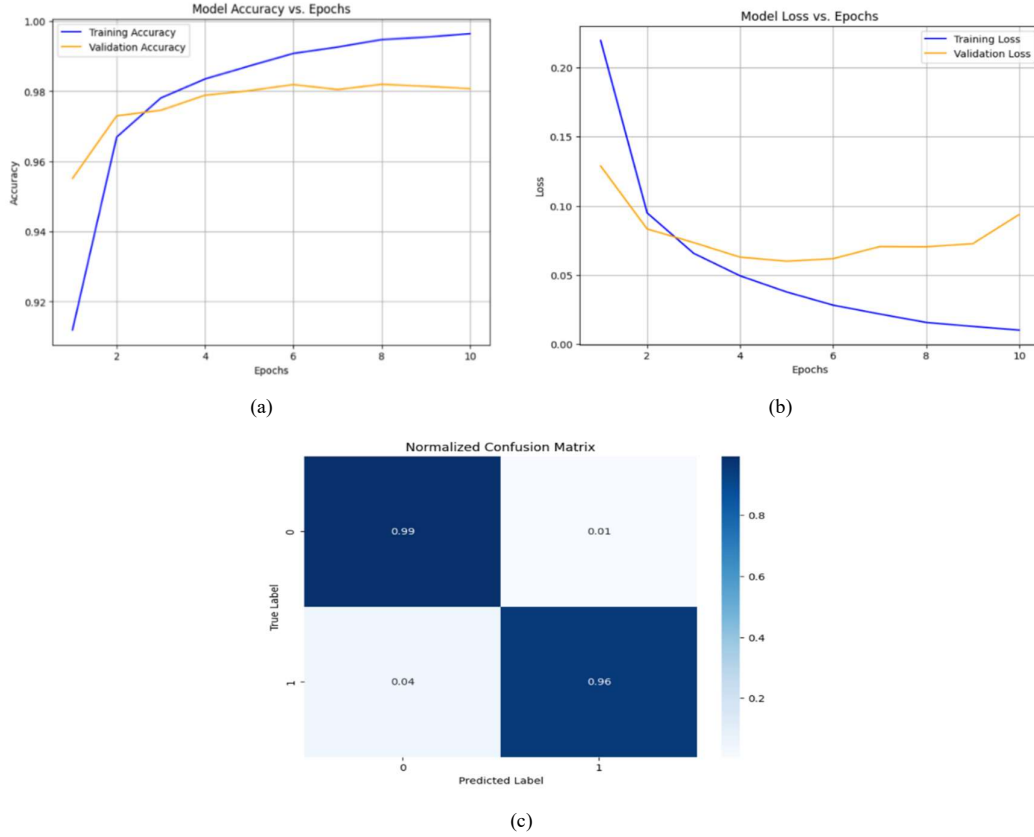


Figure 4. Results of GRU on Dataset Bitcoin_Tweets_Data Set_2

Table 1. shows the top 3 model for all 3 datasets and it can be seen as GRU, LSTM and HAN are the top 3 model in all 3 datasets showing better accuracies than other remaining models. Although, GRU and LSTM are mostly used for text processing but by this experimentation we can also see than HAN is a new model that has not been much ex-plored but it can also be used for text processing as it is also giving comparative results and can also be increased further if any further versions will come.

TABLE I: SUMMARY OF TOP 3 MODELS FOR ALL 3 DATASETS

DATASETS	MODELS	ACCURACIES	PRECISION	RECALL	F1-SCORE
Cryptocurrency Tweets	GRU	91%	83%	86%	85%
	LSTM	91%	83%	84%	83%
	HAN	91%	87%	77%	82%
Merged Twitter Dataset	GRU	96%	96%	96%	96%
	LSTM	96%	97%	96%	96%
	HAN	95%	93%	95%	94%
Bitcoin-Tweets	GRU	98%	97%	96%	97%
	LSTM	97%	96%	97%	97%
	HAN	97%	97%	94%	96%

VI. CONCLUSION AND FUTURE SCOPE

This paper provides a comprehensive assessment of numerous deep learning models for cryptocurrency sentiment analysis in tweets. Using models such as RNN, GRU, LSTM, Conv1D, Bi-LSTM+Conv1D, RCNN, SENN, CapsNet, and HAN, we aimed to classify attitudes into three distinct categories: positive, negative, and neutral. The results consistently show that LSTM and GRU models out-perform the others in terms of accuracy, precision, recall, and F1-score, with LSTM slightly exceeding the others in terms of overall robustness and generalizability. This illustrates how well recurrent architectures—LSTM—capture con-textual sentiment and

long-term dependencies in brief, casual social media texts. Additionally, the incorporation of FastText for word embedding greatly enhanced model performance and helped semantic understanding. All things considered, the suggested method offers a solid foundation for precisely identifying and analyzing market sentiment from Twitter, providing insightful information for crypto-currency researchers and investors.

Adding multilingual material and real-time Twitter streams to the dataset may enhance the model's generalizability in international markets. Accuracy and interpretability could be enhanced by using transformer-based designs such as BERT, RoBERTa, or DeBERTa, especially when used with financial text. Also, the capacity to forecast bitcoin prices may be enhanced by a hybrid model that combines sentiment analysis with price data.

REFERENCES

- [1] A. Gupta and V. K. Tayal, "Analysis of Twitter Sentiment to Predict Financial Trends," 2023 International Conference on Artificial Intelligence and Smart Communication (AISC), Greater Noida, India, 2023, pp. 1027-1031, doi: 10.1109/AISC56616.2023.10085195.
- [2] Cam, Handan, et al. "Sentiment analysis of financial Twitter posts on Twitter with the machine learning classifiers." *Heliyon* 10.1 (2024).
- [3] W. W. Bedford, N. Jackson, N. Pitts, K. Kaur and C. Tennison, "Training a Neural Network to Perform Sentiment Analysis of Stocks-Related Tweets," SoutheastCon 2024, Atlanta, GA, USA, 2024, pp. 914-921, doi: 10.1109/SoutheastCon52093.2024.10500164.
- [4] G. Popoola, K. -K. Abdullah, G. S. Fuhnwi and J. Agbaje, "Sentiment Analysis of Financial News Data using TF-IDF and Machine Learning Algorithms," 2024 IEEE 3rd International Conference on AI in Cybersecurity (ICAIC), Houston, TX, USA, 2024, pp. 1-6, doi: 10.1109/ICAIC60265.2024.10433843.
- [5] S. Ao, "Sentiment Analysis Based on Financial Tweets and Market Information," 2018 International Conference on Audio, Language and Image Processing (ICALIP), Shanghai, China, 2018, pp. 321-326, doi: 10.1109/ICALIP.2018.8455771.
- [6] M. Kulakowski and F. Frasincar, "Sentiment Classification of Cryptocurrency-Related Social Media Posts," in *IEEE Intelligent Systems*, vol. 38, no. 4, pp. 5-9, July-Aug. 2023, doi: 10.1109/MIS.2023.3283170.
- [7] A. Botta, P. V. Mohini, A. Nooka Raju and C. Balaji, "Deep Learning for Sentiment Analysis in Financial Markets," 2024 5th International Conference on Recent Trends in Computer Science and Technology (ICRTCST), Jamshedpur, India, 2024, pp. 233-238, doi: 10.1109/ICRTCST61793.2024.10578359.
- [8] O. Sattarov, H. S. Jeon, R. Oh and J. D. Lee, "Forecasting Bitcoin Price Fluctuation by Twitter Sentiment Analysis," 2020 International Conference on Information Science and Communications Technologies (ICISCT), Tashkent, Uzbekistan, 2020, pp. 1-4, doi: 10.1109/ICISCT50599.2020.9351527.
- [9] Nair, Michael & Abd-Elmegid, Laila & Marie, Mohamed. (2024). Sentiment analysis model for cryptocurrency tweets using different deep learning techniques. *Journal of Intelligent Systems*. 33. 10.1515/jisys-2023-0085.
- [10] Haritha G B and Sahana N B, "CRYPTOCURRENCY PRICE PREDICTION USING TWITTER SENTIMENT ANALYSIS" pp. 13-22, 2023. CS & IT - CSCP 2023, DOI: 10.5121/csit.2023.130302.
- [11] A. J. Oroh, Y. Bandung and L. M. Zagi, "Detection of the Key Actor of Issues Spreading Based on Social Network Analysis in Twitter Social Media," 2021 IEEE Asia Pacific Conference on Wireless and Mobile (AP-WiMob), Bandung, Indonesia, 2021, pp. 206-212, doi: 10.1109/APWiMob51111.2021.9435268.
- [12] A. Mahmood, H. U. Khan and M. Ramzan, "On Modelling for Bias Aware Sentiment Analysis and Its Impact in Twitter," in *Journal of Web Engineering*, vol. 19, no. 1, pp. 1-27, January 2020, doi: 10.13052/jwe1540-9589.1911.
- [13] G. P. M S, P. G. J, M. Mittal, P. K. H R, V. S. K and A. Jahan, "Rule Based Sentiment Analysis for Twitter(X) Data," 2024 International Conference on Advances in Modern Age Technologies for Health and Engineering Science (AMATHE), Shivamogga, India, 2024, pp. 1-5, doi: 10.1109/AMATHE61652.2024.10582181.
- [14] Gaurav Shetty, Ashwin Nair, Pradyumna Vishwanath, Ahuja Stuti, "Sentiment Analysis and Classification on Twitter Spam Account Dataset" 2020 Advanced Computing and Communication Technologies for High Performance Applications (ACCTHPA), Cochin, India, 10.1109/ACCTHPA49271.2020.9213206.
- [15] J. A. Costales, C. M. De Los Santos, J. J. E. Catulay and M. G. Albino, "Sentiment Analysis for Twitter Tweets: A Framework to Detect Sentiment Using Naïve Bayes Algorithm," 2022 4th International Conference on Computer Communication and the Internet (ICCCI), Chiba, Japan, 2022, pp. 39-44, doi: 10.1109/ICCCI55554.2022.9850257.
- [16] S. Sidharta, Y. Heryadi, L. Lukas, W. Suparta and Y. Arifin, "Sentiment Analysis of Hashtag Activism on Social Media Twitter," 2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE), Jakarta, Indonesia, 2023, pp. 73-77, doi: 10.1109/ICCoSITE57641.2023.10127833.
- [17] D. A. Damaratih, "Sentiment Analysis Of Online Lecture Opinions On Twitter Social Media Using Naive Bayes Classifier," 2021 International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE), Banyuwangi, Indonesia, 2021, pp. 24-28, doi: 10.1109/ICOMITEE53461.2021.9650135.
- [18] A. S. Bhayyar, K. P, S. Bylaiah, K. K, R. J H and P. S. R. A, "Sentiment Analysis on Karnataka election 2023 Twitter Dataset Using Machine Learning," 2023 IEEE International Conference on Blockchain and Distributed Systems Security (ICBDS), New Raipur, India, 2023, pp. 1-5, doi: 10.1109/ICBDS58040.2023.10346585.

- [19] M. Kavitha, B. B. Naib, B. Mallikarjuna, R. Kavitha and R. Srinivasan, "Sentiment Analysis using NLP and Machine Learning Techniques on Social Media Data," 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2022, pp. 112-115, doi: 10.1109/ICACITE53722.2022.9823708.
- [20] H. Thakkar, A. Patil, O. Saudagar and A. Yenikar, "Sentiment and Statistical Analysis on Custom Twitter Dataset for 2022 Russo Ukrainian Conflict," 2023 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE), Bengaluru, India, 2023, 10.1109/IITCEE57236.2023.10090995.