Forecasting NFTI Index Model: An Univariate Multivariable Regression Approach

Wending Wang[1,a,*,†], Lowan Li[2,†], Ruiqi Wu[3,†]

[1]College of Creative Studies, University of Santa Barbara, Santa Barbara, 93106, USA
[2]College of Arts and Science, NYU, New York, 10003, USA
[3]College of Engineering and Physical Sciences, University of Birmingham, Birmingham, B15 2TT, United Kingdom
a. wwd20001212@hotmail.com
*corresponding author
[†]These authors contributed equally to this work and should be considered co-first authors

Abstract: This paper proposed a forecasting model of the NFT index (NFTI) in the following year through repeating simulations applied to the univariate multivariable regression model. Beginning by choosing suitable predictors for the regression model which might affect NFTI through the best subset regression method, the team creates the multivariate regression model consisting of four dependent variables which are the log return of BTC, BTC/NFTI, NFTI spread, and NFTI volume/spread with respect to the independent variable log return of NFTI. Application of the regression model to the simulations based on the historical data generates 1000 pairs of data of log return of NFTI and log return of BTC as well as the corresponding predicted price of NFTI and BTC which are moderately correlated. Results are summarized in the cross-tabulation to quantitatively analyze the relationship between two variables which provides information for the investors about how they should formulate their own investment strategy. The results suggest that NFTI outperforms BTC unless NFTI crashes. Therefore, investment strategies can be made depending on the trend of BTC in the following year.

Keywords: forecasting model, univariate multivariable regression, NFT index

1.   Introduction

NFT Index is a digital asset index that is used to monitor the performance of 8 main tokens, which are Polygon (Matic), Enjin, Decetraland, Sand, Axie Infinity, Aavegotchi, Rarible, and Meme, in the NFT sector [1]. The index is weighted contingent on each component token's circulating supply, the total number of tokens that are actively available for trading, and utilization in the market. The NFT Index also represents the projects' developments in the decentralized financial market. These 8 tokens that are composed of the NFT Index are applied in different realms, some are used in games, some are operationalized in digital art production, and some are circulated on platforms that trade Non-fungible tokens (NFTs). Due to the unique identification codes that each NFT holds, people can trade NFTs as physical productions so that the ownership can be easily validated. For example, ownership over a digital workpiece can grant artists rights to rent them out without worrying about copyright issues.

Without a doubt, the number of wallets traded NFTs increased [2]. Even though the popularity of the NFT Index keeps increasing, there are few research papers about forecasting its future value which promotes making predictions on the NFT Index as the popularity of NFTs affects the values of those 8 tokens which directly reflect the NFT Index. The goal of our paper is to investigate the predicted values of the NFT Index through regression and simulation models. The next part of this paper will present the literature search, followed by a section elaborating on the process of collecting data. After the data collection, the data analysis will focus on describing the variable selection operation, regression model creation, and simulation procedure. At the end of the paper, conclusions about the forecasts of NFT prices would be presented.

2.   Literature Review

Many research papers are written by other scholars elaborate on associated topics and methods with our work. Firstly, Bitcoin price forecast using quantitative models is closely related to our research in a way of raising a similar question for forecasting but in a different subject and solved by using quantitative methods, which presents how to formulate arguments and demonstrate it empirically. In Multiple Regression Tests and Prediction for Ethereum Transaction Value, the writer employs similar methods such as the multiple regression model for the purpose of prediction. Prediction in Multiple Regression demonstrates the reason and use of multiple regression for predicting purposes. Moreover, in Is non-fungible token pricing driven by cryptocurrencies, the authors provide insights into the correlation between NFT pricing and other cryptocurrencies which provides ideas of selecting possible predictors associated with NFTI [3-6].

3.   Data Collection

All of the data is collected from Yahoo Finance. Since the earliest data for NFTI is from March 5th, 2021, all the historical data of each predictor are from that specific data to June 27th, 2022 which includes 70 sets of weekly data in total [7]. Since the log return of NFTI is used as the independent variable to compare the data for the week with the previous week which calculates the rate of return of NFTI investment, only 69 weeks' data are employed in the model. For the weekly data, each week consists of 5 sections – open, high, low, close and volume. "Open" represents the opening price of the security which is the price of the first trades upon the opening of an exchange for that week. "High" and "Low" stand for the highest and lowest price of the index during the week. "Close" refers to the closing price which is the last price at which a stock trades during the week. "Volume" means sum of daily amounts of trades during the week.

Based on the terms mentioned above, following data are selected as predictors:

1) Log return of BTC: ln (closing price of the week/ the closing price of previous week) which calculates the rate of return of BTC investment [8].

2) BTC/NFTI: Ratio of BTC and NFTI based on the closing price of the previous week data [7,8].

3) NFTI spread: Closing price - Opening price of previous week data [7].

4) NFTI Volume/ NFTI Spread: ratio of two variables based on the previous week's data [7].

5) NFTI Volume/(close-open): ratio of NFTI volume and the difference between previous week's closing price and opening price [7].

6) Global X cybersecurity ETF (BUG) weekly index [9].

7) Ethereum (ETH) weekly index [10].

8) KBW Nasdaq Bank (BKX) weekly index: three indexes are using the same week's data as NFTI which are selected to see the correlation of NFTI to the global cybersecurity market, performance of the digital asset Ethereum and performance of the leading banks and thrifts in U.S [11].

9) ETH/NFTI: Ratio of ETH and NFTI based on the closing price of the previous week data [7,10].

10) Log return of ETH: to see the rate of return of ETH.

11) DeFi Pulse weekly Index: to examine correlation between NFTI and tokens' performance within the Decentralized Finance industry through using the same week's data [12].

12) log return of Defi: to see the rate of return of Defi [12].

4. Data Analysis

4.1. Selecting Predictors

Using the best subset selection method and taking the adjusted R square as the standard, the team chooses the best model from all the models that can be composed of 0 to 12 factors mentioned in the previous section, moreover, balancing the interpretability and fitting degree is taken into consideration. Eventually, there are 4 suitable variables which are the log return of BTC, BTC/NFTI, NFTI spread, and NFTI volume/ NFTI Spread.

4.2. Multivariate Regression Model

Firstly, finding the correlation between the log return of NFTI and four prediction variables. Applying Multivariate Regression on the independent variable log return of NFTI and dependent variables BTC/NFTI, NFTI Spread, NFTI Volume/ NFTI Spread, and Log return of BTC, the formula shows as following:

Log return of NFTI = - 0.1048 + 1.247 Log return of BTC+0.003687 BTC/NFTI
(SE coefficient)   (0.0349) (0.177)   (0.000895)
            + 0.000013 NFTI Spread - 0.000061 NFTI Volume/ NFTI Spread
                 (0.000006)   (0.000024)

Along with the Regression equation, all predictors have p-values less than 2% except the predictor NFTI spread has a p-value that equals 5.6% which is still close to 5%. The regression has adjusted =50.35% indicating that 50.35% of the variations in the value of log return of NFTI are explained by the dependent variables. And the regression also passes the hypothesis test. Hence, the overall predictors are statistically significant and the regression is meaningful for predicting the log return of NFTI since it is reasonable to forecast NFTI based on Bitcoin and the trend of NFTI itself in the previous week's data.

4.3. Simulations for Predictions

Next, the following year's data would be generated through simulations. Now the dataset includes 69 sets of data in total which are the historical data for each variable. 52 sets of weekly data are randomly chosen from the dataset to generate the next year's prediction based on the historical data. Since each set of data has its own corresponding value for each dependent variable, plugging each data into the regression model, the result gets the predicted value of log return of NFTI as well as the corresponding value of log return of BTC for each week. By adding up the predicted log return of NFTI and the corresponding data of the log return of BTC for each week, the results of total log return of the NFTI and log return of BTC for the next year are shown. The predicted NFTI and BTC for the next year are displayed by applying log returns of NFTI and BTC to the most recent weekly data of NFTI and BTC which are $536.17 and $19297.08 on June 27th, 2022. This completes one simulation.

By performing the above simulation 1000 times, 1000 sets of the next year's Predicted Annual log return of NFTI and Predicted annual log return of BTC are obtained. The following figures present the results of the simulations:
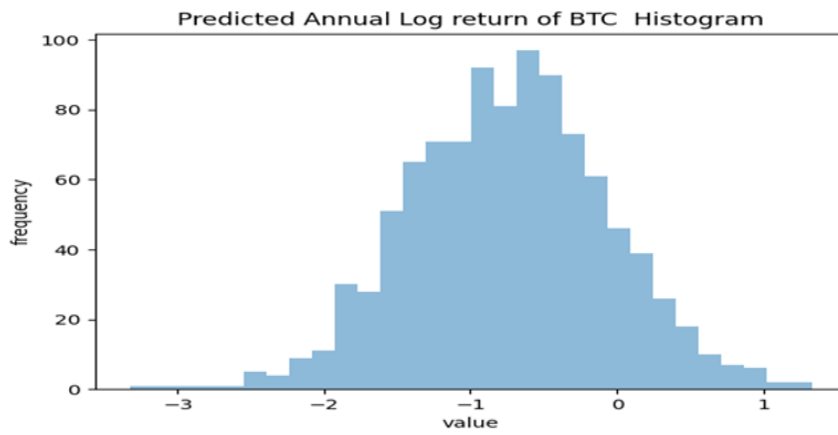
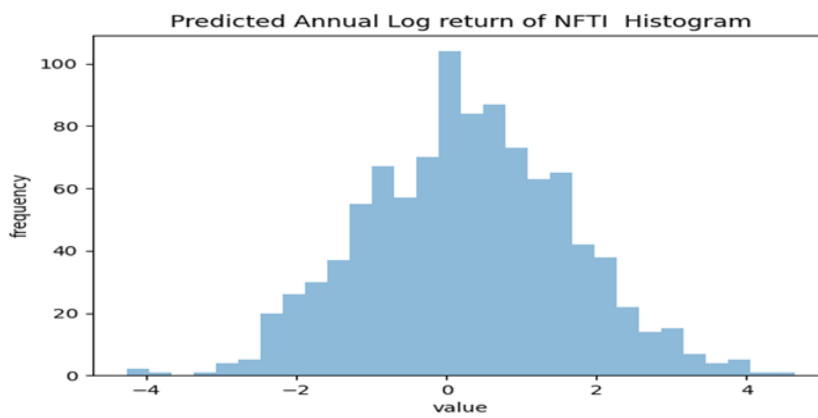Figure 1: Distribution of predicted annual log return of BTC.

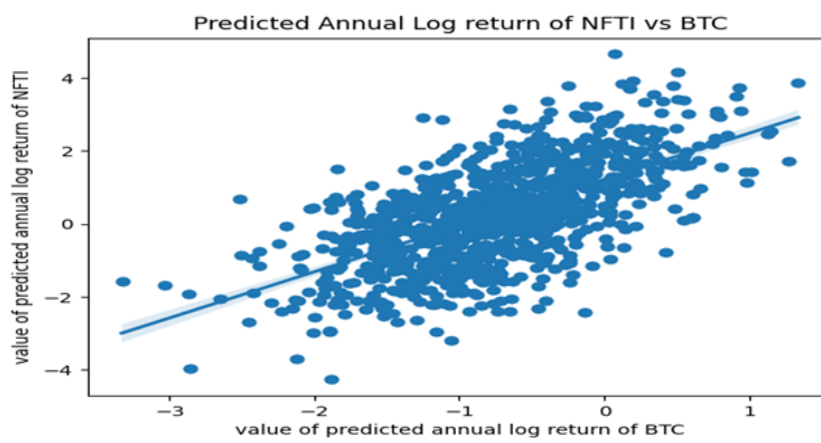Figure 2: Distribution of predicted annual log return of NFTI.

Figure 3: Regression of Predicted annual log return of NFTI on that of BTC.
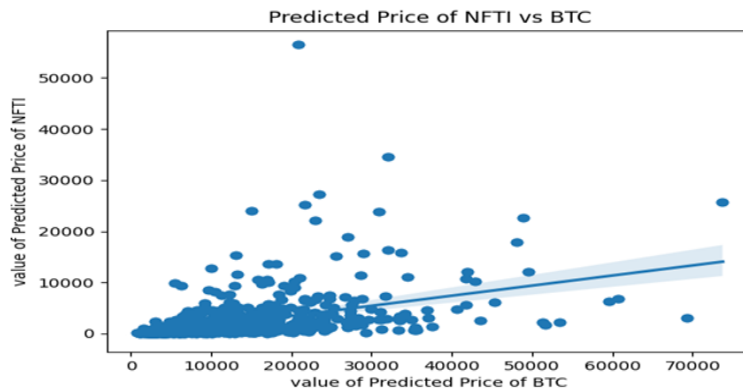
Figure 4: Regression of Predicted annual return of NFTI on that of BTC.

From figures 1 and 2, the simulated distribution results of log return of BTC and NFTI are leptokurtic and fat-tail, with mean values of -0.738 and 0.285 respectively. Besides, the standard deviation of log return of NFTI is 1.36. From figure 3 and figure 4, the conclusion includes that the log return of BTC and NFTI is moderately correlated with adjusted =40% but there is quite a bit of variation around the line. In terms of the price of BTC and NFTI, the linear correlation has adjusted =22%.

Moreover, the cross-tabulation below (see Table 1) presents the relationship between forecasting NFTI and BTC value in the next year:

Table 1: Distribution of average price (return) of NFTI and BTC.

NFTI ($) BTC ($)	[0,180]	(180,2672]	(2672)
[0,4699]	Probability: 7% NFTI: 92(-83%) BTC: 3162(-84%)	Probability: 9.6% NFTI: 526(-2%) BTC: 3717(-81%)	NA
(4698,17756]	Probability: 9.6% NFTI: 114(-79%) BTC: 7848(-59%)	Probability: 49.6% NFTI: 903(+68%) BTC: 9805(-49%)	Probability: 7.6% NFTI: 5430(+913%) BTC: 12548(-35%)
(17756,)	NA	Probability: 7.6% NFTI: 1464(+173%) BTC: 24905(+29%)	Probability: 9% NFTI: 8486(+1483%) BTC: 28708(+49%)

From the above cross-tabulation Table 1, it displays the pairs of different prices and returns of Bitcoin and NFTI and their probability of occurrence respectively. In each block, the probability is calculated by the percentage of data of predicted NFTI and BTC lying within the given range. The mean of predicted NFTI and BTC to the corresponding range is shown below the probability section. Percentages of increase and decrease with respect to the initial value of NFTI $536.17 and BTC $19297.08 are shown as well. The percentage of price changes in NFTI and BTC value and probability of occurrence provides statistics of an overall prediction for the following year.

5.   Conclusion

This paper established a univariate multivariable regression model with log return of BTC, BTC/NFTI, NFTI spread, and NFTI volume/spread as factors to predict the log return of NFTI. The simulations of 1000 possible predictions of NFTI and BTC prices in the following year are produced by applying historical data to the regression model. The results are summarized in an instructive cross-tabulation for future investment.

Possible outcomes for next year's price changes of NFTI and BTC are simulated and suggestions for investments are provided due to various conditions. If both NFTI and Bitcoin crash next year: i) there is a 7% chance of occurrence that the prices fall into the range of [$0,$180] and [$0,$4698] respectively, then NFTI and BTC have almost the same investment value; ii) there is 9.6% chance of occurrence with prices within the range of [$0,$180] and [$4698,$17756] respectively, then BTC will lose less money than NFTI. Possible investment strategies such as short sale NFT tokens and long sale BTC could be suitable under this circumstance. And there is a 9.6% probability that BTC crashes and NFTI maintains at about the same value next year, and two assets will be in the range ($4698,$17756] and ($180,$2672]. Under all the other circumstances, NFTI will outperform BTC which indicates that NFTI could be a better investment. In general, the results conclude that NFTI will have a higher investment value unless the price of BTC crashes. Investors can create their own strategies to invest in NFT based on the trend of BTC in the future. For investors who do not have enough information to predict the price changes of NFTI and BTC for the next year, the results suggest that next year's NFTI and BTC value will lie in the range ($180, $2672] and ($4698, $17756] respectively with the probability of occurrence close to 50%. In this case, NFTI is a better investment asset than BTC. For those who believe a crash of BTC in the next year is less likely, they can focus on the bottom row of the cross-tabulation which suggests that NFTI has a much higher return than BTC.

Acknowledgment

References

[1] CoinMarketCap. (2022) NFT Index Price Today, NFTI to USD Live, Marketcap and Chart. https://coinmarketcap.com/currencies/nft-index/.
[2] Dune. (2022) NFT Market Overview. https://dune.com/thomas_m/NFT-stats.
[3] Bernardi, D., Bertelli, R. (2021) Bitcoin Price Forecast Using Quantitative Models. SSRN.
[4] Wang, R. Multiple Regression Tests And Prediction for Ethereum Transaction Value. SSRN.
[5] Osborne, J.W. (2000) Prediction in Multiple Regression. Practical Assessment, Research, and Evaluation: Vol. 7, Article 2.
[6] Dowling, M. (2021) Is Non-Fungible Token Pricing Driven by Cryptocurrencies? SSRN.
[7] Yahoo! Finance. (2022) NFT Index USD (NFTI-USD) Price, Value, News & History. https://finance.yahoo.com/quote/NFTI-USD/.
[8] Yahoo! Finance. (2022) Bitcoin USD (BTC-USD) Price, Value, News & History. https://finance.yahoo.com/quote/BTC-USD/.
[9] Yahoo! Finance. (2022) Global X Cybersecurity ETF (Bug) Stock Price, News, Quote & History. https://finance.yahoo.com/quote/BUG?p=BUG&.tsrc=fin-srch.
[10] Yahoo! Finance. (2022) Ethereum USD (ETH-USD) Price, Value, News & History. https://finance.yahoo.com/quote/ETH-USD?p=ETH-USD&.tsrc=fin-srch.
[11] Yahoo! Finance. (2022) KBW NASDAQ Bank Index (^BKX) Charts, Data & News. https://finance.yahoo.com/quote/^BKX?p=%5EBKX&.tsrc=fin-srch.
[12] Yahoo! Finance. (2022) Defi Pulse Index USD (DPI-USD) Price History & Historical Data. https://finance.yahoo.com/quote/DPI-USD/history/.