Bitcoin Price Prediction Based on Machine Learning and Granger Causality Test

Mengyu Hao [1,*, †], Feiyang Su [2, †], Kaifei Wang [3, †], Xiaoqi Zheng [4, †]

[1] *School of Finance & Management, Shanghai University of International Business and Economics, Shanghai, China, 201620*
[2] *College of Sciences, Shanghai University, Shanghai, China,200444*
[3] *School of Arts and Sciences, Johns Hopkins University, Washington, DC, United States of America, 20036*
[4] *School of Computing and Artificial Intelligence, Southwestern University of Finance and Economics, Chengdu, China, 611130*
[*] *Corresponding author. Email: 19063022@suibe.edu.cn*
[†] *These authors contributed equally.*

ABSTRACT

Recently, more and more investors have seen the huge profits that the digital currency market can bring, and Bitcoin price predictions are becoming more valuable both academically and in terms of business value. In this paper, we use the daily price of bitcoin from September 12, 2016, to September 10, 2021. Data pre-processing includes moving average (MA) and BIAS. To find out the causality relationship between two factors, we use Granger causality test. Then we predict bitcoin price with Support Vector Machine (SVM) based on sliding window from machine learning methods and Autoregressive Integrated Moving Average (ARMA) method from statistical methods. The results show that there is causality relationship between gold and bitcoin. Besides, by comparing the Mean Squared errors (MSE) of 7-day-model, 14-day-model and ARMA model, we find that the ARMA model outperform the others, which reminds the investors to focus more on this model.

Keywords: *Bitcoin prediction, Granger Test, ARMA, MSE*

1. INTRODUCTION

Bitcoin is a cryptocurrency that can be used for point-to-point cash payments or investment activities around the world. The scheme was initially suggested in 2008 by Wright [1] and became operational in January 2009. Bitcoin is regarded as a financial asset. As a digital currency, bitcoin often has fierce price fluctuations. So far, some studies have confirmed that bitcoin is a speculative bubble rather than a long-term investment (Bouoiyour & Selmi [2]). However, some scholars, such as Kondor, Pósfai and Csabai [3], used the complex network framework to study bitcoin price, finding that the network characteristics of bitcoin price will fluctuate over time. In recent years, more and more investors see the huge benefits that the digital money market can bring and choose to enter the market. With this, the research on bitcoin price prediction is more and more worthy of attention.

Bitcoin price forecast can refer to the stock market forecast. In this regard, many studies have been discussed and practiced. Among them are purely mathematical methods, such as Brownian Motion Model (Yang & Aldous [4]), which simulates price fluctuations by simulating prices as Brownian motion. Some scholars also make predictions through machine learning methods, such as using SVM (Karasu, Altan, Saraç & Hacioğlu [5]). In recent years, more and more scholars have applied deep learning to bitcoin price prediction, among which the representative algorithms are Deep Neural Network (DNN) and Long Short-Term Memory (LSTM) (Chen, Li & Sun [6]). At the same time, researchers (Hiemstra & Jones [7]) applied Granger causality test to the stock market to find the main indicators of stocks. While bitcoin share considerable common features with stocks, the Granger Causality Test can also be applied to bitcoin and as a result leave us a broader horizon.

Due to the special point-to-point trading system of bitcoin, bitcoin trading can take place in the Bitcoin Exchange. Bitcoin does not need to be issued through central bank institutions, which allows people to trade bitcoin in different currencies. Bitcoin is also known as

Y. Jiang et al. (Eds.): ICEDBC 2022, AEBMR 225, pp. 342–348, 2022.
https://doi.org/10.2991/978-94-6463-036-7_51

digital gold. As an internationally recognized global currency, the price of gold is one of the basic factors affecting the price of bitcoin (Bouoiyour & Selmi [8]). Some researchers (Al-Yahyaee, Mensi & Yoon [9]) have studied the asymmetry, thick tail, long-term and short-term characteristics of price fluctuations, and the thick tail, autocorrelation and asymmetry of cryptocurrency, making it closer to general financial products such as futures or gold.

This paper aims to utilize the correlation between bitcoin and gold to make short-term predictions of bitcoin price with SVM algorithm based on the gold prices and bitcoin prices data from 2016 to 2021. We discussed bitcoin price, gold price and four other related features, namely, the moving average (MA) of bitcoin price, the BIAS of bitcoin price, the variation of bitcoin price daily and the variation of gold price daily. The above features are applied to the prediction of special currency price by SVM model and ARMA model. We use SVM model to compare the bitcoin price prediction based on past 7-day and past 14-day data, in which the mean squared error (MSE) value is considered as the assessment criterion to judge the performance. Meanwhile, the prediction results of ARMA model are presented with backtesting. The results show that the prediction performance of ARMA model is better than SVM model. On the other hand, in order to verify the effectiveness of bitcoin price prediction, we combine the machine learning method, named the SVM algorithm, with the Granger causality test to explore the effect of adding gold price on the performance of the original model. The results of Granger causality test show that it is feasible to take the gold price into account and reduce the mean square error to improve the performance of the model.

The rest of this paper is organized as follows. Section 2 shows the data and methods. Section 3 presents the results and Section 4 concludes the paper.

2. DATA AND METHODS

2.1. Data

Inspired by the Mathematical Contest in Modeling held by COMAP in February 2022, we obtained our data from the official website of the modeling contest (http://www.comapmath.com/MCMICM/index.html). The original datasets we downloaded are constituted of daily prices of gold and Bitcoin (at 0:00 every day) from September 11, 2016, to September 10, 2021, respectively.

Gold can only be traded during workdays, while Bitcoin can be traded in any day the investor favors. Although two datasets share nearly same time period, the amounts of data are different. As gold price is an indicator in this paper, which will be discussed later, we make an assumption that gold prices remain the same during none-work days. Thus, we use the gold price of the most recent workday as the gold price of the next several non-workdays.

The initial value of the Bitcoin price is much lower than the gold price, while the maximum value is about 30 times higher than the maximum value of the gold price. The variance of the Bitcoin price is nearly 70 times higher than the variance of the gold price. The standard deviation of the Bitcoin price is also much higher than the standard deviation of the gold price. All of these statistics reflect the greater volatility of the bitcoin price compared to gold. Over the five-year period encompassed by our data, Bitcoin has experienced significant gains and losses, while the gold price has remained at a relatively flat level.

To give a more comprehensive view of Bitcoin behavior, we add two more features (MA7, BIAS) as depiction of markets based on former work of Wu Xing et al [10] and two other features for further discussion about causality test. The definitions of features are defined in Table 2.

Table 1. Statistics of data

	Max	Min	Mean	Std.dev
Bitcoin	63554.44	594.08	12212.42	14041.27
Gold	2067.15	1125.70	1463.72	249.31

Table 2. definitions of features

Name of feature	Definition
Moving average	$MA = \frac{p_{n-7}+p_{n-6}+\cdots+p_{n-1}}{7}$, where p_n stands for the Bitcoin price of each day.
BIAS	$BIAS = \frac{p - MA}{MA} \cdot 100$
Δp_B	$\Delta p_B = p_{B_n} - p_{B_{n-1}}$
Δp_G	$\Delta p_G = p_{G_n} - p_{G_{n-1}}$

Table 3. correlation coefficients of features

	Bitcoin	gold	Δp_B	Δp_G	MA7	BIAS
Bitcoin	1					
gold	0.65	1				
Δp_B	0.054	0.031	1			
Δp_G	0.0056	0.031	0.02	1		
MA7	0.99	0.65	-0.0048	-0.0071	1	
BIAS	0.04	0.019	0.36	0.042	-0.039	1

According to Table 3, the correlation coefficient between MA7 and Bitcoin price is the highest, followed by that between Bitcoin price and gold price. It can be easily explained that Bitcoin price and MA7 is highly correlated due to MA7 is calculated using only Bitcoin price. However, we notice the high correlation coefficient between gold price and Bitcoin price, which will be discussed later based on Granger causality test.

2.2 Methods

In this paper, we conducted two main models to predict the price of Bitcoin. To reveal the relation between gold price and Bitcoin price, we conducted Granger Causality test. Then we apply SVM and ARMA to make price prediction.

2.2.1 Granger Causality Test

Granger causality test was first proposed by Sir Clive Granger in 1969 and is widely used in different fields. By applying Granger causality test, we can find the causality relationship between two factors. In 1994, C Hiemstra and JD Jones [11] applied Granger Causality Test to stock markets to find leading indicator of stocks. More recently, in 2019, AK Tiwari et al [12] pointed out that there is dependence between global gold market and emerging market. While Bitcoin share considerable common features with stocks and is a part of emerging market, we suppose the Granger Causality Test can also be applied to Bitcoin price and gold price.

The main process can be described as follow.[13] Two factors, namely X and Y, have causal relationship, if the prediction error of Y is significantly smaller than the error of X using only past information without Y.

$$\sigma^2(X|U) < \sigma^2(X|U - Y) \tag{1}$$

Where U stands for all past information.

$$X(t) = \sum_{j=1}^{p} \alpha_{11,j} X(t-j) + \sum_{j=1}^{p} \alpha_{12.j} Y(t-j) + E_1(t) \tag{2}$$

$$Y(t) = \sum_{j=1}^{p} \alpha_{21.j} X(t-j) + \sum_{j=1}^{p} \alpha_{22,j} Y(t-j) + E_2(t) \tag{3}$$

Equation (2) predicts current value of X based on values of X and Y with a time lag p. Equation (3) predicts the current value of Y based on values of X and Y with a time lag of p. Es are errors. Then apply F test and Chi-squared test to test if Y is C-causing X. By the theory of hypothesis testing, if p value is less than 0.05, we say the hypothesis is tested.

$$\frac{T(RRS_1 - RRS_2)}{RRS_2} \ follows\ \chi^2(p) \tag{4}$$

where

$$RRS_1 = \sum_{t=1}^{T} E_X(t)^2 \tag{5}$$

$$RRS_2 = \sum_{t=1}^{T} E_{X,Y}(t)^2 \tag{6}$$

2.2.2 Support Vector Machine

Machine learning methods are widely used in prediction in different fields. By J Gao et al [14] work, SVM and other machine learning methods can be applied to the prediction part of stock selection strategy. Because Bitcoin is a new form of asset, it shares some common features with stacks. We attempt to apply SVM algorithm to Bitcoin prediction.

The SVM algorithm is proposed to make classification by maximizing the interval between samples from different categories. And it can also be applied to make prediction. In a classification problem, given input data and a learning objective: $X = \{X_1, \cdots, X_N\}$, $y = \{y_1, \cdots, y_N\}$, where each sample of the input data contains multiple features and thus constitute a feature space:$X_i = [x_1, \cdots, x_n] \in U$. And the learning objective is a binary variable representing a negative class and positive class.

If the feature space where the input data is located has a hyperplane as the decision boundary, the learning targets are separated into positive and negative classes, and the distance from the point to the plane of any sample is greater than or equal to 1:

$$decision\ boundary: \omega^T X + b = 0 \tag{7}$$

$$point\ to\ plane\ distance: y_i(\omega^T X_i + b) \geq 1 \tag{8}$$

Then the classification problem is said to be linearly separable, and the parameters ω, b are the normal vector and the intercept of the hyperplane, respectively.

All samples above the upper interval boundary belong to the positive class, and all samples below the lower interval boundary belong to the negative class. The distance between the two interval boundaries is defined as the margin $d = \frac{2}{||\omega||}$, and the positive and negative class samples located on the interval boundaries are the support vectors.

2.2.3 Sliding Window

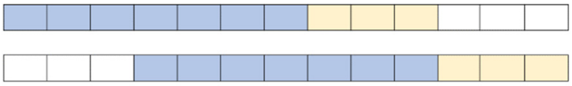

Figure 1 Sliding window

Due to the volatility of Bitcoin, the error will be huge if we use data of a considerably long time period to predict. We then attempt to use data of shorter period of time to predict. To achieve this, we use sliding windows to make prediction based on small amount of data.

In this paper, we conduct SVM with sliding window.

To make comparison, we also introduce other algorithm that has a "sliding" progress.

2.2.4 ARMA model

ARMA (Autoregressive moving average model) is an important tool for time series research. It is combined with AR model and MA model. ARMA is widely used for prediction of sales volume and market size with seasonal variation.

The data sequence formed by the predictors over time is regarded as a random sequence, and the dependencies of this group of random variables reflect the continuity of the original data in time. On the one hand, the influence of the influencing factors, on the other hand, has its own changing law, assuming that the influencing factors are$x_1, x_2, \cdots, x_k$, by regression analysis,

$$Y_t = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p + Z, \tag{9}$$

where Y is the observed value of the predicted object and Z is the error. As the prediction object Y_t is affected by its own changes, its pattern can be represented by the following for

$$Y_t = \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_p Y_{t-p} + Z_t. \tag{10}$$

The error term has dependencies in different periods, which is represented by the following formula,

$$Z_t = \epsilon_t + \alpha_1 \epsilon_{t-1} + \alpha_2 \epsilon_{t-2} + \cdots + \alpha_q \epsilon_{t-q}. \tag{11}$$

From this, the expression of ARMA model is obtained:

$$\begin{aligned} Y_t &= \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_p Y_{t-p} \\ &+ \epsilon_t + \alpha_1 \epsilon_{t-1} + \alpha_2 \epsilon_{t-2} + \cdots + \alpha_q \epsilon_{t-q}. \end{aligned} \tag{12}$$

3. RESULTS AND DISCUSSION

This paper selects data of Bitcoin Prices (at 0:00 every day) between 9/11/2016 to 9/10/2021. In order to deal with the data that we obtained; we make an important assumption. We assume that gold price does not change during weekends and holidays in order to fill the null data. From the figures of gold price and Bitcoin price and their daily variance, it can be seen that a sudden rise or drop of gold price may result in a rise or drop of Bitcoin price in the near future. After doing Granger Causality Test to daily variance of gold price and Bitcoin price, it can be determined that the exact length of the time lag between gold price and Bitcoin price is 7. Future daily variance of gold price is added as a new feature of the original data. With a comprehensive analysis, five features for Bitcoin price prediction are found: the variance of gold price 3 days later, the variance of Bitcoin price, moving average, exponential moving average and BIAS.

In the paper, we predict Bitcoin price in a very near future of nearly 3 days to get more accurate results. Predictions of Bitcoin prices are based on sliding windows. Firstly, we use last 7 days' data to predict Bitcoin price with SVM model. Secondly, we use last 14 days' data to predict Bitcoin price with SVM model. Finally, ARMA model is constructed to make prediction on Bitcoin price. After back testing, It's indicated that ARMA model is very good for prediction of Bitcoin's price. By comparing with their Mean Squared errors, we can conclude that ARMA is the best model for prediction because of the smallest Mean Squared Errors it has.

3.1. Experiments with real data

3.1.1. Granger Causality Test Between Gold and Bitcoin

Bitcoin is more like gold in an increasingly favorable macroeconomic environment.

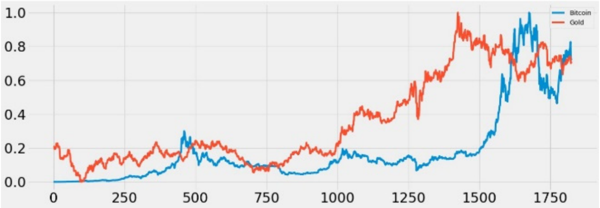

Figure 2 gold price and Bitcoin price

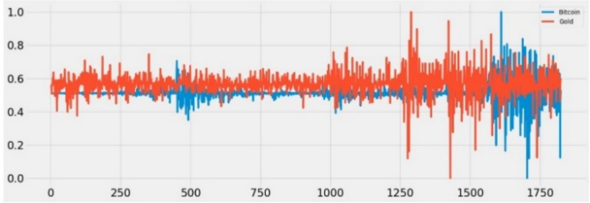

Figure 3 daily variance of gold price and Bitcoin price

From Figure 2 and Figure 3, we can see the sharp rise or drop of Bitcoin price have some connection with gold price: a sudden drop of gold price might indicate a drop of Bitcoin price in the near future. There is a time lag if we use gold price as an indicator of Bitcoin price. By applying Granger Causality Test to daily variance of gold price and Bitcoin price $(p_n - p_{n-1})$, we can find the exact length of the time lag between Bitcoin price and gold price.

Table 4. P-Values

Time lags	1	2	3	4
F test	p=0.3394	p=0.6429	p=0.7690	p=0.5885
χ^2 test	p=0.3389	p=0.6421	p=0.7679	p=0.5860
likelihood ratio test	p=0.3390	p=0.6421	p=0.7680	p=0.5864
parameter F test	p=0.3394	p=0.6429	p=0.7690	p=0.5885
Time lags	5	6	7	
F test	p=0.6982	p=0.4930	p=0.0459	
χ^2 test	p=0.6954	p=0.4880	p=0.0435	
likelihood ratio test	p=0.6958	p=0.4980	p=0.0444	
parameter F test	p=0.6982	p=0.4930	p=0.0459	

Thus, when number of lags is equal to 7, $p < 0.05$, which means that the hypothesis of that there is causality relationship between gold price and Bitcoin price holds. So, we add future daily variance of gold price as a new feature of the original data.

3.1.2. Prediction Using SVM based on sliding window

It is difficult to predict stock markets for long term, so is crypto markets. Using limited data to predict long-term behavior of Bitcoin price will be inaccurate. In order to provide more accurate prediction, we only predict Bitcoin price in a very near future, namely 3 days.

At the same time, we notice that the long-term behavior in the past of Bitcoin price might become meaningless because the price sometimes soars and drops violently. Therefore, we use data that are close to the exact date we are looking to predict at.

To evaluate the performance of our model, we calculate Mean Squared Error (MSE) with EQUATION. (13).

$$MSE = \frac{\sum i{=}1(y_i - \hat{y_i})^2}{n} \tag{13}$$

where y_i stands for real price and $\hat{y}_i$ stands for the predicted price.

Using Python and SVM regressor from Sklearn, we conduct experiments below with parameters shown in Table 5. In the SVM, we set the kernel function to a polynomial kernel with a degree of 2. In addition, we set the parameter of C to 100 and Epsilon to 0.1.

Tables 5. parameters of SVM

parameter	kernel	degree	C	Epsilon
value	poly	2	100	0.1

3.1.3. Using last 7 days' data to predict (7-day-model)

In this part, we are using data from 7 days closest to the current date to predict Bitcoin price. The predicted results are shown in the Figure 3. We can see that the curve of Bitcoin prices deviates a little from the curve of predicted price. And the Mean Squared Error of this model is 2530120.8.

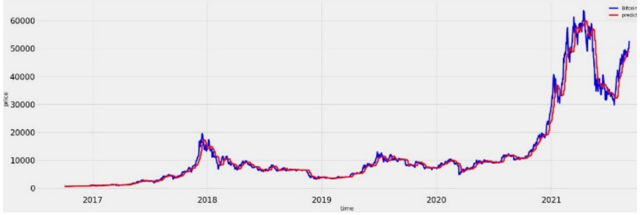

Figure 4 Using 7 days' data to predict

3.1.4. Using last 14 days' data to predict (14-day-model)

In this part, we are using data from 14 days closest to the current date to predict Bitcoin price. The predicted results are shown in the Figure 4. We can see that the curve of Bitcoin prices doesn't deviate a lot from the curve of predicted price, but more than the 7-day-model. And the Mean Squared Error of this model is 4500897.3.

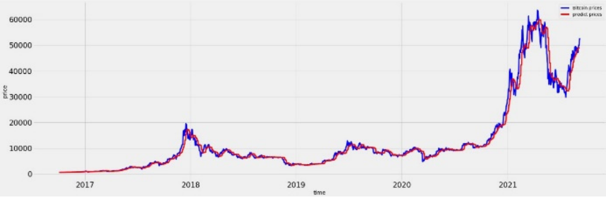

Figure 5 Using 14 days' data to predict

3.2. More Attempts using ARMA model

3.2.1. Using ARMA model to predict

The predicted results are shown in Figure 5. We can see that the curve of Bitcoin prices and the curve of predicted prices almost coincide with each other. The Mean Squared Error is 669848.2.

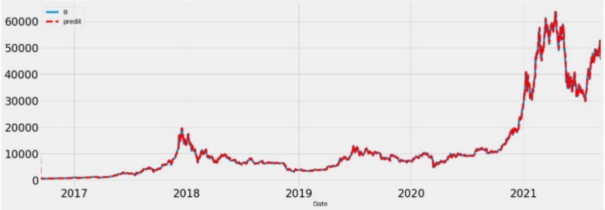

Figure 6 Using ARMA to predict

3.2.2. Back testing with ARMA model

The back testing results are shown in Figure 6 and Figure 7. Figure 6 shows the return of each trade, and figure 7 shows the cumulative return of each trade. In this back test interval, the maximum return of each trade reaches 10000, while the minimum return of each trade is about -3500. And the maximum cumulative return is above 140000. As a result, the ARMA model is good for prediction of Bitcoin price.

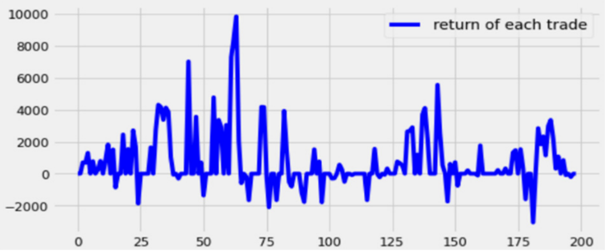

Figure 7 Return of each trade

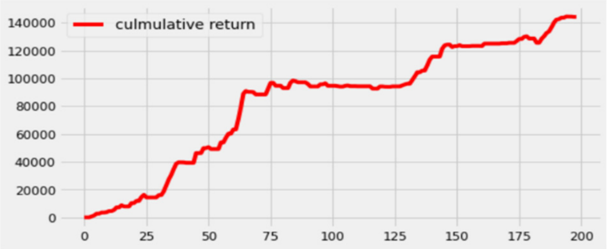

Figure 8 Cumulative return of each trade

3.3. Comparison with each kind of model

To make comparison with each kind of model and find the best one, we compute all the Mean Squared Errors. The results are shown in Table 6.

Table 6. Mean Squared Errors

model	7-day-model	14-day-model	ARMA
MSE	2530120.8	4500897.3	669848.2

The MSE of 7-day-model is much smaller than the MSE of 14-day model, which indicates that using last 7 days' data to predict Bitcoin prices has a better result than using last 14 days' data to predict Bitcoin prices. The MSE of ARMA model is the smallest among three models, showing that it is the best model to make prediction.

By observing Figure 3-6, while applying machine learning method – SVM to Bitcoin price prediction, there is an obvious time lag between the real data and predicted data. The peaks of predicted data appears later than the peaks of real data. The sharp rise or drop of Bitcoin price is hard to predict timely and precisely. However, our work can predict the trend of Bitcoin price although with time lags. In this way, it can still help crypto investors to make a wiser judgement and earn more profits.

4. CONCLUSION

Nowadays, the price prediction of virtual currencies is receiving more and more attention from scholars and business companies.
To the best of our knowledge, this paper makes the follo wing contributions to the literature, first, this paper broa dens the investigations regarding the Bitcoin market; sec ond, this paper certifies the effectiveness of traditional f orecasting models in the emerging Bitcoin market. In this paper, to give a more comprehensive view of Bitcoin behaviours, we add two more features based on former works, and the empirical results in this paper are summarized as follows. First, the causality test results show that there is causality relationship between gold and bitcoin markets. Second, based on the indicator of MSE, the prediction results show that ARMA model beats the SVM. The future work of this project can be improved through applying more statistic methods. For example, Long Short Term Memory (LSTM) and Random Forest are commonly used algorithms when predicting asset prices, selecting these deserve more attention.

AUTHORS' CONTRIBUTIONS

These authors contributed equally.

REFERENCES

[1] Wright, C. S. (2008). Bitcoin: a peer-to-peer electronic cash system. *Available at SSRN 3440802*.

[2] Bouoiyour, J., & Selmi, R. (2015). What does Bitcoin look like?. *Annals of Economics & Finance*, *16*(2).

[3] Kondor, D., Pósfai, M., Csabai, I., & Vattay, G. (2014). Do the rich get richer? An empirical analysis of the Bitcoin transaction network. *PloS one*, *9*(2), e86197.

[4] Yang, Z., & Aldous, D. (2015). Geometric brownian motion model in financial market. *University of California, Berkeley*.

[5] Karasu, S., Altan, A., Saraç, Z., & Hacioğlu, R. (2018, May). Prediction of Bitcoin prices with machine learning methods using time series data. In *2018 26th signal processing and communications applications conference (SIU)* (pp. 1-4). IEEE.

[6] Chen, Z., Li, C., & Sun, W. (2020). Bitcoin price prediction using machine learning: An approach to sample dimension engineering. *Journal of Computational and Applied Mathematics*, *365*, 112395.

[7] Hiemstra, C., & Jones, J. D. (1994). Testing for linear and nonlinear Granger causality in the stock price-volume relation. *The Journal of Finance*, *49*(5), 1639-1664.

[8] Bouoiyour, J., & Selmi, R. (2017). The Bitcoin price formation: Beyond the fundamental sources. *arXiv preprint arXiv:1707.01284*.

[9] Al-Yahyaee, K. H., Mensi, W., & Yoon, S. M. (2018). Efficiency, multifractality, and the long-memory property of the Bitcoin market: A comparative analysis with stock, currency, and gold markets. *Finance Research Letters*, *27*, 228-234.

[10] Wu, X., Chen, H., Wang, J., Troiano, L., Loia, V., & Fujita, H. (2020). Adaptive stock trading strategies with deep reinforcement learning methods. Information Sciences, 538, 142-158.

[11] Hiemstra, C., & Jones, J. D. (1994). Testing for linear and nonlinear Granger causality in the stock price-volume relation. The Journal of Finance, 49(5), 1639-1664.

[12] Tiwari, A. K., Adewuyi, A. O., & Roubaud, D.

(2019). Dependence between the global gold market and emerging stock markets (E7+ 1): Evidence from Granger causality using quantile and quantile-on-quantile regression methods. The World Economy, 42(7), 2172-2214.

[13] Rosoł, M., Młyńczak, M., & Cybulski, G. (2022). Granger causality test with nonlinear neural-network-based methods: Python package and simulation study. Computer Methods and Programs in Biomedicine, 216, 106669.

[14] Gao, J., Guo, H., & Xu, X. (2022). Multifactor Stock Selection Strategy Based on Machine Learning: Evidence from China. Complexity, 2022.