JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

RESEARCH ON BITCOIN PRICE PREDICTION USING MACHINE LEARNING

[1]Shailesh S. Mahajan, [2]Prof. S. R. Waghe, [3]Shantanu S. Dondal, [4]Vedant S. Mankar
Department of Computer Science and Engineering, Jawaharlal Darda Institute of Engineering and Technology, Yavatmal, Maharashtra, India

Abstract: Bitcoin is a kind of Cryptocurrency and now is one type of investment on the stock market. Stock markets are influenced by many factors. And bitcoin is one kind of cryptocurrency that keeps rising in recent years, and sometimes suddenly falls without knowing the influence behind it on the stock market. Because it's fluctuations, there's a need for an automation tool to predict bitcoin on the stock market. This research study learns how to create model prediction bitcoin stock market prediction using LSTM, LSTM (Long Short-Term Memory) is another type of module provided for RNN later developed and popularized by many researchers, like RNN, the LSTM also consists of modules with recurrent consistency. The Method that we apply to this research, also technique and tools to predict Bitcoin on the stock market.

Keywords- Cryptocurrency, Bitcoin prediction, Bitcoin Stock Market Prediction, LSTM.

I. INTRODUCTION

Cryptocurrency has been around for several years and has now become quite popular, widespread, and also surrounded and there is a lot of controversy from innovative developments. Cryptocurrencies are a digital currency where transactions can be done by online transactions, unlike the common currency, cryptocurrency is designed based on cryptography. Bitcoin is one kind of Cryptocurrency with no regulation from any party and decentralised. The unique characteristic of Bitcoin is daily price fluctuations and always change every day. The Bitcoin Exchange rate to (USD) is $ 12,354.73 USD on 28 June 2019 in yahoo finance stock market and sometimes keep rising and sudden fall in March the value is $ 3900 USD. The Stock markets are influenced by many uncertainties such as political issues, the economic issue is impacted to local or global levels. To interpret the key of success, to provide accurate predictions is complicated work. For the market, we can analyse with any techniques such as technical indicators, price movements, and market technical analysis. To solve the problem above, regarding the fluctuations there's a need for an automation tool for prediction to help investors decide for bitcoin or other cryptocurrency market investment. Nowadays the automation tools are usually used in common stock market predictions, and we can do the same work and strategy in this domain of cryptocurrencies. LSTM (Long Short-Term Memory) is another type of module provided for RNN. LSTM was created by Hochreiter & Schmidhuber (1997) and later developed and popularized by many researchers. Like RNN, the LSTM) also consists of modules with recurrent consistency. This project studies bitcoin and stock market predictions, method, technique, and tools from a big number of resources, paper, and other available sources.

Instead of any direct human investments, generating profit with the help of algorithms is a common practice in the stock market. Many case studies have been performed to reach the conclusion that mathematical models warrant better results than humans. Bitcoins are an eye-catching initiative in the fields of cryptography, economics, and computer sciences, as such currencies have a special character which is gained when integrating currency units with cryptographic technology. Due to the fact that

cryptocurrency has a minute history, when compared to the stock market, new and unexplored territories are thus being scouted. Structurally, both the stock market and the cryptocurrency price data are having characteristics such as time series data, but high volatility is routinely present in the latter, with heavy wavering in the prices. A cryptocurrency market differs from a trad itional stock market in the respect that the former has a lot of new features. It is required to apply new techniques for prediction suitable for the cryptocurrency market. Fewer studies have been conducted on cryptocurrency price prediction when compared to the stock market. In this project, we are predicting the Bitcoin price trend using a Long Short-Term Memory (LSTM) model.

II. LITERATURE REVIEW

[1] Numnoda et al. Have obtained highly accurate results on implementing their prediction Gated Recurrent Unit (GRU) model. However, their prototype has a large time complexity. Thus, complicating the expected results in this ever-changing environment. Additionally, the selected features aren't enough to predict the Bitcoin prices; as various factors like social media, policies, and laws that each country announces to deal with digital currency, can all play a major effect on the fluctuation of the Bitcoin prices.

[2] Mangla et al. Have compared four different price prediction models: Recurrent Neural Networks (RNN), Logistic Regression, Support Vector Machine, and Auto Regressive Integrated Moving Average (ARIMA). Their major findings are that-ARIMA performs poorly for predictions extending beyond the next day. Their RNN model can accurately predict price fluctuations for up to six days. And the logistic regression model can give accurate results only if a separable hyperplane exists.

[3] Guo et al.Have used a hybrid method consisting of multi-scale residual blocks and an LSTM network to predict Bitcoin price. Although, their work does not include comprehensive metrics which measure the investor's attention to more timely detection of bitcoin market volatility, therefore resulting in a less accurate prediction.

[4] Awoke et al. have considered basic deep learning models like GRU and LSTM. However, their research lacks further investigation to enhance the model accuracy by considering different parameters.

[5] Wei Dai proposed b-money, an anonymous and distributed electronic cash system, In that method, describes two protocols based on network that cannot be traced, where senders and receivers are identified only by digital information such as their public keys, and each message will be signed by its sender to receiver.

[6] Nick Szabo Proposed models of a new digital currency, the models based on cryptographic system puzzles, which after being solved, were sent to the Byzantine fault-tolerant public registry and assigned to the public key of the solver.

[7] Hashcash proposed by Adam Back, Haschash, a system relied on a cryptographic hash function to derive a probabilistic proof of computational work as authentication system Pow (Proof of work)

[8] Hal finney proposed a currency system based on a reusable proof of work (RPOW) in 2004.

[9] Between 2008 and 2009, Bitcoin was made the first decentralized cryptocurrency by Satoshi Nakamoto. Nakamoto published the Bitcoin whitepaper in 2008, and after January 3rd, 2009, the genesis block of the bitcoin protocol was created. Nowadays it is most successful cryptocurrency in terms of market capitalization, beside above 700 altcoins that circulated in the world (eg. Litcoin, Etherum) based on Bitcoin have been proposed since the launch of Bitcoin

[10] Greaves et al. Proposed a technique using Logistic Regression and SVM and analyzed using Graph to predict bitcoin price.

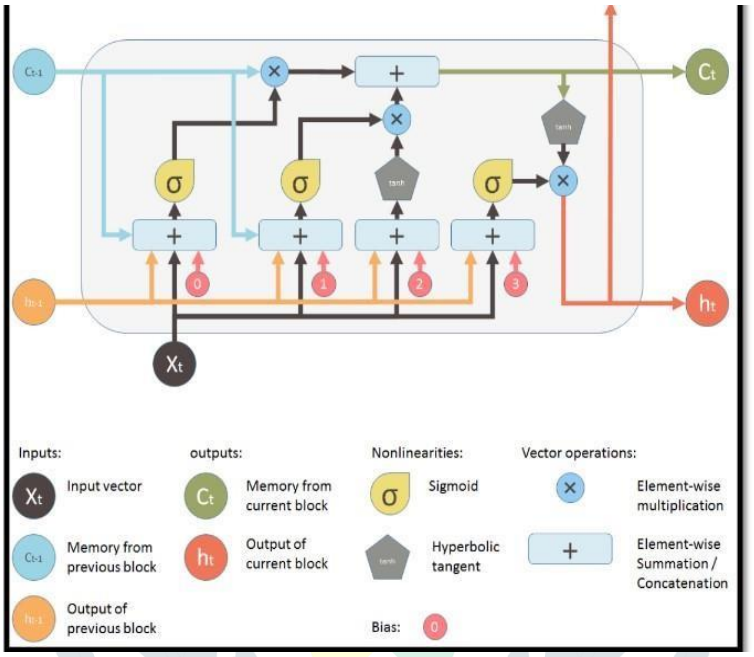
III. METHODOLOGY

LSTM ALGORITHM

1. Initialize the LSTM state: The LSTM has a hidden state vector h and a memory cell vector C, both of which are initialized to zero.
2. Process input sequence: For each time step t in the input sequence, perform the following operations:
3. Compute the input gate: Calculate the input gate i by applying a sigmoid activation function to a linear combination of the current input x_t and the previous hidden state h_{t-1}.
4. Compute the forget gate: Calculate the forget gate f by applying a sigmoid
5. activation function to a linear combination of the current input x_t and the previous hidden state h_{t-1}.

6. Generate output sequence: After processing the entire input sequence, the LSTM generates an output sequence by applying a linear transformation to each hidden state vector h_t. The output sequence can be used for tasks such as sequence prediction, text generation, or sentiment analysis.
7. Generate output sequence: After processing the entire input sequence, the LSTM generates an output sequence by applying a linear transformation to each hidden state vector h_t. The output sequence can be used for tasks such as sequence prediction, text generation, or sentiment analysis.
8. The Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN)architecture that is designed to effectively capture long-term dependencies and memory in sequential data. The LSTM cell consists of several components, including an input gate, a forget gate, an output gate, and a memory cell. The memory cell is the central component of the LSTM and is responsible for storing and updating information over time.

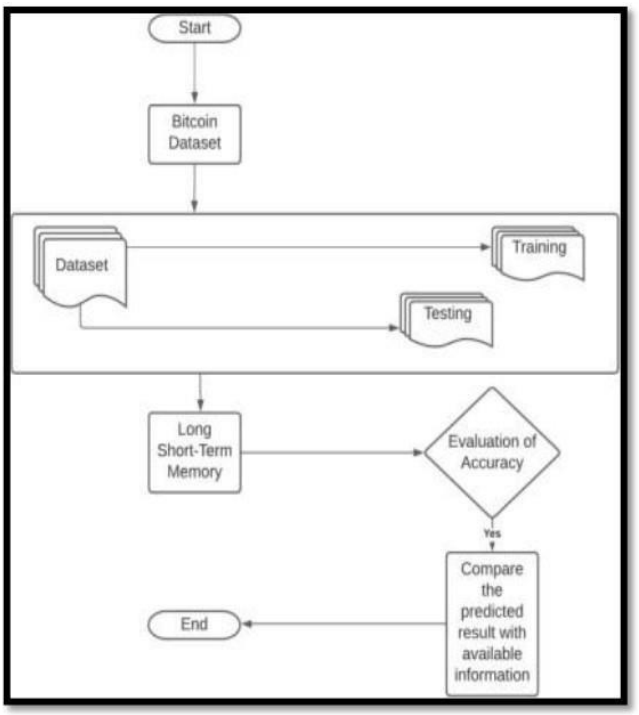
LSTM MEMORY CELL STRUCTURE OF THE HIDDEN LAYER



SHA 256

The mathematical operation that converts the input into an encrypted output of the same fixed length is called a hash. Depending on the algorithm used, the length of the output can vary, such as 32, 64, or 128 characters. Since we are talking about theSHA256 algorithm used by Bitcoin in this article, the output will be 256 bits, which is 64 characters long. Whether the input is a word, a letter, or a volume of an encyclopaedia, the SHA256 algorithm always produces a fixed length of characters (64 characters).

FLOW CHART



WORKING

This project aims to predict Bitcoin prices using a Long Short-Term Memory (LSTM) algorithm implemented in Python. Utilizing TensorFlow and Pandas, the LSTM model will be trained on a predefined dataset, with input sequences processed through input gates, forget gates, and output gates within LSTM memory cells. The system will be built using Python, running on Jupyter Notebook, and displaying results graphically using Matplotlib and Seaborn. The SHA256 algorithm, central to Bitcoin, will be utilized for data encryption. Required modules include TensorFlow, Pandas, NLTK, Scikit-learn, Plotly, and the Anaconda distribution. The project will generate price predictions displayed in graphical format, enabling insights into Bitcoin price trends over time.

IV. SYSTEM REQUIREMENT

SOFTWARE REQUIREMENT

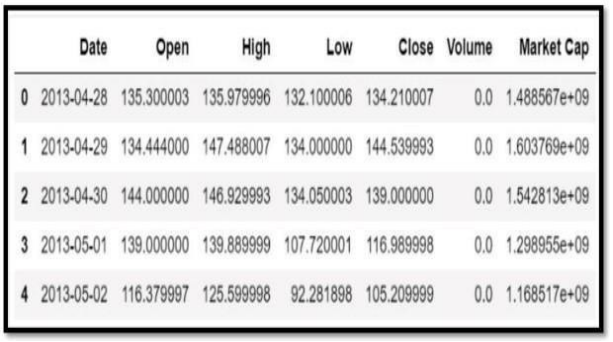
1. Python 3
2. Anaconda Distribution
3. NLTK Toolkit
4. Operating System

MODULE USED

1. Python
2. Tensorflow
3. Tensorflow
4. Pandas
5. Spyder IDE
6. Matplotlib And Seaborn
7. Data Sources
8. Sklearn
9. Plotly

IMPLEMENTATION

Once the libraries are imported and the historical data obtained, we can begin by reading the dataset as a data frame.



	Date	Open	High	Low	Close	Volume	Market Cap
0	2013-04-28	135.300003	135.979996	132.100006	134.210007	0.0	1.488567e+09
1	2013-04-29	134.444000	147.488007	134.000000	144.539993	0.0	1.603769e+09
2	2013-04-30	144.000000	146.929993	134.050003	139.000000	0.0	1.542813e+09
3	2013-05-01	139.000000	139.889999	107.720001	116.989998	0.0	1.298955e+09
4	2013-05-02	116.379997	125.599998	92.281898	105.209999	0.0	1.168517e+09

Screenshot 1: Bitcoin Dataset

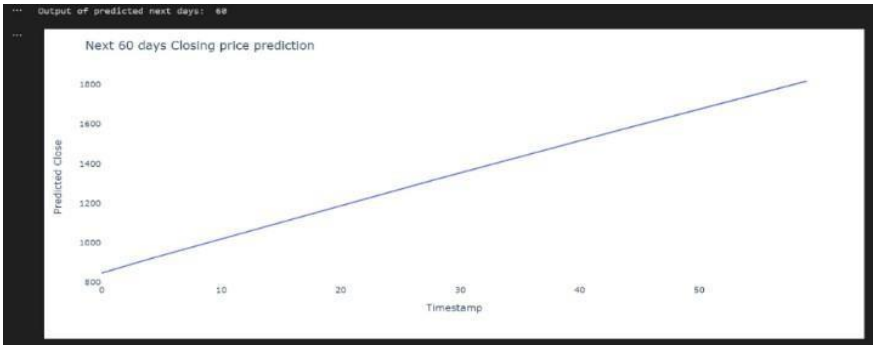
```
Date          0
Open          0
High          0
Low           0
Close         0
Volume        0
Market Cap    0
dtype: int64
```

Screenshot 2: Data Frame

	Date	Open	Close
0	January	8318.95	8389.27
1	February	9656.22	9630.72
2	March	6943.51	6871.02
3	April	7150.61	7224.48
4	May	9237.76	9263.15
5	June	9499.80	9489.23
6	July	9519.38	9589.90
7	August	11639.10	11652.39
8	September	10689.70	10660.28
9	October	11791.31	11886.98
10	November	16450.12	16645.76
11	December	21680.54	21983.14

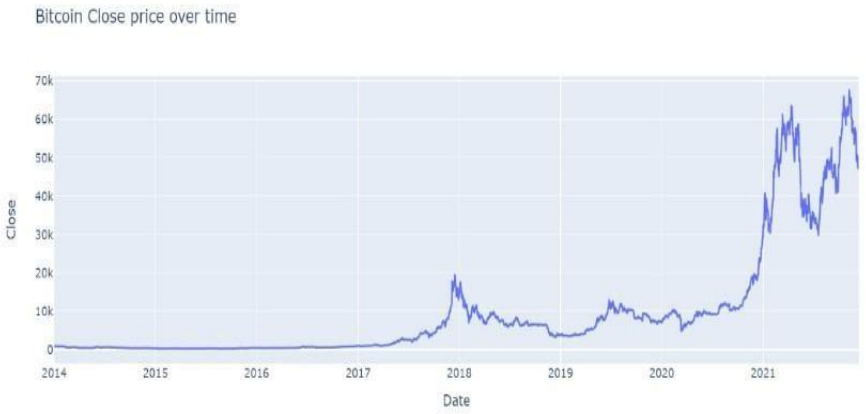
Screenshot 3: Data of Bitcoin prediction

With that, we can plot a simple bar chart to compare the average monthly opening and closing prices of Bitcoin.



Screenshot 4: Line graph that appears to show the next 60 days of predicted cleaning prices.

We can also plot a simple line chart to view the closing price of our time series.



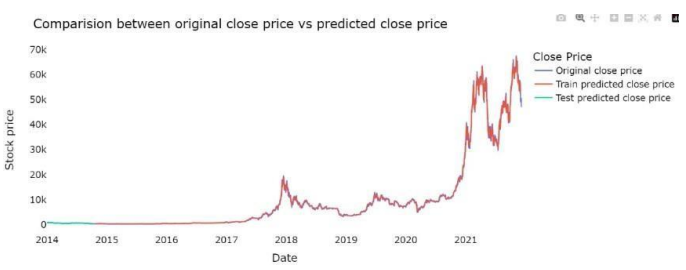
Screenshot 5: Graph of Bitcoin Close Price

LSTM MODEL

The next step of the process comes the part whereby we can start building the LSTM model for closing price prediction. Since we want to predict the closing price, we will create a new data frame to only include these data of interest for our model.This is a block of Python code that creates a new DataFrame called 'model_data' by selecting two columns, 'Date' and 'Close', from an existing DataFrame called 'df'.

	Date	Close
0	2013-04-28	134.210007
1	2013-04-29	144.539993
2	2013-04-30	139.000000
3	2013-05-01	116.989998
4	2013-05-02	105.209999
...	...	...
3158	2021-12-20	46880.278964
3159	2021-12-21	48936.614151
3160	2021-12-22	48628.509925
3161	2021-12-23	50784.538179
3162	2021-12-24	50822.196523

3163 rows × 2 columns



Screenshot 6: Matrix of Price of Bitcoin

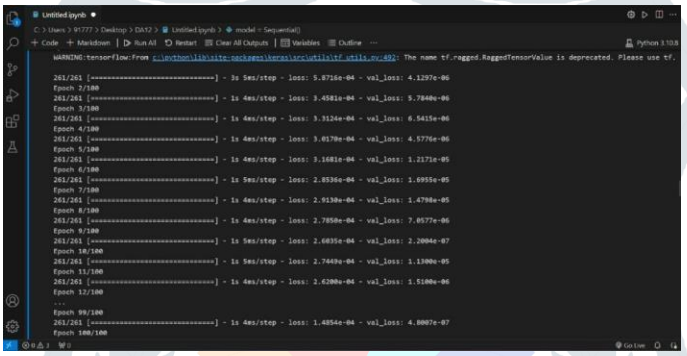
TRAINING AND TESTING

The next step would involve partitioning the data frame into a training and testing split.

```
train_data:  (2846, 1)
test_data:  (317, 1)
```

Screenshot 7: Training and testing

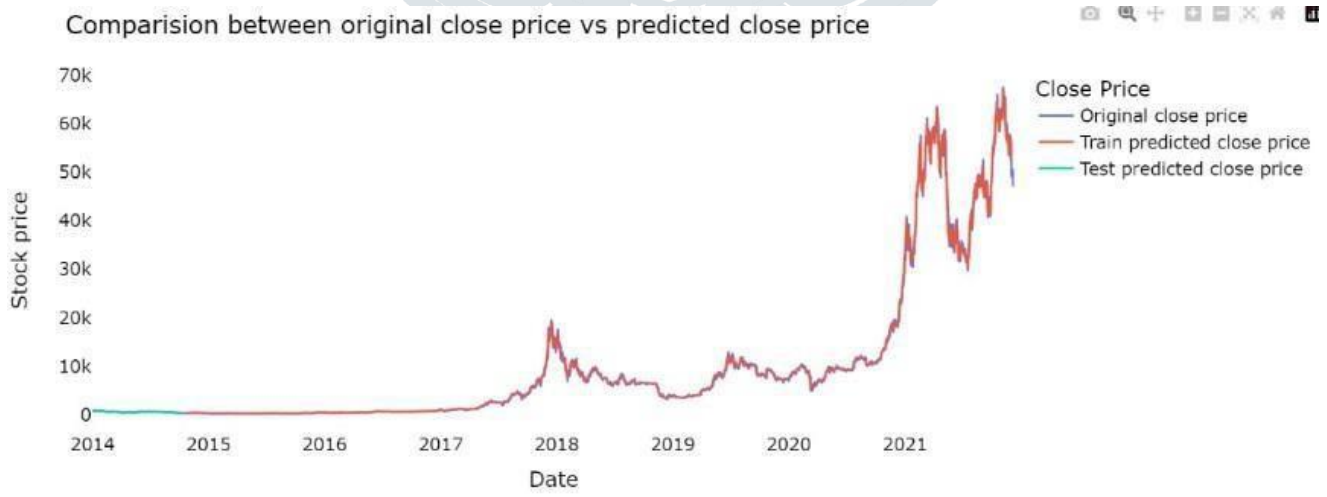
We can then compare the validation and training loss where: Validation loss indicates how well the model fits new data. Training loss indicates how well the model fits the training data



Screenshot 8: Training and validation loss

MODEL PREDICTION

Finally, we can predict closing prices from our LSTM model using parameters defined above. The snippet is used to plot the original close price and predicted close price of a stock using a trained model train_predict=model.predict(X_train) and test_predict=model.predict(X_test) are used to get the predicted values for the training and testing datasets respectively.



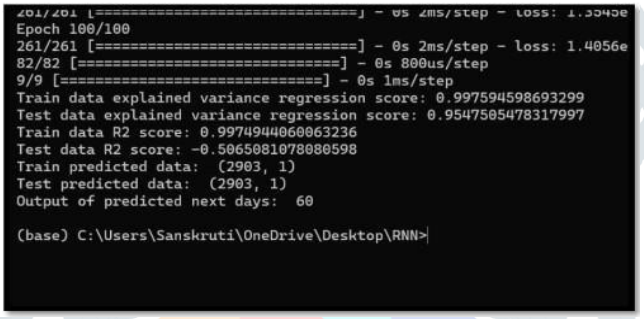
Screenshot 9: Comparison between original vs predicted close price

OUTPUT OF PREDICTED DATA

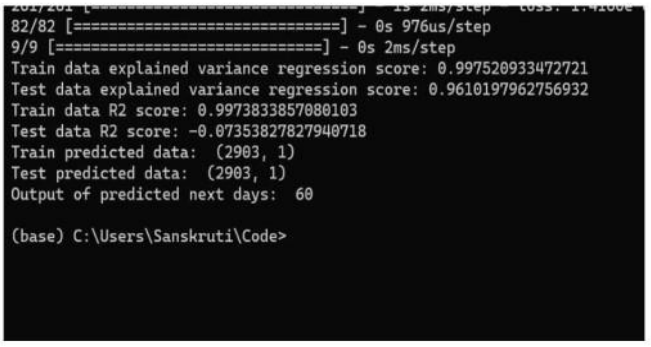
```
Train data RMSE:  371.42820914833123
Train data MSE:  137958.9145511365
Train data MAE:  191.58888747748279
----------------------------------------------------------------------------------------
Test data RMSE:  18457.691164883345
Test data MSE:  340686363.1382126
Test data MAE:  13096.214327396903
----------------------------------------------------------------------------------------
Train data R2 score: 0.9961632642208492
Test data R2 score: -2.7217571091142965
```

RESULT

RNN Model's Output



LSTM Model's Output



V. CONCLUSION

Our proposed model provides Bitcoin price predictions from Yahoo Finance stock market data. Using time series techniques, our model can produce results that predict the price for the next few days by splitting the data into training and test sets. The stock market is influenced by many uncertain factors, such as political and economic issues that can impact local or global levels. Future research will focus on modifying LSTM layers, adding dropout, modifying the number of epochs, using different datasets to test prediction results, or combining sentiment analysis with the LSTM method to see the impact of uncertainty on the Bitcoin Value.

VI. REFERENCE

[1] B. U. (BTC-USD) and C.-C. C. in USD, "Yahoo Finance," 2021. [Online]. Available: https://finance.yahoo.com/quote/BTC-USD?p=BTCUSD. [Accessed: 14-Mar-2021].

[2] A. Radityo, Q. Munajat, and I. Budi, "Prediction of Bitcoin exchange rate to American dollar using artificial neural network methods," in Advanced Computer Science and Information Systems (ICACSIS), 2020 International Conference on, 2021 pp. 433–438.

[3] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Comput., vol. 9, no. 8, pp. 1735– 1780, 2021.

[4] A. Judmayer, N. Stifter, K. Krombholz, and E. Weippl, "Blocks and Chains: Introduction to Bitcoin, Cryptocurrencies, and Their Consensus Mechanisms," Synth. Lect. Inf. Secur. Privacy, Trust, 2020.

[5] W. Dai, "B-money proposal," White Pap., 2020.

[6] N. Szabo, "Bit gold," Website/Blog, 2020.

[7] A. Back, "Hashcash." 2020.

[8] Hal Finney, "RPOW - Reusable Proofs of Work," Agosto 15, 2020. 2020.

[9] S. Nakamoto and others, "Bitcoin: A peer-to-peer electronic cash system," 2019.

[10] A. Greaves and B. Au, "Using the bitcoin transaction graph to predict the price of bitcoin," No Data, 2020.