

Relationship Analysis COVID-19 Pandemic and cryptocurrency market with Machine Learning

Haohan Wu

School of Science Purdue University West Lafayette (47906), Indiana, United States

wu1669@purdue.edu

Abstract. The global economy receives a catastrophic blow due to the COVID-19 epidemic, with long-term pessimism shown towards the global market and exponential increasing expectation for looking for a reliable safe-haven asset. Now, it seems the possible alternatives to traditional currencies issued and backed by governments have primarily emerged - in the form of Bitcoin and Ethereum, as well as several other cryptocurrencies. Cryptocurrencies have been on the market for a long time and have been controversial in the international financial markets for their unique properties. Since the outbreak of COVID-19, the price of cryptocurrencies has seen an unprecedented increase. Whether the price increase of cryptocurrencies is linked to the COVID-19 outbreak is a mystery. This paper will focus on exploring this question through a linear regression machine learning model. The data in the U.S.A are used here. Our results show that the price of bitcoin is significantly related to the price of Ethereum. There is some correlation between covid-19 new cases incensement and the price of Bitcoin and Ethereum, indicating the legitimacy of predicting cryptocurrencies' price using covid-19 new cases incensement as a factor.

Keywords: Machine learning, cryptocurrencies, commodity, covid-19, Bitcoin, Ethereum, Linear Regression.

1. Introduction

The currency has various expressions in different periods as a commonly accepted value object. The continuously growing strength of the currency becomes one of the earliest indications of the alternation of a country's prosperity and impoverishment [1]. After the end of World War II, the global economy was in a recession, and international trade became more frequent than it ever was. The United States established a world monetary system with the U.S. dollar as the core. Since then, despite some negligible fluctuation, the U.S. dollar has become the world's clearing currency and reserve currency [2]. It seems this system could hardly be challenged until many events like the financial crisis of 2007–2008 emerged.

Later, the COVID-19 outbreak triggered the pandemic recession in 1870, with an alarming speed of shock to the global economy [3]. The covid-19 pandemic led to a severe shock to the global economy and resulted in a catastrophic decline in the global economy, financial, and commodity market. Based on the data from [4], COVID-19 inevitably spread worldwide rapidly, forcing the world to take several actions to stop the coronavirus at the high cost of GDP decline. In response to such an unprecedented crisis, the U.S. Federal Reserve cut interest rates continuously, lowering them to a practical lower limit of $0 \sim 0.25\%$ [5]. However, it did little to restore people's confidence in the U.S. financial markets since this crisis was not from the financial industry itself. Unlike the 2007 U.S. subprime crisis, for the financial crisis raised by an unprecedented pandemic, there was little room for risk prevention through quantitative easing monetary policy [6].

Central banks and governments have introduced accommodative monetary and fiscal policies in an effort to stimulate the economy and minimize the impact of the New Coronavirus on the economy. However, with the loose monetary and fiscal policies and the pessimism shown towards the markets, global funds are looking for safe-haven assets. In 2020, global hedge funds poured into Bitcoin, causing its price to soar three times [7]. Bitcoin's title of "digital gold" seems tended to be more and more acceptable. Cryptocurrencies like Bitcoin and Ethereum considered an alternative asset for storing value, are becoming increasingly popular among hedge funds. Bitcoin's unique design concept

creates virtual gold in the electronic information field, which has some qualities of gold and discards the disadvantages of gold [8]. It can potentially allow people to decide on their own wealth without being limited by banks or countries [9]. Since the launch of Bitcoin, the price of bitcoin has fluctuated, but the overall trend has been extremely rapid, rising nearly 120,000 times to date.

Regarding the risk characteristics of bitcoin, researchers have mainly focused on the bubble risk and the transaction risk of bitcoin. According to [9], the platform trading model and the zealous pursuit of the public make the price of bitcoin enormously fluctuate sometimes, and simultaneously the Bitcoin price also suffers from malicious attacks on the blockchain and blocked central bank intervention, so the risk of inventing Bitcoin is relatively high. Moreover, study 10 shows that when analyzing the risk of bitcoin at the level of bitcoin returns and volatility, it is found that the time series of bitcoin prices exhibit fat-tailed, long-memory, and asymmetric properties. Among those finds, bitcoin returns and volatility show a positive asymmetry between them, i.e., when returns are positive, volatility is usually greater.

Cryptocurrencies have the advantages of traditional metal currency and modern currencies backed by governments, making numerous investors and internal capital pay exponentially increasing attention to cryptocurrencies which has resulted in the dramatic growth of the cryptocurrencies market in recent years. Among all cryptocurrencies, Bitcoin and Ethereum are relatively predominant based on the past trade history of all cryptocurrencies, making their data reliable and comprehensive. The majority of scholars currently believe that Bitcoin and other cryptocurrencies are not currencies in the true sense of the word but are branches of speculative assets, which carry a great deal of risk. However, there is a lack of research on cryptocurrencies' yield, especially under the Covid-19 pandemic.

In order to grand a precise and concise result on exploring the cryptocurrency market during the covid-19 pandemic, this paper visualizes and analyzes the price of Bitcoin and that of Ethereum with Linear Regression. Also, since the majority of cryptocurrencies nowadays are traded in the United States and held by the owners from the U.S.A according to [12], this paper focuses on the data in the U.S.A. With the complement of the research shown in this paper, the conclusion that the price of bitcoin and the price of Ethereum show a significant correlation and also there is some correlation between covid-19 new cases incensement and the price of Bitcoin and Ethereum, indicating the legitimacy of predicting cryptocurrencies' price using covid-19 new cases incensement as a factor.

2. Method

This section introduces all the necessary methods and functions used in this paper, intending to reach the goal of the data analysis about the cryptocurrencies' price during the covid-19 pandemic. First, the simple comparison and the correlation among covid-19 cases increment, the Bitcoin, and Ethereum is presented. Then, the concentration is set at the commencement of the Linear regression machine learning method among those three.

2.1. Bitcoin vs. Ethereum

According to [13], though BTC and ETH cryptocurrencies are based on blockchain technology, they still have subtle nonnegligible differences, including their block time, consensus algorithm, supply, and so on, which are concisely analyzed in Table I below.

Table 1. Bitcoin vs. Ethereum

	Bitcoin	Ethereum
Creators	Satoshi Nakamoto	Vitalik Buterin, etc.
Consensus Algorithm	Proof-of-work	Proof-of-work moving to Proof-of-Stake
Block time	10 minutes on average	15 seconds on average
Launch date	January 2009	July 2015
Transaction Throughout	7 transactions per second	30 transactions per second
Supply	21 million BTC	Infinite supply

For such a sake, it is necessary to make a thoughtful comparison to test the relevance between Bitcoin and Ethereum before using machine learning methods.

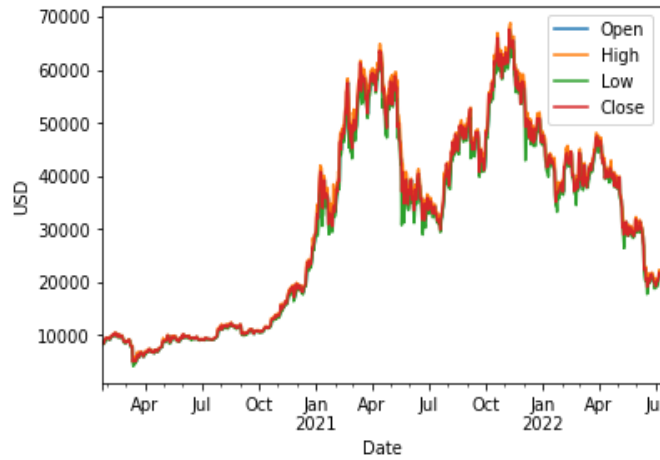


Figure 1. Bitcoin price history



Figure 2. Ethereum price history

Fig. 1 and 2 show the resulting graph generated by using Pandas [15] to pass the datasets of the history of the price of Bitcoin and that of Ethereum and then utilizing Matplotlib's plotting tool, respectively.

2.2. Correlation among cryptocurrencies and COVID-19 new cases incresement

$$\rho_{X,Y} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \quad (1)$$

As shown in (1), the Pearson correlation coefficient can be accessed through NumPy within the Python program. The input for this function in NumPy is typically a dataset in the matrix form, with the size of, where (columns) represents the total number of different variables; (rows) represents the total number of the set of samples for each variable; also, each column contains the values of a variable, and each row contains a single set of samples of random variables.

For given n variables, this function will return a square matrix with size, where such will contain the correlation coefficient between the variable and. Also, if it is equal to the correlation coefficient between and itself, all such results will be 1, implying that the diagonals of the matrix M will automatically be 1. Such logic can be interpreted as the (2).

$$M(i,j) = \begin{cases} \rho_{i,j} & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases} \quad (2)$$

In this case, only the price of Bitcoin and that of Ethereum are compared, which means the resulting matrix. M has a row size of 1 and a column size of 2. The element in M (1,2) results from

the Pearson correlation coefficient between the price of Bitcoin and that of Ethereum. Fig. 3 shows the results of applying this method in all columns of cryptocurrencies' datasets, including correlation coefficients of each variable of Bitcoin's and Ethereum's price.

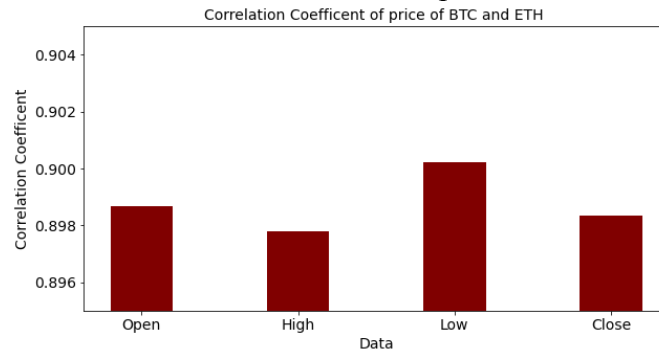


Figure 3. Correlation coefficient of price of BTC and ETH

Then, we apply the same technique of Pearson correlation coefficient to testify if covid-19 news-cases is correlate with the close price of two cryptocurrencies, with the produced Fig. 4.

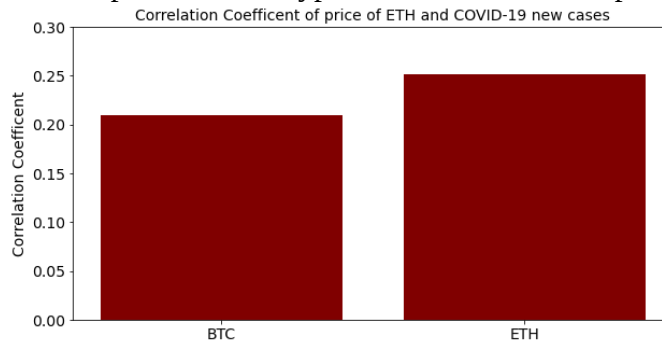


Figure 4. Correlation Coefficient of prices of cryptocurrencies and COVID-19 new cases

According to Fig. 3, with the Pearson correlation coefficient of the price of BTC and ETH, the correlation coefficient of the price of BTC and ETH is near 0.9, which can be considered significantly correlated [18]. However, in Fig. 4., the correlation coefficients between COVID-19 new cases and the price of cryptocurrencies are only roughly 0.2 and 0.25. According to [18], the correlation coefficient between 0 and 0.3 can indicate that there is existing a positive but extremely weak correlation between COVID-19 new cases and the price of cryptocurrencies.

2.3. Leaner regression prediction of cryptocurrency'price using its last 7 days data

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{d_i - f_i}{\sigma_i} \right)^2} \quad (3)$$

For any unbiased estimator, the Root-mean-square error (RMSE) defined in (3) can be used to represent the square root of the variance, which is the square root of the average of the squared differences between the actual value and the predicted values generating by ML model. In the current case, separate the original 'Close' data of Bitcoin and Ethereum into groups of any consecutive eight days from the first day. Then, we make the column of the latest day (8th day) in each group as the output y and the column of the first seven days in each group as the input x. Then x and y are separated for training and testing using the tool provided by Scikit learn. After training and testing, we use the RMSE to calculate the RMSE of the linear regression model.

$$R^2 = \left(1 - \frac{u}{v} \right) \quad (4)$$

As shown in (4), R^2 , the coefficient of determination is used to check the reliability of the linear regression model, where v is the total sum of squares and u is the residual sum of squares. The coefficient of determination has the best possible score of 1.0, and the coefficient of determination can also be negative, which means the model is arbitrarily wrong.

The resulting RMSE of the linear regression model used to predict the price of Bitcoin and Ethereum based on their last seven days' prices are shown in Figure 4. And Figure 5. According to the comparison between cryptocurrencies' prices and their average price, with the prediction score shown in Figure 6., it is clear that the prediction model is relatively accurate when predicting the price base on the past 7 days' prices.

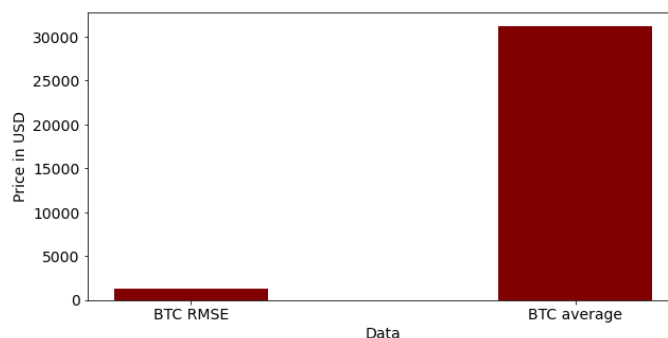


Figure 5. Bitcoin RMSE

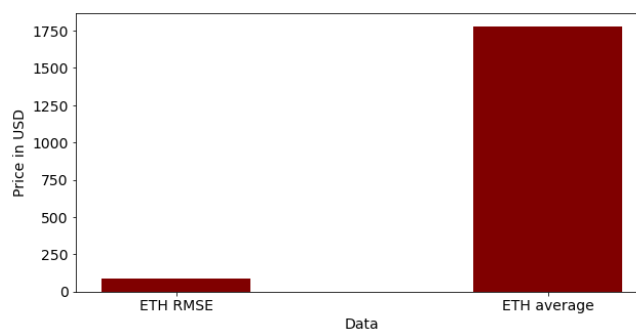


Figure 6. Thereum RMSE

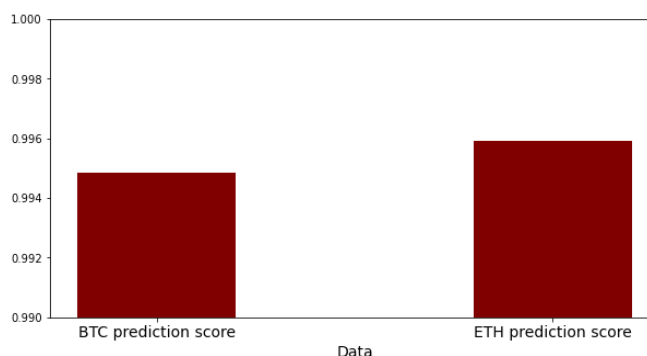


Figure 7. Prediction based on last 7-days price

2.4. Leaner regression prediction of cryptocurrency'price using last 7-days & 30-daysnew cases.

By applying the similar method in C., using the last 7-days covid new cases data and 30-days covid new cases respectively with the price history of Bitcoin and that of Ethereum.

Fig. 9 and 10 show that utilizing only the linear regression model can return the resulting RMSE of unsatisfactory. However, from both Fig. 9 and 10, the returned RMSE of the prediction is higher when more days of data were fed, which might imply a slight linear relationship between the covid-19 new cases and cryptocurrencies price.

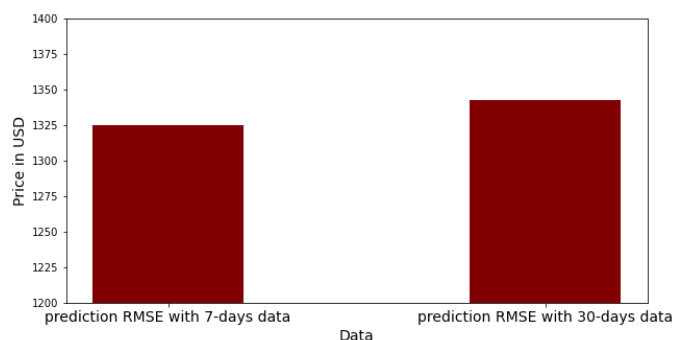


Figure 8. Ethereum price prediction with covid new cases

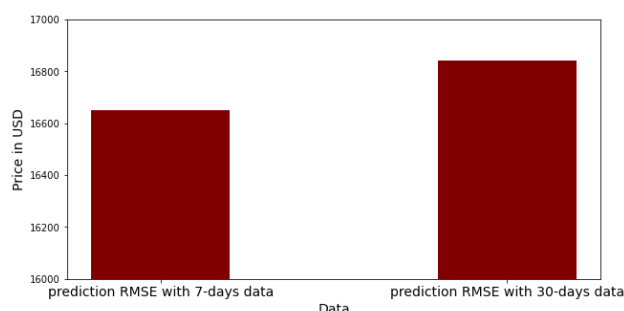


Figure 9. Bitcoin price prediction with covid new cases

3. Conclusion

In this paper, firstly, the correlation coefficient is introduced to explore the relationship between cryptocurrencies and covid-19 new cases in the U.S., which has result that indicates a weak correlation. Then the use of linear regressions machine learning model on predicting cryptocurrencies price gives a result there might be a slight linear relationship between the covid-19 new cases and cryptocurrencies price, so increasing the numbers of the label can efficiently improve the prediction result. In the future, it is worth trying different machine learning models that apply algorithms that have more accurate than linear regressions' gradient descent, which is potentially helpful in achieving a higher hit rate.

References

- [1] S. Hong-bing, "Currency War[M]," Beijing: CITIC Publishing Group (in Mandarin Chinese), 2017.
- [2] S. Arslanalp, B. Eichengreen & C. Simpson-Bell, "Dollar dominance and the rise of nontraditional reserve currencies," IMF Blog, Retrieved June 6, 2022.
- [3] "World Economic Outlook, October 2020: A long and difficult ascent," IMF, October 1 2022, Retrieved August 8, 2022.
- [4] Worldometer, Retrieved August 8, 2022, from <https://www.worldometers.info/coronavirus/>.
- [5] N. Zieminski & H. Timmons, "Federal Reserve statement - lowering federal funds rate to 0 to .25%," Reuters, Retrieved from <https://www.reuters.com/article/us-health-coronavirus-central-banks-fed-idUKKBN2121A0>.
- [6] J. Bernstein & E. Tedeschi, "Pandemic prices: Assessing inflation in the months and Years Ahead," The White House, November 30, 2021, Retrieved July 23, 2022.
- [7] M. DeMatteo, "Bitcoin price history 2009-2022 | nextadvisor with Time," Time, 2022, June 9, Retrieved August 4, 2022.
- [8] K. Kaponda, "Bitcoin the 'Digital Gold' and Its Regulatory Challenges," SSRN Working Paper, No. 3123531, 2018 Retrieved August 8, 2022.

- [9] N. Caldararo, “Bitcoin: Rube Goldberg Machine, Antique Throwback, Gigantic Distraction, Entertainment, Ripoff or New Money?”, *Interdisciplinary Description of Complex Systems*, 16, 427-445, 2018, Retrieved August 8, 2022.
- [10] T. Klein, P.T. Hien, and T. Walther, “Bitcoin Is Not the New Gold: A Comparison of Volatility, Correlation, and Portfolio Performance,” *International Review of Financial Analysis*, 59, 105-116, 2018 Retrieved August 8, 2022.
- [11] Kash, “Top 100 cryptocurrencies historical dataset,” Kaggle, July 10, 2022, Retrieved August 4, 2022, from <https://www.kaggle.com/datasets/kaushiksuresh147/top-10-cryptocurrencies-historical-dataset>.
- [12] “2021 cryptocurrency gains by country: Ethereum leads as gains skyrocket around the world,” *Chainalysis*, Retrieved August 10, 2022.
- [13] “Bitcoin vs. Ethereum: Key differences between BTC and eth,” *Cointelegraph*, December 28, 2021, Retrieved August 4, 2022.
- [14] NumPy, Retrieved August 9, 2022, from <https://numpy.org/>.
- [15] Pandas, Retrieved August 9, 2022, from <https://pandas.pydata.org/>.
- [16] Matplotlib, Retrieved August 9, 2022, from <https://matplotlib.org/>.
- [17] Scikit-learn, Retrieved August 9, 2022, from <https://scikit-learn.org/stable/>.
- [18] Z. Jaadi, “Everything you need to know about interpreting correlations,” *Medium*, October 16, 2019, Retrieved August 9, 2022, from <https://towardsdatascience.com/everything-you-need-to-know-about-interpreting-correlations-2c485841c0b8>