A Dilution Method of Bitcoin Price Prediction And Investment Strategy Based on ARIMA

Hongze Guo[1], Zhenchao Gao[1], Yingchang Liu[2], Yue Yu[3], Yuhang Zhou[1], *Zhen Zhou[1]

e-mail: 916833878@qq.com, e-mail: 2087764513@qq.com,

e-mail: 1072304530@qq.com, e-mail: 1830867761@qq.com,

e-mail: 1071565909@qq.com, *Corresponding author: 1006300202@qq.com

[1]Hebei University, HBU, Baoding, China

[2]Wuhan City College, WIC, Wuhan, China

[3]Ningbo University, NBU, Ningbo, China

Abstract—At present, bitcoin is hot, and its price prediction methods are emerging in endlessly, but there has never been a means to accurately predict mining disasters and plummets. Based on ARIMA, this paper proposes a bitcoin price prediction method based on cycle dilution method, which successfully realizes the cycle price prediction of bitcoin, with an error of less than 5%. The threshold trading experiment shows that this scheme can obtain a very high rate of return within five years, which is of great practical significance.

Keywords-ARIMA; mean value prediction; cyclic optimization; bitcoin

1. Introduction

With the continuous promotion of blockchain [1] and metaversity, more and more people value Bitcoin investment. Therefore, there is a great demand for bitcoin's price trend, price prediction and investment strategy.

However, bitcoin [2] has experienced many short-term price plunges in history, and these plunges are often without warning and difficult to predict. Modeling based on the global and international situation, the policy trends of some countries and the current new COVID [3] epidemic, we can achieve accurate prediction [4] to a certain extent, but the data that needs to be collected is massive, and the modeling process is extremely cumbersome. It is also difficult to feedback when new emergencies appear.

In view of the above situation, this paper proposes a time series-based dilution method of Bitcoin price prediction, which performs the traditional stock's limited buying and selling operation on the predicted price. Based on the comparison of annual Bitcoin price data in the past five years, the method proposed in this paper has extremely high prediction accuracy, and the final investment [5] method has a very high and reliable rate of payback.

BDEIM 2022, December 02-03, Zhengzhou, People's Republic of China

DOI 10.4108/eai.2-12-2022.2328686

2. Materials and Methods

2.1 ARIMA Introduction

ARIMA is a classic time series analysis method that is a linear combination of past errors and stationary time series past values. ARIMA is often used in short-term [6] forecasting.

ARIMA is mainly composed of AR and MA, AR is an autoregressive model, which must meet the requirements of stationarity, and the mathematical representation of the p-order autoregressive model is as follows:

$$y_t = \mu + \sum_{i=1}^{p} \gamma_i y_{t-i} + \varepsilon_t \tag{1}$$

MA is a moving average model, which mainly uses past interference and current interference to predict the real value of the model, and the mathematical representation method of the q-order moving average model is as follows [7]:

$$y_t = \mu + \sum_{i=1}^{p} \theta_i \varepsilon_{t-i} + \varepsilon_t \tag{2}$$

2.2 Model Building

At present, the existing modeling method adds parameters such as policy factors and panic factors to the model, but still maintains a certain percentage deviation, and the prediction of large-scale declines is still limited. In view of this, we chose to use the cycle prediction method, that is, to use a certain number of days as a period to predict the time series through the cycle, so as to minimize the impact of the huge decline such as mining disasters and improve the prediction accuracy of the model as much as possible. We chose to make a one-day forecast for the future price of Bitcoin, then we compare the resulting data with the actual data and perform a series of parameter corrections on the model to improve the model’s accuracy.

This is a type of time series analysis that analyzes numerical series of displayed indicator values in chronological order. It is combined with an AR model with a MA model [8] and uses the predictions obtained [9] through first-order or higher-order differential processing. It requires differential processing of the data before modeling, converting it into a stationary time series, and then modeling it [10]. This can be defined as:

$$ARIMA(p,d,q): y_t^{'} = \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i}^{'} + \varepsilon_t + \sum_{i=1}^{q} \beta_i \varepsilon_{t-i} \tag{3}$$

$$y_t^{'} = \Delta^d y_t = (1-L)^d y_t \tag{4}$$

In this formula, *p* is the order of the AR models; *d* is the order of the differences; *q* is the order of the MA models. The symbol *L* also represents a lag operator and has the following properties:

$$LC = C(C \text{ is constant}) \tag{5}$$

$$(L^i + L^j)\, y_t = y_{t-i} + y_{t-j} \tag{6}$$

$$L^i L^j y_t = y_{t-i-j} \tag{7}$$

After several predictive simulations, we found that ARIMA (0,1,1) had the highest prediction accuracy, so we chose ARIMA (0,1,1) as our predictive model, which is defined as:

$$y_t = \alpha_0 + y_{t-1} + \varepsilon_t + \beta_1 * \varepsilon_{t-1} \tag{8}$$

2.3 Rule Setting

Since the service fees for trading Bitcoin are extremely high, we set a buy-and-sell operation in this model to be performed only once a day. The daily Bitcoin price is obtained through NASDAQ.

In terms of price prediction, we mainly use diluted prediction, that is, by calculating the average value of prices over many days, and predicting the average value of prices in the next cycle through the first several average values each time, so as to dilute the impact of sudden decline to the greatest extent. At the same time, we believe that the sudden decline of prices on a certain day cannot be obtained through basic prediction.

First, take several existing past prices as samples to speculate the price of the next day. We take 3, 5 and 7 days as samples. After comparing the real price data of bitcoin from 9/11/2016 to 9/11/2021, the accuracy curve is as follows. It can be concluded that 5 days is the prediction period with the smallest average deviation through this model.

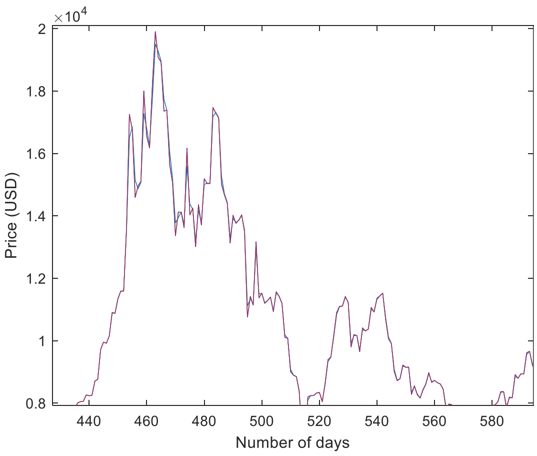

Figure 1. Precision Curve

The dilution method is implemented as follows: set the number of days 5 contained in an interval, take 7 days as a period, and take multiple samples to predict the value of the first sample in the future. This method only changes the sample from days to a multi-day average on the basis of our existing method, which is extremely feasible. Subsequently, we made predictions and statistics. We use this modified model which can achieve great precision for price predictions over a period of three days. In order to achieve maximum accuracy, we made another five-day cycle and a seven-day cycle prediction, and the number of samples collected was 7, 12, 20. Considering that the product of the period and the number of samples here is the date of the investor's initial observation, the smaller the date, the better the accuracy can be. So the 12 samples with the highest accuracy and minimal number of days in the single-day forecast are used as the method. The result is shown in the following figure:

Figure 2. Prediction Error

3. Investment Strategy Application

3.1 All In Strategy to Test the Ideal Maximum Income

We have forecasted and simplified our investment strategy, starting with predicting maximum payback by using the method of maximum risk. Because we have been able to predict the average price of bitcoin in the next cycle with 12 samples with high accuracy by taking the five-day average price as a sample, our extreme buying strategy is as follows:

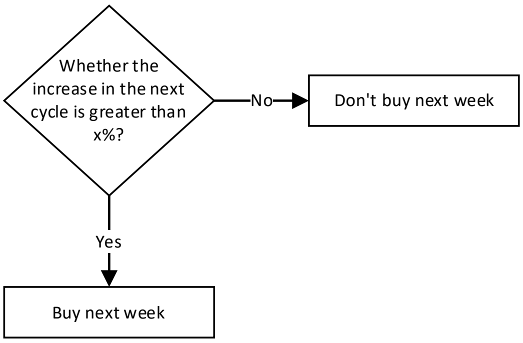

Figure 3. Extreme Buying Strategy

The chart above is a strategy that determines whether to buy or sell in a week, and when it is judged to buy in the week, the model will automatically enter the buying judgment in the cycle. Specific execution is, the model has predicted the average price during the current week, if the price of the day reaches the predicted price, buy immediately, if it does not reach, then wait for the next day. If it is already the last day of the cycle, and the price does not reach the expected average, then it means that the price of the entire cycle is lower than expected. The price prediction of the cycle is wrong, and failure information will be fed back to the model, waiting for the next time to judge whether it is allowed to buy. The same goes for the strategy for selling. The specific flow chart is as follows:

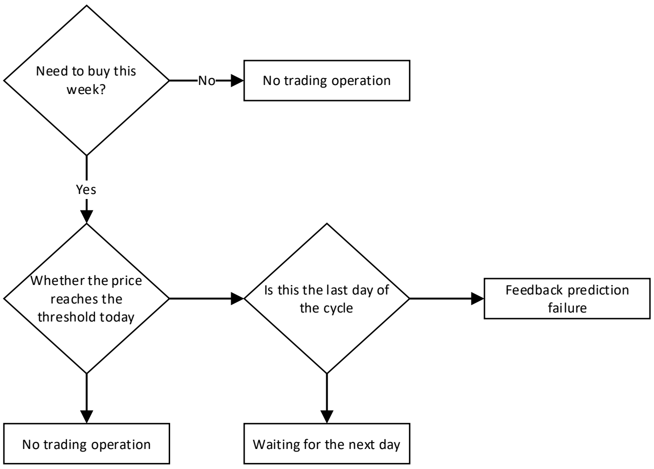

Figure 4. Buying and selling strategies

Based on the prediction data obtained above, we apply the threshold buying and selling principle in stock trading, that is, we predict to sell when the decline in the next cycle is greater than the threshold, and buy when the increase in the next cycle is greater than the threshold. Based on this method, we have obtained the following linkage diagram of income, buying threshold and selling threshold. It should be noted that X and Y in the above figure are uncertain values, which need to be iterated through the circular algorithm. After that the best value can be obtained. Therefore, we wrote a circular code, iterated the absolute values of *x* and *y* from 0 to 5, stepped 0.1, and obtained the following results:

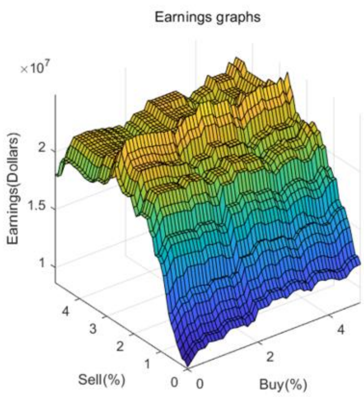

Figure 5. Yield results

According to the original data, if the ALL IN strategy is used for buying according to the optimal threshold, the principal of the cost of $1000 after five years can be worth up to $24792713.3344129. At this time, the strategy is: if it is predicted that the next cycle's increase will be 2.4% then mark the next cycle as the buying cycle, and if it is predicted that the next week's decrease will be 2.7%, then mark the next circle as the selling cycle.

3.2 Improvement Strategy

Through ALL IN's extreme strategy of completely ignoring risks, we get the buy-and-sell thresholds corresponding to the best payback (among them, the buy-and-sell thresholds are: bitcoin buying threshold: 2.7%, bitcoin selling threshold: -2.4%, gold buying threshold 0.7%, gold selling threshold: -1.5%). Then we will put it into the formula we use next:

$$btb\_ratein = 0.027 \tag{9}$$

$$btb\_rateout = -0.024 \tag{10}$$

However, keep using the ALL IN strategy obviously violates the normal situation. We then add a new rule: when the cash is less than 10% of the total assets, we don't buy. When we predict that the future change is greater than the buying threshold, we will buy bitcoin and gold with 30% of cash. If it exceeds the threshold by 1.5 times, we will buy with 50% of cash. However, if it is predicted that the future change is less than the selling threshold, half of bitcoin and gold will be sold. And we will sell all of them when future change is less than 1.5 times the selling threshold. The parameter values and definitions added are as follows:

$$money\_min = 0.10 \tag{11}$$

$$ratetime\_stop\_on = 1.5 \tag{12}$$

$$ratetime\_stop\_down = 1.5 \tag{13}$$

At the same time, in order to turn risk into visual data, we must define a risk index formula. Since the coefficient is related to the coefficient and the trading threshold, the risk index is positively correlated with the profit-stopping coefficient and the loss-stopping coefficient. In real transactions, the risk brought by stopping profit is less than stopping loss, so we use 0.2 as the coefficient of the profit-stopping coefficient. The minimum guarantee coefficient determines the amount of reserves. The smaller the minimum guarantee coefficient, the less the reserves, and the higher the risk index. Since the minimum guarantee coefficient can be 0, and the denominator cannot be 0, we add 0.5 to the denominator to standardize the result of the formula. Therefore, we define the following risk index formula:

$$\begin{aligned} &danger = [(0.2 * ratetime\_stop\_on) \\ &* ratetime\_stop\_down] / \\ &(money\_\min + 0.5) \end{aligned} \tag{14}$$

To simplify the calculation, we always keep the profit-stopping coefficient and loss-stopping coefficient the same, that is:

$$danger = (0.2 * ratetime \wedge 2) / (money\_\min + 0.5) \tag{15}$$

3.3 Verification

In terms of proof, we still use the circular algorithm to calculate the best value. The value of the risk coefficient we set is divided by the number of 1. The closer the absolute value of the risk coefficient is to 1, the lower the risk is. It should be noted that the risk coefficient is always a number bigger than 0. If the risk coefficient is less than 1, it means that the risk is caused by the profit-stopping strategy, while if it is bigger than 1, it means that the risk is caused by the loss-stopping measures. In order to make us see the risk index chart intuitively and accurately, we take the reciprocal of the risk brought by the profit-stopping strategy, so that the risk coefficients of the profit-stopping strategy and the loss-stopping strategy can be compared together. Then, we carry out a circular algorithm, and the risk index chart and the income chart after applying the strategy are as follows:

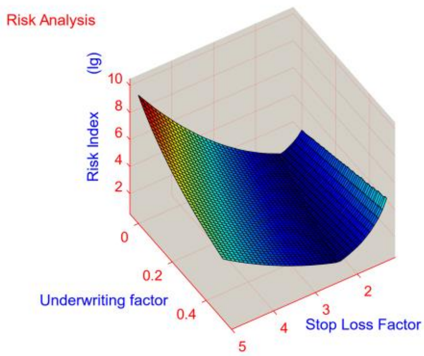

Figure 6. Risk Analysis

After the income risk is obtained, we apply it to the new strategy. After the circular nesting algorithm, the final result is shown in the figure below:

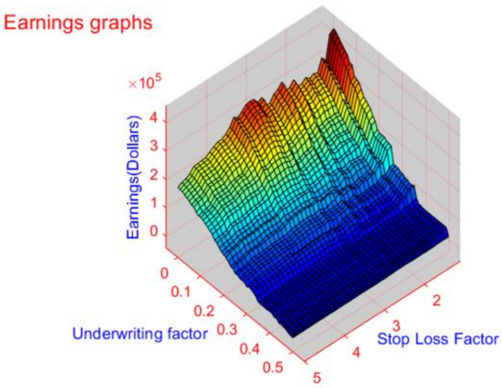

Figure 7. Income Chart

Finally, with the initial capital of $1000, the final value of a five-year investment using the method proposed in this paper is $209036.2. The result comes from investments with full consideration of risks. At the same time, the way of investment is no longer ALL IN, but a method of changing the percentage of investment amount according to the judgment of situation, taking full account of stability while maintaining accuracy.

4. Conclusions

Based on the results and discussions presented above, the conclusions are obtained as below:

(1) The dilution prediction method proposed in this paper has great accuracy.

(2) The periodic average value and the price prediction in the case of large samples, which are both obtained by dilution method, have great accuracy.

(3) The investment method proposed in this paper is highly feasible and prudent.

Acknowledgment

This work is supported by National Natural Science Foundation of China under Grant (62103127), Youth Natural Science Foundation of Hebei Province under Grant F2020201048 and Advanced Talents Incubation Program of the Hebei University under Grant 521000981366.

References

[1] Junyu Ren, Jinze Li, Huaxing Liu, Tuanfa Qin.Task Offloading Strategy with Emergency Handling and Blockchain Security in SDN-Empowered and Fog-Assisted Healthcare IoT[J]. Tsinghua Science and Technology, 2022, 27(04):760-776.

[2] Teng Long, Jiasheng Xu, Luoyi Fu, Xinbing Wang. Analyzing and De-Anonymizing Bitcoin Networks: An IP Matching Method with Clustering and Heuristics[J]. China Communications, 2022, 19(06): 263-278.

[3] YE Chao, PAN Jiawei, LIU Zhimin. The historical logics and geographical patterns of rural-urban governance in China[J]. Journal of Geographical Sciences, 2022,32(07):1225-1240.

[4] Li Yang, Yifei Zou, Minghui Xu, Yicheng Xu, Dongxiao Yu, Xiuzhen Cheng. Distributed Consensus for Blockchains in Internet-of-Things Networks[J]. Tsinghua Science and Technology,2022,27(05):817-831.

[5] Lü Jingye,Du Jingnan,Cao Ming,Fan Xiufeng. Carbon emissions trading price prediction using ARIMA-SVM model[J].Journal of Xi'an University of Science and Technology,2020,40(03):542-548.DOI:10.13800/j.cnki.xakjdxxb.2020.0323.

[6] Peng Hongxing,Zheng Kaihang,Huang Guobin,Lin Dousheng,Yang Zhichao,Liu Huanai. Vegetable price prediction based on BP, LSTM and ARIMA model [J].China Journal of Agricultural Machinery And Chemistry,2020,41(04):193-199.DOI:10.13733/j.jcam.issn.2095-5553.2020.04.031.

[7] Lin Zhengyang,Ni Xiaojie,Wang Bo. Short-term Forecast of Photovoltaic Building Integration Stock Based on Improved ARIMA Model[J].Computer Times, 2022(07):40-43. DOI:10.16644/j.cnki.cn33-1094/tp.2022.07.010..

[8] Liu Ming, Liu Wen-dong, Li Yong-xiang. Solubility, Hansen solubility parameter, molecular simulation and apparent thermodynamic analysis of 5-Fluorouracil in four binary solvents[J]. Journal of Molecular Liquids, 2022,362.

[9] Gehr Felix, Theurich Timo, Monjaraz-Tec Carlo,Gross Johann, Schwarz Stefan, Hartung Andreas, Krack Malte. Computational and experimental analysis of the impact of a sphere on a beam and the resulting modal energy distribution[J]. Mechanical Systems and Signal Processing,2022,180.

[10] Jie Tang, Mingcai Liu,Guowei Bo, Fulin Jiang, Chunhui Luo, Jie Teng, Dingfa Fu,Hui Zhang. Unraveling precipitation evolution and strengthening function of the Al-Zn-Mg-Cu alloys with various Zn contents: Multiple experiments and integrated internal-state-variable modeling[J].Journal of Materials Science & Technology,2022,116(21):130-150.