Proceedings of 2018 Eleventh International Conference on Contemporary Computing (IC3), 2-4 August, 2018, Noida, India

Forecasting Price of Cryptocurrencies using Tweets Sentiment Analysis

Arti Jain, Shashank Tripathi, Harsh Dhar Dwivedi, Pranav Saxena
Department of Computer Science, Jaypee Institute of Information Technology, Noida, Uttar Pradesh, India
arti.jain@jiit.ac.in, shashankt865@gmail.com, harshdwivedi124@gmail.com, psaxena130@gmail.com

***Abstract*— The problem is to find a method to predict the two-hour price of cryptocurrencies on the basis of the Social Factors, which are increasingly used for online transactions worldwide. The few previous methods proposed to predict price of cryptocurrency are inefficient because they fail to take into consideration the differences in the attributes between real currencies and cryptocurrencies. In this paper, we focus on two cryptocurrencies, namely Bitcoin and Litecoin, each with a large market size and user base, and attempt to predict their future prices using multi-linear regression model.**

Keywords—bitcoin; cryptocurrency; exchange rates; litecoin; miners; price prediction; twitter

I. INTRODUCTION

The presence of an internet access has triggered a new breed of digital currency [4] [20], known as cryptocurrency which is quite distinct from what is being used in a prevalent monetary system. An advent of cryptocurrency is based on mining and has brought about significant changes in online economic activities of users' worldwide. This digital currency system is built on computer cryptology and decentralized (peer-to-peer) network architecture [21]. Nowadays, cryptocurrencies are often used in online transactions, and their usage is on increase at drastic pace every year. Various cryptocurrencies have emerged since 2008 when Bitcoin, one of the prominent cryptocurrency, has been first introduced.

Bitcoin is considered as the first decentralized [1] [5] [12] digital currency where the transaction occurs between users directly, with no intermediary. Bitcoin is essentially a large, distributed public ledger of validated transactions. The ledger is organized as a blockchain [11] [13] of validated transactional records to track an ownership of every Bitcoin. Each transactional record contains the receiver's public key [10]. In a Bitcoin transaction, an owner validates his ownership using the private key and sends a transaction instruction encrypted with this key. Digital system then records the transaction instruction that contains the public key of the new owner, in a new block. To protect the ledger's integrity, the system labels and thus protects each block with a unique hash [10]. The hash is generated based on information on the block and an integer key. The generated hash needs to meet a hash-rate criterion. A new block that documents recent transaction is confirmed and added to the blockchain only when a valid hash is found. The system crowd sources the hash-generating process from specialized users, also known as miners [14]. By having a large number of miners investing a large amount of computational power on hash-generation, Bitcoin makes it difficult for ill-intentioned users to find a valid hash for their altered blocks[1] before other miners find a valid hash for the block that contains real transaction. Miners are incentivized to contribute computational power in generating the hash and validating blocks by receiving new Bitcoins [15]. By design, the number of Bitcoins that are generated per block starts at 50 and decreases by a half in every 210,000 blocks. A block is generated approximately every ten minutes. To control block generation speed- the hash-rate criterion and the mining difficulty adjusts every 2,016 blocks. Higher mining difficulty is associated with more computing power investment per Bitcoin. Under this mechanism, Bitcoin mining yield starts at around 7,200 blocks per day and reduces by one-half approximately in every 4 years [10].

Mathematically, the maximum supply of Bitcoins for initial four years comes out to be 21 million. Since a new block gets generated every 10 minutes. So, the number of blocks generated in one hour is 6 blocks, and the number of blocks generated in 4 years is as is seen in equation (1).

$$6 * 24 * 365 * 4 = 210{,}240 \cong 210{,}000 \qquad (1)$$

For every block mining, the miner receives a reward of 50 Bitcoins initially, but the reward reduces by half after each reward. Thus, maximum Bitcoins that can be mined from a single block are as is seen in equation (2),

$$50 + 25 + 12.5 + 6.25 + \cdots + 50/2^{\infty}$$

$$= \sum_{i=0}^{\infty} \frac{50}{2^i} = \frac{50}{1-\frac{1}{2}} = 2 * 50 = 100 \qquad (2)$$

Maximum number of Bitcoin that can be mined = (maximum blocks that are created) * (maximum of Bitcoin per block) = 210,000 * 100 = 21 million

Cryptocurrency serves as a metaphor for traditional currency system and is primarily characterized by the fluctuation in its price and the number of transactions [7]. For instance, Bitcoin begins to garner worldwide attention by the end of the year 2013 and witnesses a significant rise and

[1] https://en.bitcoin.it/wiki/difficulty

978-1-5386-6835-1/18/$31.00 ©2018 IEEE

fluctuation in its price and number of transactions. Other cryptocurrencies, namely- Ripple [22] and Litecoin [23] are also showing significant unstable fluctuations since the end of December 2013. Such unstable fluctuations serve as an opportunity for the speculation, for some users while hindering most others from using cryptocurrencies. Figure 1 shows the share percentage of various cryptocurrencies as is seen April 2018.

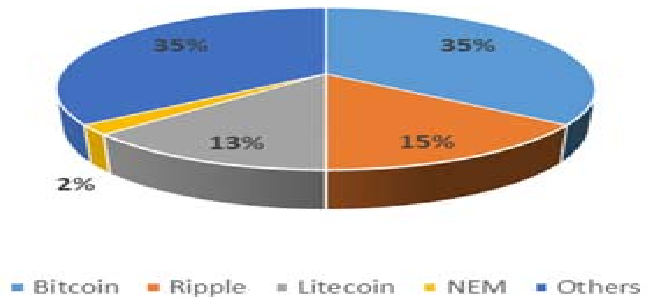

Fig. 1. Share Percentage of Various Cryptocurrencies

Microblogging sites such as Twitter [6] [16] serve as a forum where people share opinions regarding topics of common interests. Hence such sites mirror users' feelings to cryptocurrency on a daily basis. Therefore, researchers use this platform to analyze users' sentiments that are related to cryptocurrency [2]. Previous studies are limited to Bitcoin because a large amount of data that it provides eliminates the need to build a model to predict fluctuations in price and number of transactions of other cryptocurrencies [8]. They are largely traded online, where many users rely on the web-based information and make decisions about selling or buying them. Thus, there arises a need to have a prediction model, so that cryptocurrency traders can trade these commodities efficiently and effectively. In case of Litecoin, there is a limited supply of 84 million Litecoins for initial four years which is comparatively large to the Bitcoin which is 21 million for initial four years. Litecoin is very similar to Bitcoin except it has four times as many coins than Bitcoin.

This paper proposes a methodology to predict the price of cryptocurrencies, namely- Bitcoin and Litecoin using multi-linear regression model. Our methodology extracts and analyzes tweets that are tagged with the name of cryptocurrencies (Bitcoin, Litecoin), also concurrent price data which is extracted from Coindesk[2], and explores significant features that are mapped with the concurrent prices of the cryptocurrencies to build prediction curves which can predict the cryptocurrencies prices in the near future.

The paper is organized as follows. Section II gives background study related to cryptocurrency. Section III describes our proposed framework. Section IV mentions dataset details. Section V discusses experimentation and results. Section VI concludes the paper.

[2] http://coindesk.com/price

II. Background Study

Famous researcher[3] in the past has proposed an electronic transaction system for a peer-to-peer network that is based on cryptographic proof instead of trust. The network is robust as it uses proof-of-work to record the public history of transactions. A node can leave; again, join the network at will, and votes with CPU power. It can also express acceptance of valid blocks and rejection of invalid blocks. To serve as a payment medium and value storage, Bitcoin, Litecoin, and others create a decentralized authentication system [10] to deal with counterfeits and double-spending problems, whereas modern fiat monetary system and early digital payment system require central institutions to authenticate transactions and serve as repositories. Over the past several years, more than 300 academic and research articles are published on varied aspects of Bitcoin, Litecoin and other cryptocurrencies [3] [17]. Current unique active users of cryptocurrency wallets are somewhere between 5.8 million to 11.5 million. At least 2,000 people are working as full-time employees in the cryptocurrency industry. Although cryptocurrency mining facilities are geographically dispersed, a significant concentration is observed in Chinese provinces. Researchers [18] have surveyed Bitcoin key management techniques to cover the majority of deployed Bitcoin software. Several gambling sites such as Satoshi Dice[4] are turned to Bitcoin in order to protect customer privacy and to receive funds from customers who are unable to use any other payment methods. In this game, a player wins if dice roll is lesser than the player's chosen number. Researchers [19] have stated that several websites such as Overstock.com, online retailer, is receiving payments by Bitcoin since January 2014 instead of credit or debit card networks which significantly affects revenue gain, larger order size, and customer demographics. Several other merchants, such as travel- Expedia, electronics- Newegg, TigerDirect, restaurant delivery and takeout- Foodler, gift cards- Gyft are subsequently accepting cryptocurrency as a payment processor. Several models [25] such as Recurrent Neural Network (RNN), Auto-Regressive Integrated Moving Average (ARIMA), and Long Short-Term Memory (LSTM) are explored to predict Bitcoin prices that are sourced from the Bitcoin Price Index (BPI). Due to voluminous computations, Graphics Processing Unit (GPU) is considered to train models. Among them, RNN, LSTM, and ARIMA have 5.45%, 6.87%, and 53.47% respectively which indicate that RNN is the best to perform Bitcoin forecasting. Researchers [26] have compared several regression models- baseline, linear, SVM and ANN for Bitcoin price prediction, based on past information. Results indicate that the linear regression method is the most suitable for prediction, and Bitcoin price classification can be predicted with an accuracy of about 55%. Researchers have considered Multi-Layer Perceptron (MLP) based Non-linear AutoRegressive with eXogeneous inputs (NARX) for forecasting the Bitcoin price, using the past prices- opening, closing, minimum and maximum, together with the Moving Average (MA) indicator [9]. Particle Swarm Optimization

[3] Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. Accessed on 31 March 2018.

[4] https://www.journalism.co.uk/press-relaeses/10-facts-about-the-rise-and-fall-of-bitcoin-and-other-top-10-cryptocurrencies-from-2014-to-2018/s66/a716307/. Accessed on 10 Jan 2018.

(PSO) method is used to optimize the number of hidden units, input and output lags of the NARX model. Results demonstrate that the model predicts Bitcoin prices accurately while passing through validation tests. Bitcoin as virtual currency serves interest to economists having the potential to disrupt other payments and banking methods [15]. Users need to safely manage Bitcoin keys; otherwise there may be a loss of funds and a poor reputation for Bitcoin which can deter user adoption [18]. Table I shows Bitcoins recent activity as on 09 April 2018.

TABLE I. BITCOIN RECENT ACTIVITY AS ON 09 APRIL 2018

No. of Bitcoins minted	1,69,66,825
No. of reachable Bitcoin nodes	11,105
No. of transactions	15,46,81,372
No. of accounts ever used	2,15,06,448
No. of blocks to date	5,17,346
Blockchain size	149 GB
Estimated transaction volume	$486,772,289.21

Rest other currencies such as Lisk, Steem, Siacoin and others when are combined contribute 20% of the total market share. Researchers have observed that Litecoin is also increasing its popularity, along with Bitcoin, due to its low transaction fees and faster transactions [24]. Table II shows Litecoin's recent activity as on 09 April, 2018.

TABLE II. LITECOIN RECENT ACTIVITY AS ON 09 APRIL 2018

No. of Litecoins minted	6,007,783
No. of reachable Litecoin nodes	18
No. of blocks to date	1,400,388
Block chain size	15.85GB

Even though, currently, at its early stage, cryptocurrency provides valuable insights about market flow, and perhaps behavior of both buyers and sellers [15]. Users may face market risk via fluctuations in exchange rates between Bitcoin and other currencies. They may dismiss short-term spikes in price but sharp movements are really a source of concern, both for users who take currency for transactions or for storage. User's seeking trade with the volume of such currency cannot do so even quickly without affecting the market price.

III. PROPOSED FRAMEWORK

The proposed framework (Figure 2) works in two phases- Training phase and Detection phase. The training phase is a one-time activity. For carrying out training phase, we have collected Twitter data and the concurrent Bitcoin and Litecoin prices. The collected Twitter data and prices data are not in same format, the former being in JSON format and the latter being in the CSV format. So, in order to make synchronization in between these two, the Twitter data is converted into CSV format. The process of conversion of JSON file to CSV file is highlighted in Figure 3. The tweets in the data are analyzed for their sentiment polarity. The tweets having polarity above 0 are tagged as positive tweets. The tweets having polarity equals to 0 are tagged as neutral tweets. The tweets having polarity less than 0 are tagged as negative tweets. All the tagged tweets are stored and the stored data is broken into chunks containing tagged tweets which are posted in the time duration of two hours. The number of positive tweets, neutral tweets and negative tweets present in one chunk, are counted. These counted numbers are then mapped with the average of the prices that occur in corresponding two hours' time duration. The count of positive tweets, neutral tweets, and negative tweets are the features of the dataset, and the mapped average price is the label of the dataset. Model is validated with the original labels of the given dataset. If the result of validation is acceptable, then the model is ready to be used for predicting future price by analyzing real-time tweets. If not, then a new model is to be formed. The training and testing process is repeated until an acceptable model is formed. Once the acceptable model is formed, the detection phase starts. In the detection phase, real-time tweets are inputted to the model, and the model predicts the average price for the duration of two hours.

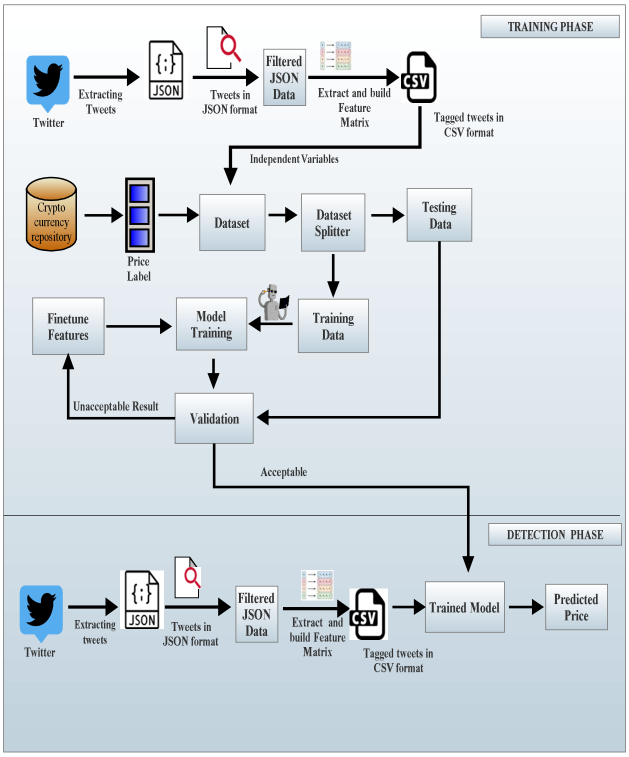

Fig. 2. System Architecture

IV. DATASET USED

This section discusses the dataset that is taken into consideration for the determination of cryptocurrencies exchange rates.

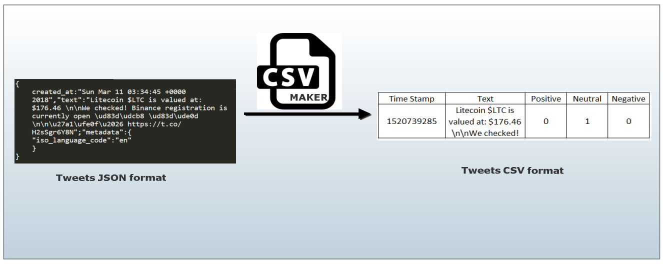

Fig. 3. JSON to CSV Converter

A. *Collection of tweets and prices*

For this research work, data is extracted as tweets with the name of cryptocurrencies- Bitcoin, Litecoin; also, concurrent price data which is extracted from Coindesk. For Bitcoin, data is collected from (30 days) March 2018 using the REST API[5] of Twitter. For Litecoin, data is collected from Feb–March 2018. At the same time, per minute price of the Bitcoin and Litecoin are also collected as is shown in Table III and IV. The collected tweets are obtained in JSON format, and the prices are obtained in .csv format. We tagged tweets as positive, neutral, and negative. For this, Textblob sentiment polarity is used for knowing tweet's sentiments. The value returned by "Textblob.sentiment.polarity" is in between -1 and 1. The tweets whose polarity value is 0 are tagged as neutral. The tweets whose polarity value is in between -1 and 0 are tagged as negative. The tweets whose polarity value is in between 0 and 1 are tagged as positive. After collection of tweets and prices, we have counted the number of tagged tweets in two hours duration and the final dataset comprises of the total count of positive, negative and neutral tagged tweets at the end of every two hours. We have considered the count of the tagged tweets for two hours because the efficiency of the model increases up to the second hour and after that, it starts decreasing, making two-hour duration an ideal for the consideration.

TABLE III. TWITTER TWEETS COLLECTION

Target Crypto – currency	**Opinion Topics**		
	Source	**Time Period**	**Tweets**
Bitcoin	Twitter	01-03- 18 to 11-03-18	13,43,252
Litecoin	Twitter	01-03- 18 to 11-03-18	1,49,735

TABLE IV. PRICE COLLECTION

Target Cryptocurrencies	**Opinion Topics**	
	Source	**Time Period**
Bitcoin	www.coindesk.com	01-03- 18 to 11-03-18
Litecoin	www.coindesk.com	01-03- 18 to 11-03-18

B. *Model Selection and Training*

1) Multiple Linear Regression: Multiple Linear Regression (MLR) [27] is a form of linear regression analysis. The model forms the relationship between two or more independent variables and a dependent variable by fitting a linear equation to a trained dataset. The independent variables can be continuous or categorical (dummy variables). The assumption of MLR is satisfied by Statsmodel to find the relation between price and sentiment of the tweet. The main assumptions of MLR are linearity, lack of multicollinearity, equality of variance, normality, independence of error. All the assumption is tested using the Statsmodel. Since our dataset where cross sectional form, no independent error are assumed by default. Equation (3) shows the equation of MLR as is used in the model

$$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + ... + b_n X_n \quad (3)$$

Here, b_0, b_1, b_2, …, b_n are the coefficients, X_1, X_2, X_3, …, X_n are the independent variables, and Y is the dependent variable.

There are several advantages of using MLR- it is used to identify the strength of the effect that the independent variables have on a dependent variable. MLR can be used to forecast effects or impacts of changes i.e. MLR helps us to understand how much the dependent variable changes, when there is a change in the independent variables. MLR predicts trends and future values with ease.

V. EXPERIMENTAL ASSESSMENT

The model predicts price regarding two cryptocurrencies, Bitcoin and Litecoin. The coefficient of determination (R2_score) and the accuracy are used to evaluate the performance of the proposed model. Table V and VI the number of positive tweet collected between 2 hour and the impact on the price of Bitcoin and Litecoin. R2_score is calculated and the value is 44% for Litecoin and 59% for Bitcoin. Statsmodel is use to verify that the assumption of MLR are met or not.

Statsmodel uses Ordinary Least Square (OLS) Method to find various results. The Statsmodel checks whether the independent variables are linear, not correlated, and all the assumption for MLR are full-filled. Statsmodel give a lot of test values and statistical information and out of those the test set which we focus verifies the assumption of multiple linear regression, and the terms are- coef, p-test, t-test, Prob (Omnibus), Durbin-Watson.

TABLE V. SAMPLE BITCOIN DATA

Time Period (hh:mm:ss)	**#of Positive tweets**	**#of Neutral tweets**	**#of Negative tweets**	**Average price ($)**
18:00:00*-20:00:00	4535	4541	1258	9141.29
20:00:00*-22:00:00	4151	4115	1159	8995.64
22:00:00*-00:00:00	3337	3302	949	8813.82
00:00:00-2:00:00	2801	3060	878	8663.99
2:00:00-4:00:00	2997	3148	888	8647.42
4:00:00-6:00:00	359	341	81	8601.66

[5] https://www.w3resource.com/API/twitter-rest-api/

TABLE VI. SAMPLE LITECOIN DATA

Time Period (hh:mm:ss)	#of Positive tweets	#of Neutral tweets	#of Negative tweets	Average price ($)
06:00:00-08:00:00	466	631	51	183.385
08:00:00-10:00:00	555	658	52	186.054
10:00:00-12:00:00	479	531	33	185.836
12:00:00-14:00:00	531	597	50	182.334
14:00:00-16:00:00	601	531	126	183.960
16:00:00-18:00:00	635	602	233	178.609
18:00:00-20:00:00	537	597	253	177.522

Figure 4 and 5 highlights the comparison of MLR to calculate the predicted price and the actual price of Litecoin and Bitcoin respectively for the time duration from 08-03-2018 6AM to 11-03-2018 1AM. The graphs show that curve of predicted price mostly matches with the curve of original price.

Fig. 4. Comparison of Multiple Linear Regression calculated predicted prices and actual prices of Litecoin for duration (08-03-2018 6AM to 11-03-2018 1AM).

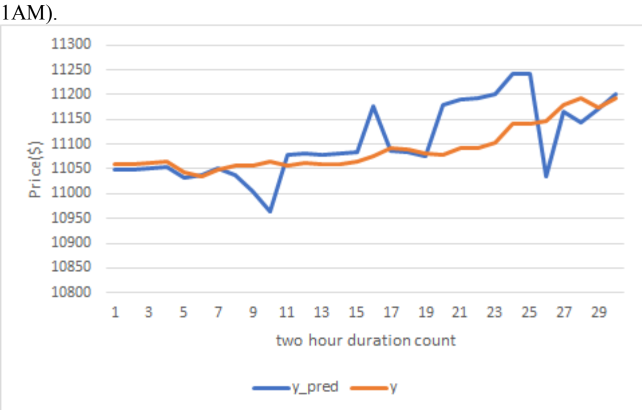

Fig. 5. Comparison of Multiple Linear Regression calculated predicted prices and actual prices of Bitcoin for duration (08-03-2018 6AM to 11-03-2018 1AM).

A. Brief about data given in Table VII and IX:

1) *Dependent Variable:* Dependent variable tells the price of Bitcoin and Litecoin in dollars.

2) *Models:* Statsmodel uses OLS model which is used for estimating the unknown parameter in multiple linear regression to find the various statistical test and R squared.

3) *R-Squared:* It tells how close the dataset is to the fitted regression line. It is the percentage of response (dependent variable) variation explained by linear regression as is seen in equation (4). Higher the R-squared the better the model fits the data.

$$\text{R} - \text{squared} = \frac{\text{Explained Variation}}{\text{Total Variation}} \quad (4)$$

4) *Adjusted R –squared:* It tells how close the dataset is to the fitted regression line, but adjusts for the number of terms in a model i.e if we add more unnecessary variables then the Adjusted R- squared will decrease and if we add valuable variables then Adjusted R- squared will Increase. Always the Adjusted R-squared is less than R-squared.

5) *Prob(Omnibus):* This test is used to check whether data is normal or not. If the Prob(Omnibus) is very small that is the mean <0.05 (Standard Value) then our data is probably not normal. This is a more precise way than graphing our dataset to determine if our data is normal or not. In both Bitcoin and Litecoin data is normal as the value of Prob(Omnibus) is greater than 0.05.

6) *Durbin-Watson:* This test is used to check the presence of correlation among the residual. The Durbin-Watson (DW) can have a value ranging from 2 to 4. The range 1.5 to 2.5 are considered normal

If the value is 2 then no correlation exists. If (0 to <2) then positive autocorrelation. If (>2 to 4) then negative autocorrelation. The residual part of Bitcoin and Litecoin has a positive correlation.

7) *Cond. No.:* This is a test for multicollinearity. If Cond. No. > 30 indicates unstable results.

8) *Coef:*It tells the coefficient of each term in MLR. The coefficient of const (b_0) is the intercept on the y-axis whereas the coefficients of positive, negative and neutral tweets tell the effect of the predictor on response variable by keeping all the other variables constant as is shown in equation (5).

$$Y=b_0+b_1 X_1+b_2 X_2+b_3 X_3 \quad (5)$$

9) *p test:* The p-value for each term tests the null hypothesis that the coefficient is equal to zero (no effect). A low p-value (< 0.05) indicates that you can reject the null hypothesis. In other words, a predictor that has a low p-value is likely to be a meaningful addition to your model because changes in the predictor's value are related to changes in the response variable.

10) *t statistic test:* The t statistic tests the hypothesis that a population regression coefficient is 0. H0: The regression coefficient is 0. So now equation (6) is modified since H0 is zero.

Since we have assumed that the hypothesis that H0 is 0 so it is the ratio of the sample regression coefficient (Coef) to its standard error (std err).

$$t - statuistic = \frac{Coeficient(Ho)}{stderr} \tag{5}$$

B. *Multiple Regression Output for Bitcoin*

Table VII and Table VIII determine the Statsmodel and residual output respectively, for Bitcoin cryptocurrency.

TABLE VII. STATSMODEL INFORMATION AND RESIDUAL OUTPUT FOR BITCOIN

Dep. Variable:	Y	**R-squared:**	0.444
Model:	OLS	**Adj. R-squared:**	0.426
Method:	Least Squares	**F-statistic:**	1.597
Date:	Fri, 23 Mar 2018	**Prob (F-statistic):**	0.286
Omnibus:	0.493	**Durbin-Watson:**	0.466
Prob(Omnibus):	0.781	**Cond. No.**	4.15e+04

TABLE VIII. STATS MODEL OUTPUT FOR BITCOIN

	Coef	**std err**	**T**	**P>|t|**
Const	8370.4	697.4	12.001	0.000
#+tweet	-0.1760	1.152	-0.153	0.084
#neutraltweet	0.5256	0.693	0.758	0.077
#negativetweet	-0.6941	1.502	-0.462	0.060

C. *Multiple Regression Output for Litecoin*

Table IX and Table X determine the Statsmodel and residual output respectively, for Litecoin cryptocurrency.

TABLE IX. STATSMODEL INFORMATION AND RESIDUAL OUTPUT FOR LITECOIN

Dep. Variable:	Y	**R-squared:**	0.591
Model:	OLS	**Adj. R-squared:**	0.549
Method:	Least Squares	**F-statistic:**	13.98
Date:	Fri, 23 Mar 2018	**Prob (F-statistic):**	8.08e-06
Omnibus:	1.414	**Durbin-Watson:**	0.604
Prob (Omnibus):	0.493	**Cond. No.**	3.49e+03

TABLE X. STATS MODEL OUTPUT FOR LITECOIN

	Coef	**std err**	**t**	**P>|t|**
Const	174.7343	4.275	40.876	0.000
#+tweet	0.0228	0.008	2.729	0.011
#neutraltweet	0.0104	0.004	2.559	0.016
#negativetweet	-0.0771	0.014	-5.514	0.000

The graphs given in Figure 6 and Figure 7 shows the residual error calculated for the Multiple Linear Regression applied to the data of Bitcoin and Litecoin respectively, both graphs highlight that the data points are uniformly distributed and are majorly found near the line, positioned at residual zero. This means that the residual error of the data points is less since the most of the dataset are distributed within the range of [-1000, 1000] for Bitcoin and [-50, 50] for Litecoin respectively.

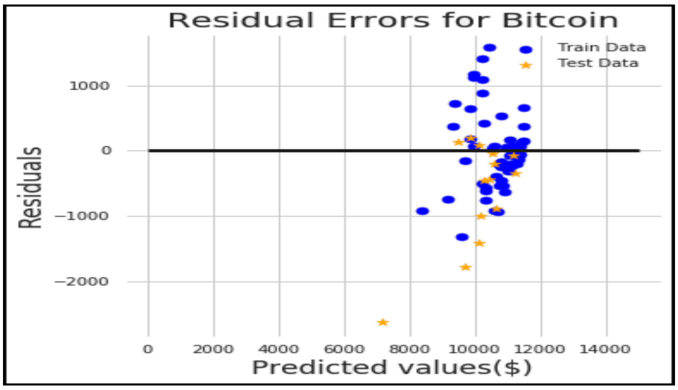

Fig. 6. Residual Error of MLR Applied on Bitcoin.

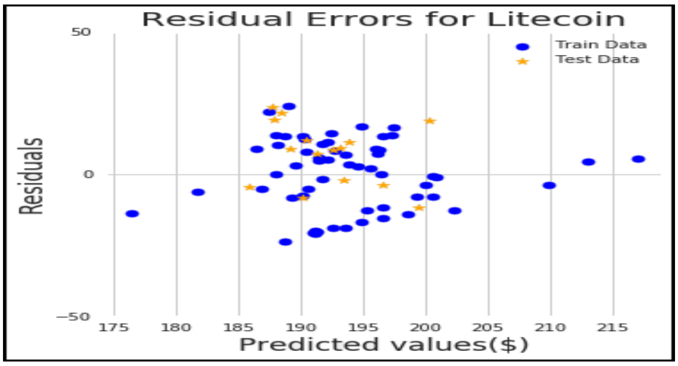

Fig. 7. Residual Error of MLR Applied on Litecoin.

Figure 8 and 9 illustrate comparison of tagged tweets for Bitcoin and Litecoin. The graphs show that number of positive tweets and neutral tweets have significant count and thus have more impact on the price of cryptocurrency whereas negative tweets are very low in comparison to the number of positive and neutral tweets posted in a time duration, so they do not have much impact on price of cryptocurrency. The red line is almost horizontal which indicates that it is better than green.

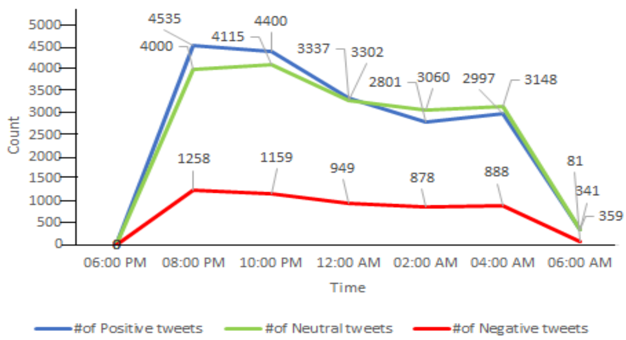

Fig. 8. Comparison of tagged tweets for Bitcoin.

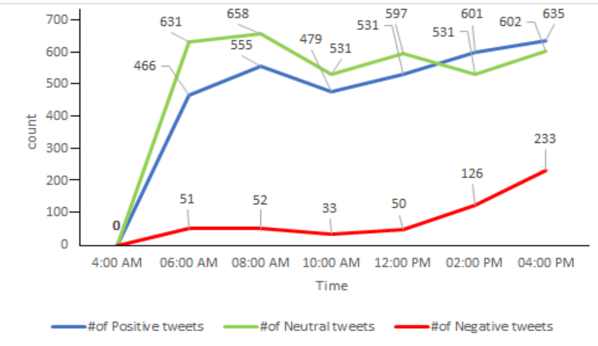

Fig. 9. Comparison of tagged tweets for Litecoin.

VI. Conclusion And Future Works

This paper predicts the price of the two cryptocurrencies- Bitcoin and Litecoin on the basis of sentiments of users' tweets which are related to these cryptocurrencies. Multiple Linear Regression predicts the price of the Bitcoin and Litecoin with R2_score of 44% and 59% respectively. From these scores, we can infer that Bitcoin's price does not get much affected by the sentiments of tweets in comparison to the Litecoin's price. Fluctuation in the price of Bitcoin has a dependency on other factors like mining cost, economic factor. Price of the cryptocurrencies for every 2 hours is predicted and a dependency of cryptocurrency price on the number of positive tweets in this duration are reflected. It is noted that the social factors play a major role in deciding the price of a cryptocurrency. In future, credibility of the user, popularity of user, user network are some other social factors that can be considered to measure the price of cryptocurrency and for increasing the accuracy of the price prediction model.

References

[1] V. Y. Naimy, and M. R. Hayek, "Modelling and predicting the Bitcoin volatility using GARCH models," *International Journal of Mathematical Modelling and Numerical Optimisation*, vol. 8, no. 3, pp. 197-215, 2018.

[2] Y. B. Kim, J. G. Kim, W. Kim, J. H. Im, I. H. Kim, S. J. Kang, and C. H. Kim, "Predicting fluctuations in cryptocurrency transactions based on user comments and replies," *PloS one*, vol. 11, no. 8, e0161197, 2016.

[3] S. Ahamad, M. Nair, and B. Varghese, "A survey on crypto currencies," In *4th International Conference on Advances in Computer Science, AETACS*, pp. 42-48, Citeseer, 2013.

[4] D. Garcia, C. J. Tessone, P. Mavrodiev, and N. Perony, "The digital traces of bubbles: feedback cycles between socio-economic signals in the Bitcoin economy," *Journal of the Royal Society Interface*, vol. 11, no. 99, 20140623, 2014.

[5] M. Matta, I. Lunesu, and M. Marchesi, "Bitcoin spread prediction using social and web search media," In *UMAP Workshops*, 2015.

[6] J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," *Journal of Computational Science*, vol. 2, no. 1, pp. 1-8, 2011.

[7] R. Böhme, N. Christin, B. Edelman, and T. Moore, "Bitcoin: economics, technology, and governance," *Journal of Economic Perspectives*, vol. 29, no. 2, pp. 213-238, 2015.

[8] R. Grinberg, "Bitcoin: An innovative alternative digital currency," *Hastings Science & Technology*, LJ, 4, 159, 2012.

[9] N. I. Indera, I. M. Yassin, A. Zabidi, and Z. I. Rizman, "Non-linear autoregressive with exogeneous input (NARX) Bitcoin price prediction model using PSO-optimized parameters and moving average technical indicators," *Journal of Fundamental and Applied Sciences*, vol. 9, no. 3S, pp. 791-808, 2017.

[10] X. Li, and C. A. Wang, "The technology and economic determinants of cryptocurrency exchange rates: The case of Bitcoin," *Decision Support Systems*, vol. 95, pp. 49-60, 2017.

[11] M. H. Miraz, and M. Ali, "Applications of blockchain technology beyond cryptocurrency," arXiv preprint arXiv:1801.03528 (2018).

[12] J. H. Ziegeldorf, R. Matzutt, M. Henze, F. Grossmann, and K. Wehrle, "Secure and anonymous decentralized Bitcoin mixing," *Future Generation Computer Systems*, 80, pp. 448-466, 2018.

[13] B. Biais, C. Bisiere, M. Bouvard, and C. Casamatta, "The blockchain folk theorem," 2018.

[14] J. Ma, J. S. Gans, and R. Tourky, "Market structure in Bitcoin mining, No. w24242, *National Bureau of Economic Research*, 2018.

[15] C. Decker, and R. Wattenhofer, "Information propagation in the Bitcoin network," In *13th International Conference on Peer-to-Peer Computing (P2P)*, pp. 1-10, IEEE, 2013.

[16] A. Jain, A. Gupta, N. Sharma, S. Joshi, and D. Yadav, "Mininig application on analyzing users' interests from Twitter," In *Proceedings of 3rd International Conference on Internet of Things and Connected Technologies (ICIoTCT)*, IEEE, March, 2018.

[17] G. Hileman, and M. Rauchs, "2017 global cryptocurrency benchmarking study," 2017.

[18] S. Eskandari, J. Clark, D. Barrera, and E. Stobert, "A first look at the usability of bitcoin key management," arXiv preprint arXiv:1802.04351, 2018.

[19] R. Sidel, "Overstock CEO sees Bitcoin sales rising more than expected," *Wall Street*, 2014.

[20] R. Grinberg, "Bitcoin: An innovative alternative digital currency," *Hastings Science & Techchnology*, LJ, 4, 159, 2012.

[21] I. Alqassem, and D. Svetinovic, "Towards reference architecture for cryptocurrencies: Bitcoin architectural analysis," In *International Conference on Internet of Things (iThings), and Green Computing and Communications (GreenCom), and Cyber, Physical and Social Computing (CPSCom)*, pp. 436-443. IEEE, 2014.

[22] F. Armknecht, G. O. Karame, A. Mandal, F. Youssef, and E. Zenner, "Ripple: overview and outlook," In International Conference on Trust and Trustworthy Computing, pp. 163-180. Springer, Cham, 2015.

[23] D. Bradbury, "Bitcoin's successors: from Litecoin to Freicoin and onwards," *The Guardian*, 25, 2013.

[24] N. Gandal, and H. Halaburda, "Competition in the cryptocurrency market," 2014.

[25] S. McNally, J. Roche, and S. Caton, "Predicting the price of Bitcoin using Machine Learning," In *26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)*, pp. 339-343, IEEE, 2018.

[26] A. Greaves, and B. Au, "Using the Bitcoin transaction graph to predict the price of Bitcoin," *No Data*, 2015.

[27] S. Asadi, S. S. Amiri, and M. Mottahedi, "On the development of multi-linear regression analysis to assess energy consumption in the early stages of building design," Energy and Buildings, 85, 246-255, 2014.