e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:01/January-2023 Impact Factor- 6.752 www.irjmets.com

CRYPTOCURRENCY PREDICTION USING SENTIMENT ANALYSIS

Govind Singh Bisht[*1], Tushar Monga[*2], Dr. Amita Goel[*3], Mrs. Nidhi Sengar[*4], Dr. Vasudha Bahl[*5]

[*1,2]B.Tech, Information Technology , Maharaja Agrasen Institute Of Technology, Delhi, India

[*3,4,5]Associate Professor, Department of I.T, Maharaja Agrasen Institute of Technology, Delhi, India

DOI : https://www.doi.org/10.56726/IRJMETS33096

ABSTRACT

Millions of individuals today utilize cryptocurrencies, which have a strong open-source community and payment network. The first study to predict cryptocurrency prices using news and social media emotion was published in [1]. In this paper, we apply sentiment analysis and machine learning principles to find the correlation between "public sentiment" and "market sentiment". We use twitter data to predict public mood and use the predicted mood and cryptocurrency financial news to predict the market movements.

Keywords: Sentiment Analysis, cryptocurrency, natural language processing.

I. INTRODUCTION

Recent years have seen a huge increase in the use of social media platforms, partly due to the constant and innate human need for interaction and the simplicity of creating and sustaining these online ties. Utilizing this data has been a top goal in many language processing and research projects since social network groups today offer a wealth of information and viewpoints on a variety of topics. The virtual currency is an alternate form of exchange made up of numerous decentralized cryptocurrency kinds, each with unique properties and values. The information flow surrounding virtual money is distinguished by diversity, with the nature of opinion and mood changing depending on a wide range of factors. It is in the interest of many people and companies to have the price prediction movement and direction for the Cryptocurrency. Cryptocurrency traders need to predict trends in the market to determine when to sell or buy a cryptocurrency.

In this paper, we test a hypothesis based on the premise of behavioral economics, that the emotions and moods of individuals affect their decision-making process, thus, leading to a direct correlation between "public sentiment" and "market sentiment". We perform sentiment analysis on publicly available Twitter data to find the public mood. We use these moods and cryptocurrency financial news to predict the market movements. Technically speaking, sentiment analysis is a component of natural language processing that entails training a classifier on a labeled corpus to find and establish qualities and features that will later be applied to processing fresh input data and determining the nature of it.

It is very difficult to be precise with language because it is interpretable, filled with ambiguities and presumptions, and subject to many different individual reactions, perceptions, insinuations, and idioms. But because it carries such a broad and important meaning, it must be understood. In general, a person is much better at understanding this meaning than a machine is. Contrary to the predisposition for inconsistencies, ambiguities, gaps, and contradictions, the comprehension and transmission of language can be considerably influenced by correctness when using the appropriate processing mechanisms.

This makes use of the crucial processing tool NLTK. The Natural Language Toolkit, or more simply NLTK, is a collection of Python-coded tools and applications for symbolic and statistical natural language processing (NLP) of English. It was created by Steven Bird and Edward Loper at the University of Pennsylvania's Department of Computer and Information Science. NLTK contains sample data and graphical demos. In addition to a cookbook, it comes with a book that describes the fundamental ideas behind the language processing jobs that the toolkit supports.

II. PROBLEM STATEMENT

To use sentiment analysis which can help us to predict the price of the crypto currency, with high precision of accuracy and a low error rate. The outcome of sentiment analysis will not tell the future, but it might predict the general trend and the direction to expect the prices to move.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:01/January-2023 Impact Factor- 6.752 www.irjmets.com

Several different data sources are taken into consideration as potential inputs to the model in order to forecast changes in bitcoin prices. Sentiment analysis of compiled tweets about cryptocurrencies is the first input used. Data from financial news is the second. The methods used to collect each of these data sources are described in this section.

TWEETS

Twitter is a well-known social media platform where users can communicate with one another. The tweepy library allows for data mining. We need keys, which we can apply for on Twitter's developer website. Since many experts think that public opinion influences the cryptocurrency market, they frequently use Twitter as a source.

FINANCIAL NEWS

More than 600 million people read newspapers online, compared to almost 2.5 billion who regularly read newspapers in hard format. The scope of textual analysis has expanded beyond linguistic studies due to the rising audiences of news media, social media, and blogging websites.

III. METHODOLOGY

The project will first fetch data from twitter and the news headlines about cryptocurrency. The data will then be processed and using the sentiment analysis, a polarity will be calculated. On the basis of the calculated polarity, a forecasting will be made for the cryptocurrency.

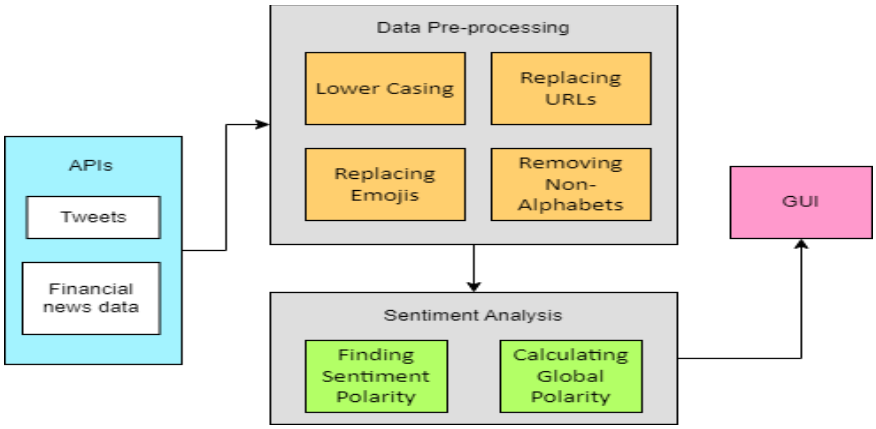

Figure 1: Algorithm Diagram

Step 1: Data

Step 1.1: Fetching Tweets:

Using tweepy library, fetching the tweets based on cryptocurrency asked by the user.

Step 1.2: Scraping News Data:

Fetching news data about cryptocurrency

```
def retrieving_tweets_polarity(symbol):
    stock_ticker_map = pd.read_csv('Yahoo-Finance-Ticker-Symbols.csv')
    stock_full_form = stock_ticker_map[stock_ticker_map['Ticker'] == symbol]
    symbol = stock_full_form['Name'].to_list()[0][0:12]

    auth = tweepy.OAuthHandler(ct.consumer_key, ct.consumer_secret)
    auth.set_access_token(ct.access_token, ct.access_token_secret)
    user = tweepy.API(auth)
    tweets = tweepy.Cursor(user.search, q=symbol, tweet_mode='extended',
                           lang='en', exclude_replies=True).items(ct.num_of_tweets)
```

Figure 2: Fetching Data

Step 2: Data Pre-processing

Step 2.1: Lower Casing: Each text is converted to lowercase.

Step 2.2: Replacing URLs: Links starting with “Http” or “https” or “www” are replaced by “URL”.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:01/January-2023 Impact Factor- 6.752 www.irjmets.com

Step 2.3: Replacing Emojis: Replace emojis by using a pre-defined dictionary containing emojis along with their meaning. (e.g.: ":)" to "EMOJIsmile").

Step 2.4: Removing Non-Alphabets: Replacing characters except Digits and Alphabets with space.

```
    blob = TextBlob(tw)
    polarity = 0  # Polarity of single individual tweet
    for sentence in blob.sentences:

        polarity += sentence.sentiment.polarity
        if polarity > 0:
            pos = pos+1
        if polarity < 0:
            neg = neg+1

        global_polarity += sentence.sentiment.polarity
    if count > 0:
        tw_list.append(tw2)

    tweet_list.append(Tweet(tw, polarity))
    count = count-1
if len(tweet_list) != 0:
    global_polarity = global_polarity / len(tweet_list)
else:
    global_polarity = global_polarity
neutral = ct.num_of_tweets-pos-neg
```

Figure 3: Polarity Calculation

Step 3: Results of Sentiment Analysis

Step 3.1: Finding Sentiment Polarity: We will quantify each tweet and news headline sentiment with a positive or negative value, called its sentiment polarity.

Step 3.2: Calculating Global Polarity: Based on all the sentiment polarities, a global sentiment polarity is calculated for the asked cryptocurrency.

Step 4: Cryptocurrency Movement Prediction

Step 4.1: Movement Prediction: Using the results of Sentiment Analysis, the movement of cryptocurrency will be forecasted and a buy or sell recommendation will be given by the application.

Subheading should be 10pt Times new Roman,

IV. RESULTS AND DISCUSSION

The data used for this study is fetched from twitter and yahoo finance ticker. We are fetching 100 tweets regarding the cryptocurrency and then the polarity of each tweet is calculated which later is used to calculate the overall polarity to indicate the movement of the price of the cryptocurrency. The data from the yahoo finance ticker is used to display the current price of the cryptocurrency.

```
##########################################################################
Today's BTC-USD Stock Data:
          Date         Open         High          Low        Close    Adj Close       Volume
731 2022-12-07 17086.486328 17106.65625 16767.955078 16812.189453 16812.189453 21507817472
##########################################################################
##########################################################################
Positive Tweets : 20 Negative Tweets : 10 Neutral Tweets : 70
##########################################################################
```

Figure 4: Bitcoin

In figure 4, current day's BTC-USD data is shown and polarity is calculated by fetching the recent 100 tweets related to the price of BTC-USD.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:01/January-2023 Impact Factor- 6.752 www.irjmets.com

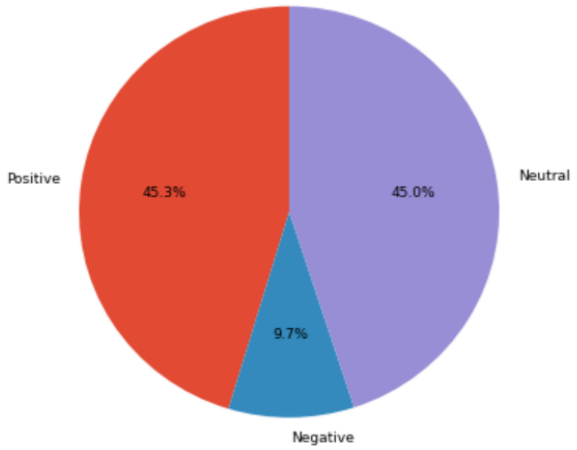

Figure 5: Sentiment Analysis for BTC-USD tweets

V. CONCLUSION

The literature evaluation covered the application of sentiment analysis to forecast cryptocurrency prices across a variety of social media platforms, including Twitter and Yahoo-Finance Ticker. Some academics utilize social media to forecast bitcoin values since numerous studies show that using sentiment analysis along with machine learning technologies increases efficiency greatly compared to using only machine learning methods.

For a better adaptation of the language processing findings to the jargon of cryptocurrencies as a future extension of the research. Additionally, the data might be added as a parameter to the LSTM model in order to more accurately predict future growth in the major currencies and investment opportunities.

VI. REFERENCES

[1] Lamon, C., Nielsen, E., Redondo, E.: Cryptocurrency Price Prediction Using News and Social Media Sentiment (2016). http://cs229.stanford.edu/proj2017/ final-reports/5237280.pdf

[2] Phillips, R.C., Gorse, D.: Predicting cryptocurrency price bubbles using social media data and epidemic modelling. In: 2017 IEEE Symposium Series on Computational Intelligence (SSCI), pp. 1–7. IEEE (2017)

[3] Mai, F., Shan, Z., Bai, Q., Wang, X., Chiang, R.H.: How does social media impact Bitcoin value? a test of the silent majority hypothesis. J. Manag. Inf. Syst. 35(1), 19–52 (2018)

[4] Vytautas, K., Niels, D., Jochen, D.W.: Using sentiment analysis to predict interday Bitcoin price movements. J. Risk Fin. 19(1), 56–75 (2018)

[5] Phillips, R.C., Gorse, D.: Mutual-excitation of cryptocurrency market returns and social media topics. In: Proceedings of the 4th International Conference on Frontiers of Educational Technologies, pp. 80–86. ACM (2018)

[6] Kushwanth, R., Sachin, A., Shambhavi, B. and Shobha, G. (2014), Sentiment Analysis of Twitter Data, International Journal of Advanced Research in Computer Engineering and Technology (IJARCET 2014), vol. 3, no. 12, pp. 4337-4342.

[7] Pak, A. and Paroubek, P. (2010), Twitter as a Corpus for Sentiment Analysis and Opinion Mining, Proceedings of the 7th Conference on International Language Resources and Evaluation (LREC'10), vol. 10, pp. 1320-13.