Journal of Engineering Sciences

Vol 13, Issue 03, MARCH /2022

ISSN NO:0377-9254

CRYPTOCURRENCY PRICE PREDICTION USING DEEP LEARNING

Rama Devi Ponnaganti ,Asst. Professor, M.Tech,Department of CSE, rponnaga@gitam.edu

Sreeja Killamsetty, Department of CSE, sreejakillamsetty@gmail.com

Mummidivarapu Satya Samhita, Department of CSE.samhita2661@gmail.com

G.V. Tejavardhan Reddy, Department of CSE,gvtejavardhan@gmail.com

N. Eswar Sai, Department of CSE,eswarsaijai@gmail.com

ABSTRACT: cryptocurrency is one of the main phenomena in recent times together with other. Due to the redefinition of the money term and its price fluctuations. Moreover, scientists are increasingly recognizing Twitter's predictive power for a wide range of events, and particularly for financial markets. This article examines to what degree cryptocurrency returns can be estimated using public opinion on Twitter. Using a sentiment analyzer on cryptocurrency -related tweets and financial data, the Twitter sentiment was found to have predictive power for cryptocurrency results. Once again, our findings confirm the presence of a correlation between them. We observed 73% accuracy when making predictions based on cryptocurrency-related tweet sentiment and historical cryptocurrency price.

Keywords- *Cryptocurrency, price prediction, Deep learning.*

1. INTRODUCTION

The Bitcoin is a highly cryptic and virtual currency used by many investors throughout the world. Satoshi Nakamoto was the inventor of Bitcoins in 2009[1]. Thus, Bitcoin is a blockchain based currency that encompasses a public records of all the transactions performed under monitoring. Many researchers have worked in this field to predict and analyze the trends and patterns of the Bitcoin prices. Initially, with very less data and limited scope in algorithms and tools the accurate representation and factual prediction of values was difficult but with the advancements in technology and higher scopes in domains like machine learning and deep learning researchers have been keen on developing models that can provide an insight to the estimation of monetary values. A literature survey that consisted of some prominent work in the respective domain provided quite notable results. There is high volatility in the market, and this provides an opportunity in terms of prediction as mentioned in [1]. The writer of [2] proposes a solution to double spending problem using peer to peer distributed server. The authors of [3] assert that Bitcoin is the world's most valuable cryptocurrency and is traded on over 40 exchanges worldwide accepting over 30 different currencies The study in the paper [4] reveals that the authors have executed the result of Bayesian Neural Networks (BNNs) by analyzing the time series of Bitcoin process. The paper [5] proposes that the model for prediction of time series data based on the concept of sliding window using Artificial Neural Network (ANN) technique which is Radial Basis Function Network (RBFN). It depicts certain limitations such introduction of hybrid or ensemble techniques with new features. The paper [6] attempts to identify and understand daily trends in Bitcoin market by gathering optimal features surrounding Bitcoin prices and plot a graph using normalization. The authors of the work [7], use rolling window Long Short- Term Memory (LSTM) model to predict Bitcoin price

by selecting the input features such as macroeconomics, global currency ratios, and block chain information. In the work [8], the authors explore Neural Network ensemble approach called Genetic Algorithm based Selective Neural Network Ensemble by using back tracking strategy where the author suggested that some input information might be missing as more processing of the data was required.

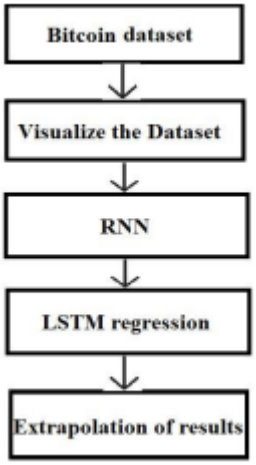

Fig.1: Deep learning model

Neural Network (RNN) model using Long Short-Term Memory (LSTM) regression algorithm on the acquired Cryptocurrency dataset for predicting the prices of cryptocurrency (Bitcoin) by analyzing the dataset and applying deep learning algorithms. Thus, for this research the dataset used consists of various parameters of Bitcoins data values . The goal of this research is to design a model that will consistently be able to predict the price of Bitcoin. Predicting the exact price is very hard. Therefore, we simplify the problem; we only try to predict whether the price will increase, decrease or stay the same within certain thresholds. The prediction analysis would be carried out based on the resultant values from the given algorithms. The objectives of proposed model is to create model that leads to the Bitcoin price prediction accuracy by incorporating RNN elements.

2. LITERATURE REVIEW

2.1 Impact of Social Media on Cryptocurrency Trading with Deep Learning:

Due to the revolutionary increase in the amount of available data, the rise of high performance GPUs and the novel scientific results of neural networks, deep learning has received high attention among machine learning scientists. The numerous layers of deep architectures are able to extract and learn different abstractions of the input data and model them more efficiently than previous machine learning methods in many application fields. Deep learning is one of the main technology of language understanding, natural language processing and time series analysis. One of the promising theoretical question is the feasibility of deep learning for weak coherent signals, where the feature extraction and therefore the analysis and modeling is difficult with classical methods. (For example market demand or price movement of a financial asset and the corresponding news sentiment, or signals of an IoT sensor network are weak coherent signals. Digital currencies with an underlying peer-to-peer, decentralized payment system have seen a large growth recently. Most of the coins reward the so called miners, who are offering their computational capacity to enable to run and secure the decentralized services. This earning has a real values on open market, therefore the growing popularity of cryptocurrencies has an influence on the growing of the price tag of this rewards. The aftermath of this considerable presence on the market is the appearance of the exchange services for these currencies. These exchanges execute a huge amount of tradings with an opened API.The market analysis is difficult, because the volatility is high, the signal is noisy, and the lack of classical features cause the need of rethinking the well-tried strategies. Furthermore, cryptocurrency markets are not regulated. Due to the large amount of public data from the blockchain and cryptocurrency exchanges, data driven machine learning approaches may explore pricing anomalies. Besides the trading data additional sources of information (like political, or financial news) may enhance the predictive performance of a

machine learning model. Twitter like short texts are typically within the modeling capacity of such analytical systems. As social media, including Twitter, become one of the primary source of news (eg. political, technological, financial news), it has a great influence on masses. Thus, it is critical to be able to model, classify, cluster these texts and eg. to extract sentiment and detect fake news.

2.2 Virtual currency, tangible return: Portfolio diversification with Bitcoin:

Bitcoin is a major virtual currency. Using weekly data over the 2010-2013 period, we analyze a Bitcoin investment from the standpoint of a U.S. investor with a diversified portfolio including both traditional assets (worldwide stocks, bonds, hard currencies) and alternative investments (commodities, hedge funds, real estate). Bitcoin investment has highly distinctive features, including exceptionally high average return and volatility. Its correlation with other assets is remarkably low. Spanning tests confirm that Bitcoin investment offers significant diversification benefits. We show that the inclusion of even a small proportion of Bitcoins, say 3%, may dramatically improve the risk-return trade-off of well-diversified portfolios.

2.3 Bitcoin: A Peer-to-Peer Electronic Cash System:

A purely peer-to-peer version of electronic cash would allow online payments to be sent directly from one party to another without going through a financial institution. Digital signatures provide part of the solution, but the main benefits are lost if a trusted third party is still required to prevent double-spending. We propose a solution to the double-spending problem using a peer-to-peer network. The network timestamps transactions by hashing them into an ongoing chain of hash-based proof-of-work, forming a record that cannot be changed without redoing the proof-of-work. The longest chain not only serves as proof of the sequence of events witnessed, but proof that it came from the largest pool of CPU power. As long as a majority of CPU power is controlled by nodes that are not cooperating to attack the network, they'll generate the longest chain and outpace attackers. The network itself requires minimal structure. Messages are broadcast on a best effort basis, and nodes can leave and rejoin the network at will, accepting the longest proof-of-work chain as proof of what happened while they were gone.

2.4 Predicting the Price of Bitcoin Using Machine Learning:

The goal of this paper is to ascertain with what accuracy the direction of Bitcoin price in USD can be predicted. The price data is sourced from the Bitcoin Price Index. The task is achieved with varying degrees of success through the implementation of a Bayesian optimised recurrent neural network (RNN) and a Long Short Term Memory (LSTM) network. The LSTM achieves the highest classification accuracy of 52% and a RMSE of 8%. The popular ARIMA model for time series forecasting is implemented as a comparison to the deep learning models. As expected, the non-linear deep learning methods outperform the ARIMA forecast which performs poorly. Finally, both deep learning models are benchmarked on both a GPU and a CPU with the training time on the GPU outperforming the CPU implementation by 67.7%.

2.5 An Empirical Study on Modeling and Prediction of Bitcoin Prices With Bayesian Neural Networks Based on Blockchain Information:

Bitcoin has recently attracted considerable attention in the fields of economics, cryptography, and computer science due to its inherent nature of combining encryption technology and monetary units. This paper reveals the effect of Bayesian neural networks (BNNs) by analyzing the time series of Bitcoin process. We also select the most relevant features from Blockchain information that is deeply involved in Bitcoin's supply and demand and use them to train models to improve the predictive performance of the latest Bitcoin pricing process. We conduct the empirical study that compares the Bayesian neural network with other linear and non-linear benchmark models on modeling and predicting the Bitcoin process.

Our empirical studies show that BNN performs well in predicting Bitcoin price time series and explaining the high volatility of the recent Bitcoin price.

2.6 Time Series Data Prediction Using Sliding Window Based RBF Neural Network:

Time series data are data which are taken in a particular time interval, and may vary drastically during the period of observation and hence it becomes highly nonlinear. Stock index data are time series data observed daily, weekly or even monthly. Prediction of these types of data is very challenging. For accurate prediction of time series data different intelligent techniques are being used by the researchers, on the other hand, prediction of next day close price on the basis of current day price is not appropriate, instead an average of a particular range of stock data known as window may be suitable for prediction of highly nonlinear stock data. This paper explores an Artificial Neural Network (ANN) technique: Radial Basis Function Network (RBFN) for data prediction using the concept of sliding window, which produces data for current day using historical data of earlier days calculated by Weighted Moving Average (WMA). Experiments were carried out using 10-fold cross validation technique with MATLAB written code for BSE30 Index data. Result produced through RBFN were measured through MAPE, MSE, MAD and RMSE and found satisfactory.

2.7 Bitcoin Price Prediction using Machine Learning:

The goal of this paper is to ascertain with what accuracy the direction of Bitcoin price in USD can be predicted. The price data is sourced from the Bitcoin Price Index. The task is achieved with varying degrees of success through the implementation of a Bayesian optimised recurrent neural network (RNN) and a Long Short Term Memory (LSTM) network. The LSTM achieves the highest classification accuracy of 52% and a RMSE of 8%. The popular ARIMA model for time series forecasting is implemented as a comparison to the deep learning models. As expected, the non-linear deep learning methods outperform the ARIMA forecast which performs poorly. Finally, both deep learning models are benchmarked on both a GPU and a CPU with the training time on the GPU outperforming the CPU implementation by 67.7%.

2.8: Predicting Bitcoin price using Rolling Window LSTM model:

Bitcoin has recently received considerable interest in the fields of economics and cryptography as its system is based on the latest blockchain technologies. In this study, we used the rolling window long short-term memory (LSTM) model to predict bitcoin price. We selected the input features of the model, which are relevant to macroeconomics, global currency ratios, and blockchain information. We used blockchain information because it has a direct relationship with the supply and demand of bitcoin in the currency market. In order to predict the bitcoin prices, we conducted an empirical study to compare the proposed model with various machine learning models. Our results indicate that the proposed model outperforms other models, and that the rolling window LSTM can accurately predict recent bitcoin prices. We also suggested a new approach to maintain the stable price of the cryprocurrency by adjusting the monetary volume based on the estimated inflation rate.

2.9: Bitcoin price prediction using ensembles of neural networks:

This paper explores the relationship between the features of Bitcoin and the next day change in the price of Bitcoin using an Artificial Neural Network ensemble approach called Genetic Algorithm based Selective Neural Network Ensemble, constructed using Multi-Layered Perceptron as the base model for each of the neural network in the ensemble. To better understand the practicality and its effectiveness in real-world application, the ensemble was used to predict the next day direction of the price of Bitcoin given a set of approximately 200 features of the cryptocurrency over a span of 2 years. Over a span of 50 days, a trading strategy based on the ensemble was compared against a

"previous day trend following" trading strategy through back-testing. The former trading strategy generated almost 85% returns, outperforming the "previous day trend following" trading strategy which produced an approximate 38% returns and a trading strategy that follows the single, best MLP model in the ensemble that generated approximately 53% in returns.

2.10: Automated Bitcoin Trading via Machine Learning Algorithms:

In this project, we attempt to apply machine-learning algorithms to predict Bitcoin price. For the first phase of our investigation, we aimed to understand and better identify daily trends in the Bitcoin market while gaining insight into optimal features surrounding Bitcoin price. Our data set consists of over 25 features relating to the Bitcoin price and payment network over the course of five years, recorded daily. Using this information we were able to predict the sign of the daily price change with an accuracy of 98.7%. For the second phase of our investigation, we focused on the Bitcoin price data alone and leveraged data at 10-minute and 10-second interval timepoints, as we saw an opportunity to evaluate price predictions at varying levels of granularity and noisiness. By predicting the sign of the future change in price, we are modeling the price prediction problem as a binomial classification task, experimenting with a custom algorithm that leverages both random forests and generalized linear models. These results had 50-55% accuracy in predicting the sign of future price change using 10 minute time intervals.

3. IMPLEMENTATION

In this paper, we present a solution to predicting crypto currency price changes. To accomplish this, methods utilizing sentiment analysis of tweets are reviewed. This is involved utilizing Twitter's API and a Python library called "Tweepy" to collect and store tweets which mentioned Bitcoin or Ethereum. The tweets were then analyzed to create a sentiment score by day and compared to the price changes to that day to determine if a relationship between Twitter sentiment and crypto currency price changes could be determined.

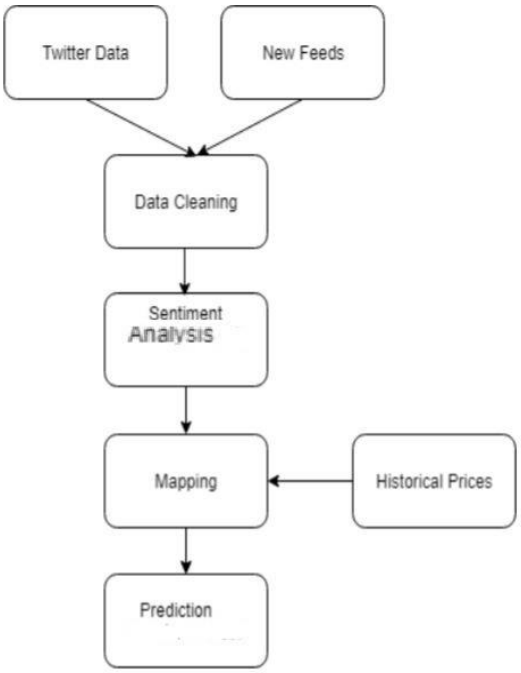

Fig.2: System architecture

MODULES:

Data:

To solve the problem of predicting cryptocurrency price changes several different data sources are considered as possible inputs to the model. The first input considered is sentiment analysis of collected tweets about Bitcoin or Ethereum. The second was Google Trends data, and the third was tweet volume. This section details how each of these data sources were gathered, cleaned, and adjusted when necessary.

Collecting Tweets from Twitter's API:

The first step in collecting the desired tweets was to find the hashtag for the cryptocurrencies. For this we utilize Tweepy - an open-source Python library for accessing the Twitter API, to collect Twitter data. Tweepy allows for filtering based on hashtags or words. There are multiple ways in which the cryptocurrencies of interest may be referred to in tweets. The most direct

way is by using a hastag(”#”) followed by ”bitcoin” or ”ethereum”. Other likely possibilities are using a hashtag and either currencies abbreviation (”#btc” for Bitcoin, and ”#eth” or Ethereum). Early collections of tweets using only the ”#bitcoin” and ”#ethereum” hashtags quickly provided a large data set. As these hashtags had little ambiguity they were selected as the only ones we would use to collect tweets.

Cleaning the Tweets for Analysis:

With the tweets collected further processing is required. Tweets come in a format with characters which do not provide ”information” for a sentiment analysis. VADER (Valence Aware Dictionary for sEntiment Reasoning) sentiment analysis was used to analyze the collected tweets. VADER analysis provides several benefits including the fact that it not only classifies text as positive, negative, or neutral, but also measures the intensity, or polarity, of words used. For our purposes, we also benefit from the fact that the words and scores used in VADER are specifically tuned to microblog and social media contexts. To eliminate noise from the analysis we first clean the collected tweets. Tweets contain a large amount of noise, such as hashtags, URLs, and emotions. These characters make Twitter sentiment analysis a challenging assignment. Preprocessing of the data is a very important step as it decides the efficiency of the other steps down in line for sentiment analysis.

Not all tweets are posted by humans. A substantial number of users and tweets are actually from bots. Twitter has estimated that as many as 23 million of its active users are actually bots. If the bots were sending tweets which contained positive or negative sentiment about the cryptocurrencies then those tweets may still have an influence on people’s demand to own cryptocurrencies, and therefore impact the prices. However, many of the tweets do not contain any sentiment and instead provide only facts or are serving the function of advertising. Beyond bots there is concern of the subject matter. Conversations about cryptocurrencies can be very neutral in nature. What the current USD price of a single bitcoin is a fact and does not carry any sentiment. Therefore, sentiment analysis of the tweets may provide limited information to the model. After pre-processing the collected tweets, algorithm results showed the information gained from the tweets through sentiment analysis is still of limited value. Overall tweets were generic, generated by bots, or advertisements.

4. IMPLEMENTATION

The findings of our analyses show that sentiment analysis is less effective for cryptocurrency price changes in an environment in which prices are falling. This is because tweets about cryptocurrencies tend to be objective in nature (not having a clear sentiment) or positive regardless of price changes. While their value has exploded in recent years, they still pale in use compared to traditional fiat currencies. In addition, cryptocurrencies are actually a part of a larger technology (the blockchain). As such, Twitter activity about them can be driven by people with a special interest in the currency or the technology rather than just a store of value, as a traditional stock may be viewed. Google Trends data and tweet volume better reflects the overall interest in owning cryptocurrencies as they increase and decrease with prices.

Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture[1] used in the field of deep learning. Unlike standard feedforward neural networks, LSTM has feedback connections. It can not only process single data points (such as images), but also entire sequences of data (such as speech or video).

For example, LSTM is applicable to tasks such as unsegmented, connected handwriting recognition,[2] speech recognition[3][4] and anomaly detection in network traffic or IDSs (intrusion detection systems). A common LSTM unit is composed of a cell, an input gate, an output gate and a forget gate. The cell remembers values over arbitrary time intervals and the three gates regulate the flow of information into and out of the cell. LSTM networks are well-suited to classifying, processing and making predictions based on

time series data, since there can be lags of unknown duration between important events in a time series. LSTMs were developed to deal with the vanishing gradient problem that can be encountered when training traditional RNNs. Relative insensitivity to gap length is an advantage of LSTM over RNNs, hidden Markov models and other sequence learning methods in numerous applications.

Autoregressive Integrated Moving Average Model:

An ARIMA model is a class of statistical models for analyzing and forecasting time series data. It explicitly caters to a suite of standard structures in time series data, and as such provides a simple yet powerful method for making skillful time series forecasts. ARIMA is an acronym that stands for AutoRegressive Integrated Moving Average. It is a generalization of the simpler AutoRegressive Moving Average and adds the notion of integration.

5. EXPERIMENTAL RESULTS

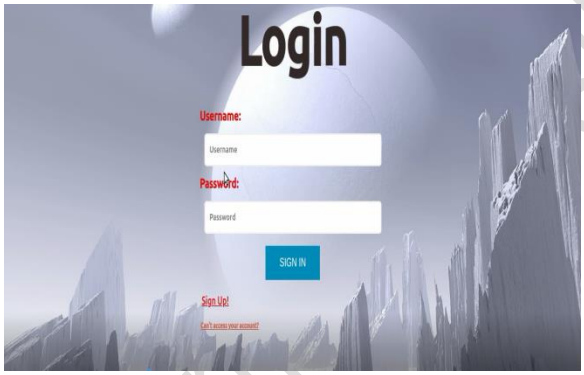

Fig.3: Login screen

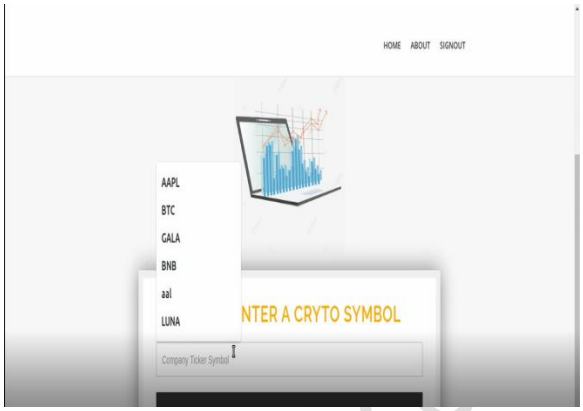

Fig.4: Home screen

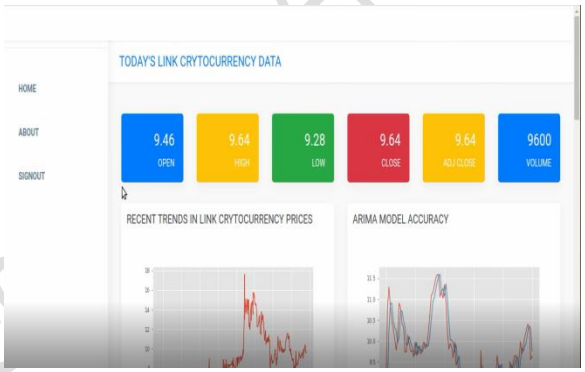

Fig.5: Crypto currency calculation

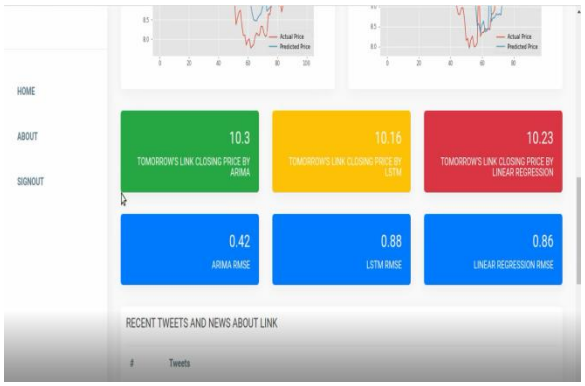

Fig.6: Crypto currency calculation

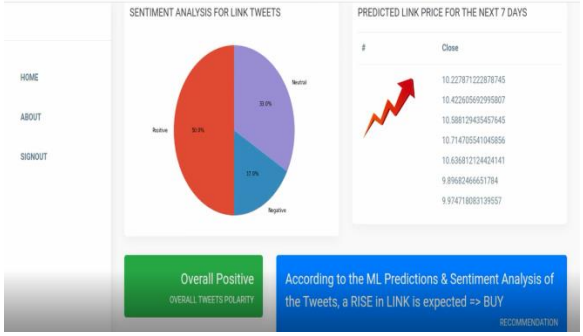

Fig.7: Prediction result

6. CONCLUSION

Analysis of tweet sentiment is an important field of price forecasting studies. Because of the large amount of news updates per minute about crypto currency, using Twitter in the sentiment analysis for crypto currency is becoming a significant step for most researchers. Text mining and classification techniques on Twitter data are therefore needed which can predict the best sentiment. This paper's main contribution is to find a partial correlation between the price fluctuation of cryptocurrency and the fluctuation of the sentiment classes using a machine learning algorithm. Observed that there is a strong correlation between the Bitcoin percentage shift and Twitter sentiment. As future research, developing a special cryptocurrency sentiment lexicon could improve the connection between the sentiment analysis and the cryptocurrency fluctuation, considering other features such as hashtags, Twitter users, number of tweets, and emoticons.

7. FUTURE SCOPE

In the future, we will work on cryptocurrencies with more than one inter-dependency using the proposed model. Moreover, we will add sentimental factors, such as Twitter and Facebook posts and messages, to the proposed model to improve the accuracy of the prediction results. Traditional commodities such as gold and oil prices can also be considered to enhance the prediction results.

REFERENCES

[1] M. Briaere, K. Oosterlinck, and A. Szafarz,Virtual currency, tangible return: Portfolio diversification with Bitcoins, Tangible Return: Portfolio Diversification with Bitcoins , 2013.

[2] S. Nakamoto, ”Bitcoin: A Peer-to-Peer Electronic Cash System”, Available at: https://Bitcoin.org/Bitcoin.Accessed on 2008.

[3] McNally, Sean & Roche, Jason & Caton, Simon. (2018). Predicting the Price of Bitcoin Using Machine Learning. 339343. 10.1109/PDP2018.2018.00060.

[4] H. Jang and J. Lee, "An Empirical Study on Modeling and Prediction of Bitcoin Prices With Bayesian Neural Networks Based on Blockchain Information," in IEEE Access, vol. 6, pp. 5427-5437, 2018

[5] Hota HS, Handa R & Shrivas AK, “Time Series Data Prediction Using Sliding Window Based RBF Neural Network”, International Journal of Computational Intelligence Research, Vol.13, No.5, (2017), pp.1145-1156

[6] Siddhi Velankar,Sakshi,Valecha,Shreya Maji, aBitcoin Price Prediction using Machine Learninga,20thInternational Conference on Advanced Communication Technology(ICACT)on,vol.5,pp.855- 890,2018

[7] Jang Huisu,Jaewook Lee,Hyungjin Ko,Woojin Le, aPredicting Bitcoin price using Rolling Window LSTM modela ,DSF, ACM ISBN 123-4567-24- ˆ567/08/06,vol.4,pp.550-580,2018.

[8] Sin, Edwin & Wang, Lipo. (2017). Bitcoin price prediction using ensembles of neural networks. 666-671. 10.1109/FSKD.2017.8393351.

[9] Isaac Madan, Shaurya Saluja, Aojia Zhao, Automated Bitcoin Trading via Machine Learning Algorithms, Stanford:Department of Computer Science, Stanford University, 2015.

[10] John Mern1; Spenser Anderson1 ; John Poothokaran1 ,a Using Bitcoin Ledger Network Data to Predict the Price of Bitcoin

[11] Ruchi Mittal;Shefali Arora;M.P.S Bhatia; "Automated cryptocurrencies price prediction using machine learning" a,ICTACT JOURNAL ON SOFT COMPUTING, VOLUME: 08, ISSUE: ˆ04 JULY, 2018

[12] Amin Azari,a Bitcoin Price Prediction: An ARIMA Approach ˆ a,Available at: https://www.researchgate.net/publication/328288986, 2018

[13] http://wp.firrm.de/index.php/2018/04/13/building-a-lstmnetwork-completely-from-scratch-no-libraries/

[14] A. Mikhaylov. Asset Allocation in Equity, Fixed-Income and Cryptocurrency on the Base of Individual Risk Sentiment. Accessed: 2019. [Online]. Available: https://pdfs.semanticscholar.org/df78/f60a84c2a17f47bce27578746c6313251%a88.pdf

[15] D. L. K. Chuen, L. Guo, and Y. Wang, ''Cryptocurrency: A new investment opportunity?'' J. Alternative Investments, vol. 20, no. 3, pp. 16–40, 2017.