2019 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS)

Bitcoin price prediction using Deep Learning Algorithm

Engr. Muhammad Rizwan
Department of Information Technology
Mehran UET Jamshoro
rixwankhan04@gmail.com

Dr. Sanam Narejo
Department of Computer Systems Engineering
Mehran UET Jamshoro
sanam.narejo@faculty.muet.edu.pk

Dr. Moazzam Javed
Department of Computer Systems Engineering
Mehran UET Jamshoro
moazzam.jawaid@faculty.muet.edu.pk

Abstract*— The world has more than 5000 digital-currencies, bitcoin is one of it, which has more than 5.8 million dynamic client and approximately more than 111 exchanges throughout the world. So, the aim for this paper is to do the near prediction of the price of Bitcoin in USD. Precious details are taken from the price index of Bitcoin. A Bayesian recurrent hierarchical (RNN) neural network and a long-term memory (LSTM) network can accomplish this function. The total identification accuracy of 52% and an 8% RMSE is obtained by the LSTM. In contrast to the profound training systems, the common ARIMA method for the prediction of time series. This model have not much efficient as deep learning model can be performed. The deep learning methods were predicted to outperform the poorly performing ARIMA prediction. So here we used Gated Recurrent Network model (GRU) to forecasting Bitcoin price Eventually, all deep learning models have a GPU and CPU that beat the GPU implemented by 94.70 percent for their GPU training time.

***Keywords*— Autoregressive integrated moving average (ARIMA), Recurrent Neural Network (RNN), Long Short Term Memory (LSTM), blockchain, cryptocurrency, deep learning, predictive model, time series analysis**

I. Introduction

Bitcoin drives the digital currency market with 58% piece of exchanges; comparing to $4.9 Billion USD exchange volume and more than 5.8 million dynamic clients. In October 2008, Bitcoin was first presented by Satoshi Nakamoto through his white paper entitled "Bitcoin: peer-to- peer Electronic Cash System" [1]. Bitcoin is the first decentralized cryptographic money while other advanced monetary forms (otherwise known as Altcoin or option virtual monetary forms) are made by cloning or modifying the instrument of Bitcoin [2]. All exchanges constrained by cryptography make them secure, approved, and put away in "blockchain" by a decentralized arrange [3]. With the idea dependent on the new electronic money framework, online installment exchanges should be possible straightforwardly between any two consenting partakers without the requirement for a trusted outsider, for example, a money related foundation. Bitcoin was the biggest and generally well known in digital currency showcase estimated by advertise capitalization in March 2017. Bitcoin accounts involved 72% of the all-out cryptographic money in advertise and number of exchanges were 286,419 in January – February 2017 which are more than all different digital forms of money [2]. In 2013, the cost of Bitcoin was at 1,000 USD and went up to16,000 USD in December 2017. This makes Bitcoin's costs amazingly hard to foresee.

The Bitcoin's worth fluctuates simply like a stock yet in an unexpected way. There are various calculations utilized on stock advertise information for value expectation. Be that as it may, the parameters influencing Bitcoin are unique. Along these lines it is important to foresee the predictionof Bitcoin with the goal that right speculation choices can be made. The predictionof Bitcoin doesn't depend on the business occasions or mediating government not at all like stock advertise. In this manner. Thus, to predict the value we feel it is necessary to leverage deep learning to foresee the worth we feel it is important to influence AI innovation to anticipate the prediction price of Bitcoin.

Bitcoin has more than 40 exchanges worldwide, with over 30 different currencies[1] approved and exchanged. It has a current market capitalisation of $9 billion with over 250,000 trades weekly, according to Block Chain Intelligence. In terms of exchange, Bitcoin offers a new chance of market forecasts and therefore uncertainty much greater than fiat currencies owing to its relatively young age[4].It is also unique in its open nature in relation of traditional fiat currencies; no complete data on cash or fiat currencies are available. The Bitcoin is first ever crypto-currencies that was effectuated in 2009, eventually that became modish in 2012. The Crypto-currency is just codes which have some financial estimation Whereas Bitcoin is the first decentralized digital currency it is neither governed by any government nor any central bank in world.

The predictions for the global financial markets as the stock market have been studied[5][6] for a long time. Bitcoin has an interesting parallel to this because in a market still in its transitional period, it's a prediction issue for time series. Current forecast approaches for time sets like Holt-Winters are

978-1-7281-4956-1/19/$31.00 ©2019 IEEE

focused on static expectations which involve information that can be separated into averages, seasonal models and sound to be accurated [7]. It is the type of approach is better suited for a role like predicting revenue in situations where seasonal impacts exist. Since Bitcoin's economy lacks seasonality and low uncertainty, such strategies are not very productive for this mission. Because the problem is complicated, deep learning is the foundation of its success in a similar field for a fascinating technological solution. Because of the Bitcoin data's period existence, the repeating NN and Long-Term Memory (LSTM) are preferred above the generic MLP (traditional multilayer perceptron).

II. RELATED WORK

This researched work[8] based on blockchain toward characterized cryptocurrency for high accuracy in blockchain, in blockchain which connected to other block it is decentralize network this paper defined that it gather data and analyze user network end monitoring activity of user that has changed over time to time which relate to economic so recognize the features that estimate the requirement of cryptocurrency so for this purpose used machine learning predict attributes on user activiry like number of wallet, unspent transaction, output block size, income transaction per day and wallet and unique address estimate how many user join network each day. The all correlation co-related to cryptocurrency market analyze data over 20 month and calculated most significant features to predict bitcoin accurately and improve result up to 99.4% and gave 0.0113 (RMSE). In research [9]work proposed the Bitcoin price prediction comparison using Machine Learning, in this paper discussed about Bitcoin. Bitcoin is actually digital currency that is used in globally to pay in economical system, so that Bitcoin prediction is important to predict, the Bitcoin is fluctuated which make them difficult to predict So this proposed work aim to uncover most efficient and highest reliability techniques to predict Bitcoin Price from various machine learning tools. This researched used multiple regression model and compared with last model which name was Huber regression it has less time compute prediction, but GRU model has more accuracy in result. But it also took more time to predict as compare to Huber model, GRU model produced accuracy in prediction by using features of low high etc. it also use some parameter of Keras modeling like epochs, batch size and layer etc. That gave us 99% accurate prediction using machine learning. In research paper[10] work on predicting the price accurately by used of machine learning algorithm in this proposed work two part considered, the first is to recognize daily trend in Bitcoin market while obtaining insight optimal attributes surrounding Bitcoin price collected dataset to consist of various features relating payment network, Bitcoin price over five year recorded and second part of this is to investigate of using available data. So first part is to collect data from different source Quandle and Coinmarket Cap and the second part is to normalize the all information using normalization techniques like Z-score, normalization and boxes normalization etc. This proposed work using different techniques to predict Bitcoin price with time series that is Bayesian regression which break data to all feasible consecutive interval of size then implement K-mean clustering and other one is to GML/Random forest that also construct three time series data then apply on it. The result is acquired better in these two techniques after normalization of information. In research[11] proposed work shown how can predict Bitcoin return using high dimension technical indication, the name as reverse engineering of prediction that predict the bitcoin price return or remain same that is another technique is to split domain regularly, Bitcoin return into 21 interval and aim is to forecast these all interval of next day return, so for this we built return prediction/model using 124 indicator that related to historical price. Here used decision tree classification it works is to selected a non parametric supervised learning method that made chain of all possible chance of outcome for bit coin price return, and it also used Classification and Regression Tree(CART) method. CART method based on binary tree and main difference lies in CART support numerical target regression does not compute rule set due to binary tree used to target information given to each node. That's why here used decision tree to find result, these techniques built top down root node and splitting data into subset to return Bitcoin price. This work proposed[12] the prediction of bitcoin using machine learning prediction is estimated in term of USD, the source of price data taken from bitcoin price index, the accuracy improved with changing degree of success to implemented two techniques to increase efficiency in prediction so he used Bayesian optimized recurrent network and Long Short term Memory (LATM). Result produced highest classification accuracy 52% and Root Mean Square Error (RMSE) of 8% on LSTM and RNN achieved 7.15% error (RMSE) and classification accuracy 52.78% improved, but RNN was effectively using temporal length 50 days although LSTM performed more in the 50 to 100 days range with 100 days it gave you good performance.

III. METHODOLOGY

This paper reflects the CRISP technique of data mining. The CRISP-DM motivation for the traditional KDD[26] focuses on the company-level of the forecasting task. The data set is used by Bitcoin covers the period 19 August 2013 to 19 July 2016. Figure 1 displays a time series graph of this. Data is omitted from prior to August 2013 as they no longer represent the network correctly. Dataset is used in bitcoin Ethereum Historical Data and Bitcoin Historical Data. CSV files for bitcoin exchanges from Jan 2014 to July 2019, with by-the-minute updates of OHLC (Open, High, Low, Close), Volume in BTC and currency, as well as weighted bitcoin price. The information was also normalized in order for it to have a mean of 0 and default of 1. The start first five sample are shown in table 1.

TABLE I: FIRST FIVE SAMPLE OF DATASET.

Timestamp	Open	High	Low	Close
1325317920	4.39	4.39	4.39	4.39
1325317980	4.39	4.39	4.39	4.39
1325318040	4.39	4.39	4.39	4.39
1325318100	4.39	4.39	4.39	4.39
1325318160	4.39	4.39	4.39	4.39
1325318220	4.39	4.39	4.39	4.39
1325318280	4.39	4.39	4.39	4.39
1325318340	4.39	4.39	4.39	4.39
1325318400	4.39	4.39	4.39	4.39
1325318460	4.39	4.39	4.39	4.39

Standardization was chosen through standardization because it best suits deep learning models ' activation functions.

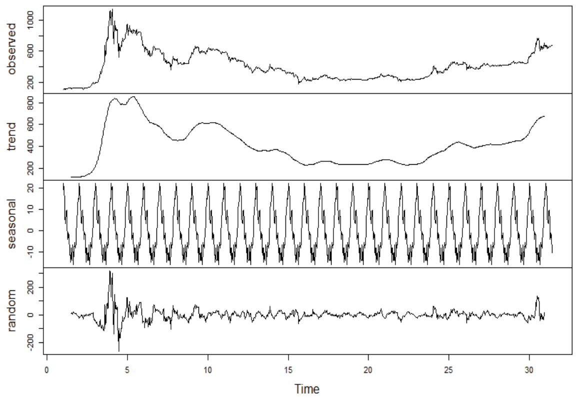


FIGURE I: Bitcoin Time Series Information Disintegration

THE GIVEN FIGURE SHOWN TRAINING ERROR WHICH IS GIVEN BELOW.

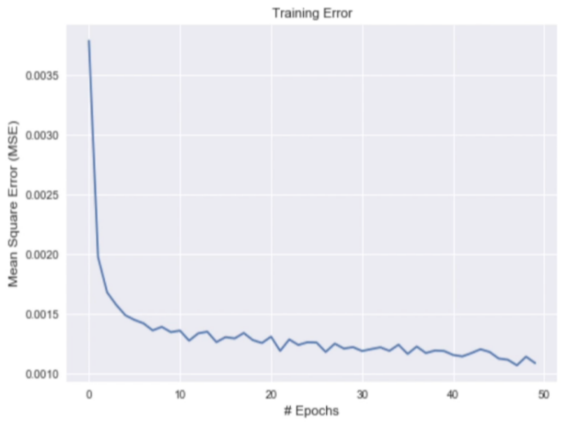


FIGURE II: Training Error of Dataset

IV. PROPOSED METHODOLOGY

The strategy proposed in this paper can be pictured on Fig. 3, which abridges the: (1) sources where the information were gathered; (2) changes utilized for information pre-handling; (3)the information apportioning for preparing and testing purposes; (4) kind of property determination strategies applied; (5) utilization of deep learning methods to forecast the cost and (6) execution assessment measurements is utilized.

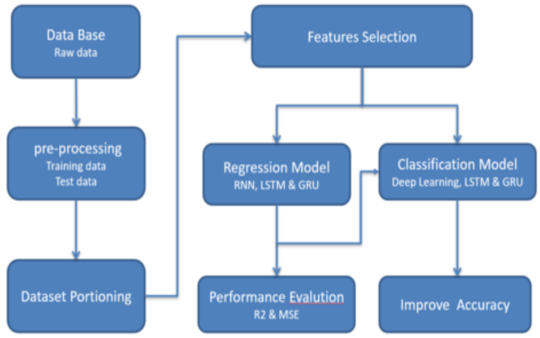


FIGURE III. Overview of the Proposed Methodology

A. Collected data

The sources of information can be categorized into internal (the behavior of different parameters of Bitcoin) and external (the economic factors, external demand or information obtained from social networks or specialized forums, also named as public). Thisinformation includes opening, maximum, minimum, and closing Bitcoin exchange rate (OHLC), the volume of trades, total transaction fees, number of transactions, cost per transaction and average hash rate. As a contribution of the present paper to the identification of relevant attributes for the prediction of the Bitcoin price trend,external information was considered and obtained from international economic indicators.

B. Data pre-processing

In this paper, the data pre-handling stage proposed by was utilized, i.e., the slack time frame idea and the smoothing of the information. Subsequently, in the pre-preparing stage, the worth "1" was doled out to the class if the end conversion scale of Bitcoin at a Day (D) is more prominent than or equivalent to the earlier day (D − 1). Else, it was relegated the worth "0".

C. Data partitioning

So as to contrast the acquired outcomes and the system proposed by, same dataa interval(named as interval of 1) was considered, which considers similar information parceling (70%of the data for preparing and the staying 30% latest data for approval/test). Additionally, a bigger information interim was additionally considered and used to produce a standard for future researches.

D. Features Selection

Features of the datasets from A are as follows:

Highlights Definition Close most recent exchange Open opening exchange High most elevated exchange during day Low least exchange during day Weighted value mean Bitcoin value Volume_(BTC) all out exchange volume of y BTC

Volume_(Currency) all out exchange volume of day USD Timestamp information recorded time In this exploration, the scikit-learn library is utilized to make models with just highlights Close, Open, High, and Low when the Weighted value is to foresee.

E. Modeling

In this research, we took regression AI because of nonstop estimations of Bitcoin cost. With the scikit-learn library, the best two regression models. Recurrent Neural Network(rnn), Long Short Term Memory(lstm). For profound learning-based relapse models, Keras library was utilized to make LSTM and GRU models.

V. DEEP LEARNING MODELS

A proper deep learning design model is essential for their success in terms of network parameters. Random selection, matrix quest and heuristic mixture strategies such as genetic algorithms are three principal choices for choosing deeper learning models parameters. This study was conducted by manual matrix quest and Bayesian optimisation. The quest for Grid, carried out for the Elman RNN, requires choosing two minimum and maximum hyper parameters for each hyper parameter. You then look for the right parameters for that function space. This approach was used for parameters that could not be used for optimization in Bayesia. In the python programming language, this template was developed with Keras[13]. Compared to the RNN the LTSM parameters are used for Bayesian optimization are chosen for the possible scenarios. This is a heuristic query approach which takes the function as Gaussian sampled and preserves a corresponding distribution for this feature because the outcomes of several hyper parameter choices are detected. The predicted change can be improved over the best results in the next experiment in selecting hyper parameters [14]. The RNN and LSTM network quality are tested on validation information utilizing over-fitting prevention measures. Drop-out is applied in two layers and, unless its validity failure in five epochs changes, they automatically stop template learning.

VI. MODELLING & SIMULATION

A. RNN

RNN is a time series based neural network algorithm [15]. It typically retains hidden state information. In the previously hidden state, the current state is determined and the new hidden state is used to determine the corresponding state as shown in equations below.

When “t” is time, σ is activation function, omeans output data, xmeans input data,h is a hidden state,,W, U, Vis an output weight, previously hidden state and existing hidden state estimation matrix, respectively. Prediction of RNN model can beseen infigure1.

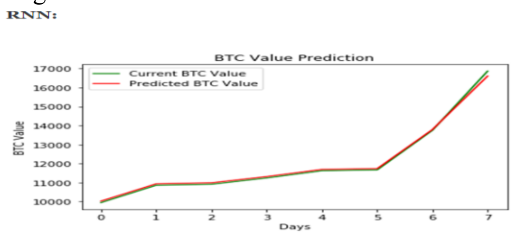


FIGURE I: RNN Prediction model

The number of predictions are close to the actual value but RNN is ideal for sequence or time serial data, but it has a long-term dependence data problem that causes discontinuity.
When predictions were done on a set or around 100 dates, the number of predictions close to the actual value were as shown in the table. The table1 can be read as around 84 predictions made by the RNN model were as close as 90–100% of the actual value.

TABLE I: PREDICTION MADE BY RNN

Accuracy	RNN
90-100%	84
80-90%	16
70-80%	-

B. LSTM

In order to resolve the disappearance problem, the LSTM is developed in RNN[16]. Information is obtained and sent to be stored in the next state using the cell state and the secret state. The gates to add, exit, and forget should decide whether or not the data can move according to the priority of the data. We can see the prediction of bitcoin price throughout training in the given figure. As defined in the following formula, the extinct gradient can thus be covered.

$$h_t = \sigma(x_t W + h_{t-1} U)$$

$$o_t = \sigma(h_t V)$$

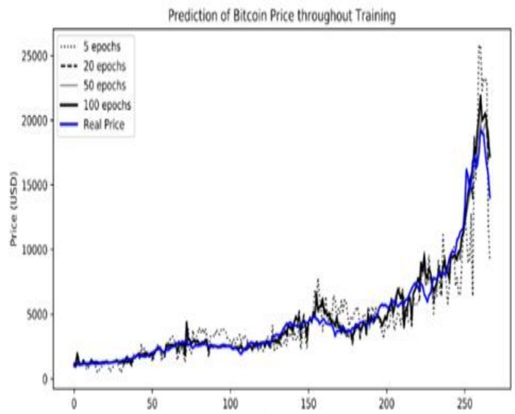


FIGURE II: LSTM Prediction model throughout training.

$$i = \sigma(x_t W_i + h_{t-1} U_i)$$
$$f = \sigma(x_t W_f + h_{t-1} U_f)$$
$$o = \sigma(x_t W_o + h_{t-1} U_o)$$
$$c_t = (c_{t-1} \times f) + (i \times \sigma(x_t W_c + h_{t-1} U_c))$$
$$h_t = \sigma(c_t) \times o$$

Where,f is forget gate, i is input gate c is cell state, his hidden state,ois output gate,σis

Activation function, Wand Uare weight matrix, and t is time. Description of RNN node structure can be seen in figure 3.

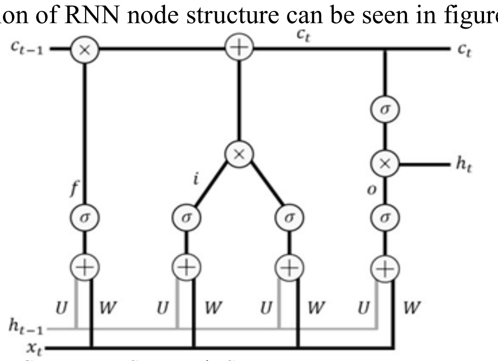


FIGURE III. LSTM node Structure

Keras modeling parameters are described by LSTM first.

C. GRU

The GRU, which has been developed from LSTM, was proposed by Kyunghyun Cho et al[10] in 2004. By adjusting the key to reset and upgrade the door in LSTM, GRU has a less complicated structure than LSTM. A reset door defines the number of prior state data for current data, while a fix gate calculates the value of the previous state. In every layer the regularization is added to lessen over fitting. The regularization function used is dropout which permits few neurons to be detached to avoid gradient.

VII. CODE FOR GRU LAYER

```
regressorGRU = Sequential()
First GRU layer with Dropout regularisation
regressorGRU.add(GRU(units=50,    return_sequences=True,
input_shape=(X_train.shape[1],1)))
regressorGRU.add(Dropout(0.2))
Second GRU layer
regressorGRU.add(GRU(units=50, return_sequences=True))
regressorGRU.add(Dropout(0.2))
Third GRU layer
regressorGRU.add(GRU(units=50, return_sequences=True))
regressorGRU.add(Dropout(0.2))
Fourth GRU layer
regressorGRU.add(GRU(units=50))
regressorGRU.add(Dropout(0.2))
The output layer
regressorGRU.add(Dense(units=1))
```

FIGURE I: GRU Prediction Model

VIII. RESULTS

The site was tested by running the code on a local facilitated server. Utilizing Anaconda command prompt, "MakeMigrations" was accumulated to ensure all information documents arranged appropriately. Dataset Testing: tested running numerous varieties of similar informational index to distinguish which was gave the best result. Network Tuning: Testing proficiency of yields of the neural system by tuning the quantity of neurons, expanding and diminishing the quantity of layers. Analysis: Testing code on various conditions to identify potential hole and situations of failure. Modify the code to suit the need of the facilitating condition.

Testing with various layers GRU

- Testing with 2 layers - Two layers didn't give acceptable outcomes as the model had predicted number.

with over 30% contrast rate structure the first qualities. This caused because of inadequate number of neurons and layers. Condenses the presentation of the entirety of the machine learning models worried at the Bitcoin cost expectation. From the outcomes, we can mention the accompanying objective facts. True to form, the consequences of the various techniques RNN, LSTM and GRU are testified yet GRU model is better in between these models. The normal exactness of the GRU techniques is 94.70%, higher than the normal precision of the RNN and LSM models (52.0%), Which based on previous worked.

The two most common metrics are the R-square (R2) and mean squared error (MSE) for calculating precision of continuous variables. Table 1 demonstrates the R2 and MSE for all of our deployed models and Table 2 indicates the deployment time of our models. The results show the progress of the deep-learning regression models: GRU and LSTM. At 0.00002 and 0.992 or 99.2% of GRU, MSE's best results are given. In comparison, GRU's measured time is lower than LSTM.

TABLE I. MSE AND R2 OF IMPLEMENTED MODELS

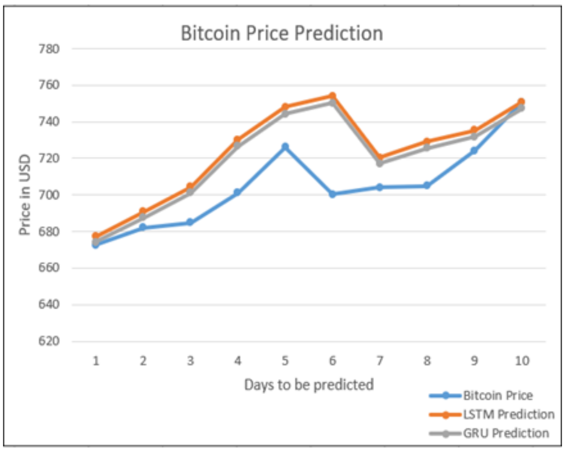
Model	R2	MSE
GRU	0.992	0.00002
LSTM	0.992	0.00043

TABLE II. Implemented MODULE Time

Model	Time (sec)
GRU	85.2718
LSTM	111.0601

TABLE III. First 10 days comparison between the predicted prices of Implemented Modules.

Bitcoin Price in (USD)	LSTM Price Prediction in (USD)	GRU Price Prediction in (USD)
672.625	677.108	674.093
682.195	690.657	687.543
684.804	704.314	701.049
701.076	730.109	726.588
726.058	747.952	744.239
700.377	754.145	750.336
704.181	720.418	716.989
705.016	728.981	725.423
724.068	735.318	731.687
750.929	750.929	747.14



Graphical Representation of 10 days Comparison of Price prediction.

The proposed model has been compared with the existing literature including RNN, LSTM and ARIMA as shown in Table 4. Accordingly, it can be observed that the proposed model outperforms the existing methods in both Accuracy and Error rate.

TABLE 4: Analytical Comparison

S.No	Models	Accuracy	Error Rate
1	RNN	85.40%	14.60%
2	LSTM	92.30%	7.70%
3	ARIMA	91.00%	9.0%
4	GRU	94.70%	5.30%

IX. CONCLUSION

Our aim of study is to develop model which predict bitcoin price using deep learning. Since deep learning is used to select the parameter to get successive outcomes in developing model. In this latter we implemented for three proposed model RNN, LSTM and GRU we found that total parameter and dataset can influence result. The Previous model developed using RNN and LSTM which had less predicted accuracy that is 52% approximately. Whereas in our comparative analysis GRU model result better as comparatively LSTM model. The Optimal model for GRU result accuracy is 94.70%. The proposed model shows about 42.3% of accuracy improvement. The tests of all our GRUs show the highest detailed outcomes which take time. Furthermore, Selected features: Low, High, Close and Open can't be enough to predict the Bitcoin value, as various factors, including social media responses, legislation and laws each country advertises for handling the digital currency won help to increase and lower the Bitcoin price. Therefore, modified information should always be gathered and applied for the best results of all models.

REFERENCES

1. S. Nakamoto, “Bitcoin: A peer _to_peer electronic cash system,” 2008.
2. G. Heleman, M. Rauchs, “Global crypto-currency bench-marking case study,” Cambridge Centre for Alternative Finance (2017).
3. G. Neil, H. Halaburda. " Can we predict the winner in a market with network effects?? Competition in crypto-currency market," Games, vol.7 no.3 , 2016, p.p: 16.
4. M. Brie`re, K. Oosterlinck, and A. Szafarz, “Virtual currency, tangible return: Portfolio diversification with bitcoins,” Tangible Return: Portfolio Diversification with Bitcoins (September 12, 2013), 2013.

5. I. Kaastra and M. Boyd, “Designing a neural network for forecasting financial and economic time series,” Neurocomputing, vol. 10, no. 3, pp. 215–236, 1996.
6. H. White, “Economic prediction using neural networks: The case of ibm daily stock returns,” in Neural Networks, 1988, IEEE International Conference on. IEEE, 1988, pp. 451–458.
7. C. Chatfield and M. Yar, “Holt-winters forecasting: some practical issues,” The Statistician, pp. 129–140, 1988.
8. M. Saad and A. Mohaisen, "Towards characterizing blockchain-based cryptocurrencies for highly-accurate predictions," IEEE INFOCOM 2018 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Honolulu, HI, 2018, pp. 704-709.(2018)
9. T. Phaladisailoed and T. Numnonda, "Machine Learning Models Comparison for Bitcoin Price Prediction," 2018 10th International Conference on Information Technology and Electrical Engineering (ICITEE), Kuta, 2018, pp. 506-511.(2018)
10. S. Velankar, S. Valecha and S. Maji, "Bitcoin price “Top 10 cryptocurrency (2017), — best cryptocurrency to invest.”
11. Roy, S., Nanjiba, S., & Chakrabarty, A. (2018, December). Bitcoin Price Forecasting Using Time Series Analysis. In 2018 21st International Conference of Computer and Information Technology (ICCIT) (pp. 1-5). IEEE.
12. S. McNally, J. Roche and S. Caton, "Predicting the Price of Bitcoin Using Machine Learning," 2018 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP), Cambridge, 2018, pp. 339-343.(2018)
13. M. Daniela and A. BUTOI, “Data mining on Romanian stock market using neural networks for price prediction”. informatica Economica, 17,2013.
14. U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, “The kdd process for extracting useful knowledge from volumes of data,” Communications of the ACM, vol. 39, no. 11, pp. 27–34, 1996.
15. A. Bernal, S. Fok and R. Pidaparthi. "Financial Market Time Series Prediction with Recurrent Neural Networks." 2012.
16. S. Hochreiter, J. Schmidhuber. "Long short-term memory." Neural computation vol.9 no.8, 1997, pp. 1735-1780.
17. Jang, H., & Lee, J. (2018). An empirical study on modeling and prediction of bitcoin prices with bayesian neural networks based on blockchain information. IEEE Access, 6, 5427-5437.
18. G. Hileman and M. Rauchs,(2017), “Global cryptocurrency benchmarking study,” Cambridge Centre for Alternative Finance,.