The Formulation of Investment Strategy Based on ARIMA Model

Wenkai Cui[1], Jilong Zhou[2], and Yubo Shan[3]

[1]1916020106@s.upc.edu.cn, [2]1909010203@s.upc.edu.cn, [3]1903040217@s.upc.edu.cn

[1]College of Oceanography and Space Informatics, China University of Petroleum (East China), West Changjiang Road, Qingdao, China

[2]College of Science, China University of Petroleum (East China), West Changjiang Road, Qingdao, China

[3]College of Chemistry and Chemical Engineering, China University of Petroleum (East China), West Changjiang Road, Qingdao, China

Abstract: With the development of international financial market, financial investment has become an indispensable part of human life. More and more people want to invest in the financial market to obtain income. However, benefits are often accompanied by risks. Meanwhile, gold and bitcoin are two important investments in financial investment, with great volatility. In order to reduce the occurrence of tragedy and help people better manage their money, this paper first establishes a time series model (ARIMA) based on the five-year gold bitcoin transaction price from 2016 to 2021 to predict the price trend and development of gold and bitcoin. Then the Apriori algorithm [1] is used to make decisions, and the rate of return is calculated through the decision model. The goodness of fit is 97.78% of the total income curve.

Keywords: Bitcoin, ARIMA, Investment portfolio, Apriori algorithm

1 Introduction

In the trading market, investors choose the target in the market, and different targets will have their corresponding transaction prices at different times, and income can be earned by buying or selling at different transaction prices. For market traders, it is extremely important to predict the subsequent price through the price change of the underlying over a period of time, and to operate the assets held in their hands to maximize the transaction.

Gold trading attracts a large number of investors with its large-capacity market and low transaction costs. At the same time, Bitcoin, as an emerging currency, has also become one of the giants of market transactions.

This paper first established a time series model, and through the scoring process and ADF test [2], as well as ACF and PACF graph and residual test, this paper obtained the forecast data of gold and Bitcoin, and then, this paper established a decision model based on data preprocessing.

ICIDC 2022, June 17-19, Qingdao, People's Republic of China

DOI 10.4108/eai.17-6-2022.2322763

And give the optimal trading strategy. By changing the position adjustment amount, the model analyses the change trend of the subsequent position adjustment amount. And by calculating the Sharpe ratio and the Treynor ratio, it shows that our model is optimal. This paper then perform an analysis of model sensitivity, and finally, this paper evaluate and draw conclusions from the models this paper have built, and based on those conclusions, explain our strategies, models, and results to traders in the form of a memorandum.

2 Time Series Analysis Model Based on ARIMA

Time series analysis is a method of establishing mathematical models, curve fitting and parameter estimation methods by observing the overall time series data and fitting the later data by observing the data.

2.1 Data Stationarity Test

Firstly, this paper tests the stability of gold and bitcoin transaction prices in five years by ADF method. When the difference order is 1 and 2, the P-Value and T-test Result values of the daily transaction price data of gold and bitcoin meet the requirements, and the null hypothesis is rejected, and both are stationary time series in this case.

2.2 Determining the Model Parameters of ARIMA

It can be seen from the previous step that when the difference order is 1 or 2, the requirements for stationary time series are met.

According to the calculation results, the daily transaction price data of gold corresponds to p=5, q=5, and the daily transaction price data of Bitcoin corresponds to p=3, q=3. Accordingly, the ARIMA models [3] for the daily transaction price data of gold and the daily transaction price of Bitcoin are initially determined as ARIMA (5, 1, 5) and ARIMA (3, 1, 3) respectively.

2.3 Data inspection and processing of ARIMA Model

In order to make the daily trading price data of gold and the daily trading price data of bitcoin become a stable time series, the data is transformed into a stable time series after a first differential processing [4]. Then, the residual test is carried out. The residual distribution obtained conforms to the normal distribution, indicating that the residual of gold and bitcoin transaction price data does not have autocorrelation and partial correlation [5].

2.4 Next Day Trading Price Forecast

After the above data processing, ARIMA can be calculated for the daily transaction price data of gold and the daily transaction price of Bitcoin.

As predicted by the model, the transaction prices of gold and bitcoin on the next day (excerpts) are shown in the following Table1.

Table1: Prediction Results of Time Series Analysis Model Based on ARIMA

Date	p^{k_g}	p^{p_g}	p^{k_b}	p^{p_b}
9/11/2016	1324.6	1323.80965	621.65	774.3601
9/12/2016	1323.65	1322.46565	609.67	774.07493
…	…	…	…	…
9/9/2021	1794.6	1819.377689	46368.69	47474.94051
9/10/2021	/	1804.195805	/	45558.66243

Note: p^{k_g} is known transaction price data of gold on the next day; p^{p_g} is predicted transaction price data of gold on the next day; p^{k_b} is known transaction price data of bitcoin on the next day; p^{p_b} is predicted transaction price data of bitcoin on the next day.

From the Table1, we can calculate the relative error between the predicted transaction price and the actual transaction price of gold and bitcoin and visualize it.

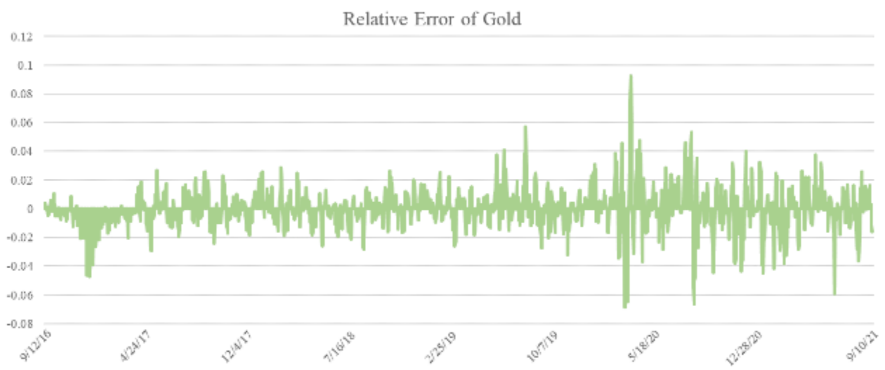


Figure 1 The Relative Error of The Next Day's Gold's Transaction Price

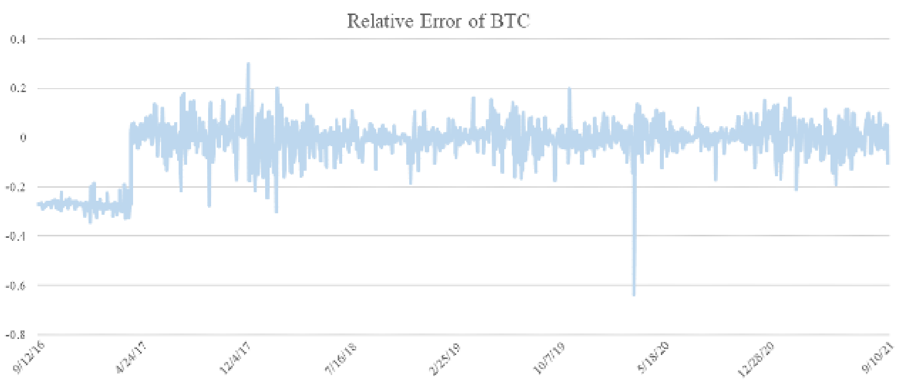


Figure 2 The Relative Error of The Next Day's Bitcoin's Transaction Price

As can be seen from Figure 1 and Figure 2, the relative error of the next day's bitcoin transaction price is larger than that of the gold's transaction price, and the change is more unstable.

3 The Decision Model BASED ON Apriori algorithm

3.1 Data Preprocessing

As a measure of the center of a distribution, the triple mean has an advantage in that it combines the median and the median number to reflect both the center value of the distribution without losing attention to extreme values [6]. Therefore, we find the triple mean of the increase to replace most of the increase in gold and bitcoin. If the increase is greater than the triple mean, we can assume that it will continue to rise.

The triple mean is determined by the following Eq. (1):

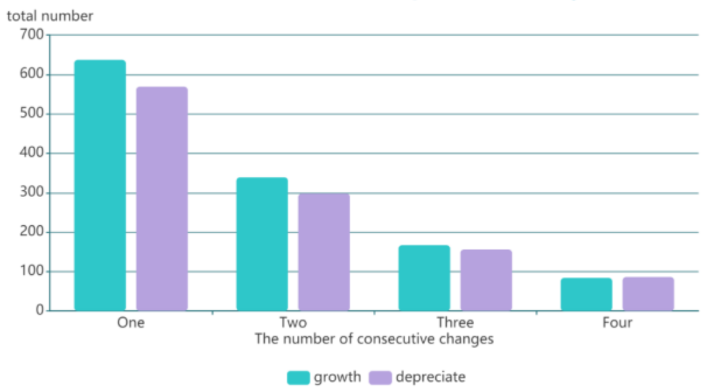
$$M_t = \frac{1}{4}Q_{0.25} + \frac{1}{2}Q_{0.5} + \frac{1}{4}Q_{0.75} \quad (1)$$

In Table2, we regard $Q_{0.1}$ as the biggest drawdown of the day. From Table 2, we can get Gold M_t=0.004725, Bitcoin M_t=0.0188, Gold $Q_{0.1}$=-0.0135, Bitcoin $Q_{0.1}$=-0.0643.

Table2: Quantiles of Daily Transaction Prices for Gold and Bitcoin

	$Q_{0.1}$	$Q_{0.25}$	$Q_{0.5}$	$Q_{0.75}$	$Q_{0.9}$
Gold's Price Rise	0.0007	0.0018	0.0041	0.0089	0.0134
Gold's Price Falls	-0.0135	-0.0080	-0.0044	-0.0018	-0.0009
Bitcoin's Price Rise	0.0020	0.0058	0.0158	0.0378	0.0134
Bitcoin's Price Falls	-0.0643	-0.0347	-0.0150	-0.0055	-0.0021

According to the above rules, the distribution chart of the consecutive rise and fall days of gold and bitcoin transaction prices is calculated, as shown in the Figure 3:



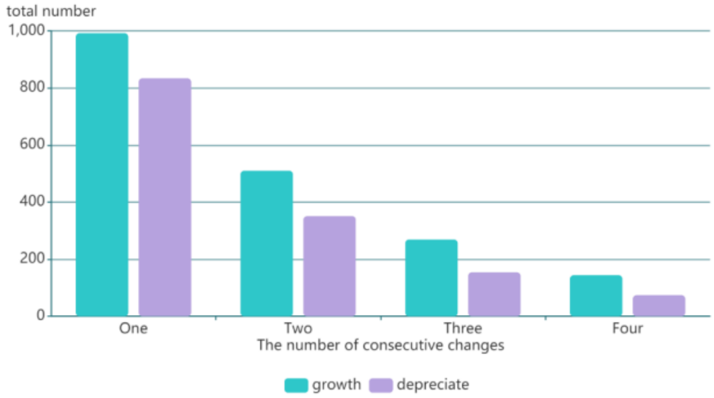


Figure 3 The Number of Rise and Falls in Transaction Prices of Gold and Bitcoin

According to the calculation results, we can count the proportion of different days of consecutive increases and decreases in the transaction prices of gold and Bitcoin (making the proportion of the single rise and fall as 100%).

After the above preprocessing, we get the required data and start to build the decision model.

3.2 Model Building

Firstly, the notations used in this paper are displayed in Table4.

From the Table3, it can be seen that the number of consecutive days of rise for four days or less accounts for 90% of the total number of rise days, then we can think that if the rise has continued for four consecutive days, the next day continues to grow is a small probability event, which can be ignored.

Table3: The Proportion of Different Times of Consecutive Increases and Decreases in the Transaction Prices of Gold and Bitcoin

Transaction Item	Situation of Changing	Two Days	Three Days	Four Days
Gold	Rise	53.218%	79.435%	92.622%
	Fall	52.373%	79.789%	94.903%
Bitcoin	Rise	51.411%	78.528%	93.044%
	Fall	42.086%	60.552%	69.424%

Table 4: Notations Used in this Paper

Symbol	Description
M_t	triple mean
$Q_{0.1}^{GF}$	0.1 quantile of gold decline
$Q_{0.5}^{GR}$	median gold increase

$Q_{0.1}^{BF}$	0.1 quantile of bitcoin decline
$Q_{0.5}^{BR}$	median bitcoin increase
α_g	commission for gold (1%)
α_b	commission for bitcoin (2%)
t_g	the highest consecutive increase in the gold transaction price
t_b	the highest consecutive increase in the bitcoin transaction price
S_g^i	the adding to positions in the gold transaction price that has fallen i consecutive times
S_b^i	the adding to positions in the bitcoin transaction price that has fallen i consecutive times

Through the above analysis of the rise and fall, we can add or cover the number of times of rise or fall. By reading literature [7], we can get that when the number of consecutive rise is higher, the probability of the next fall is relatively large.

Therefore, our decision is based on the fact that if the predicted price of tomorrow will fall, we will increase the position in time, and if the predicted price of tomorrow will rise, we will replenish the position in time.

Due to the small probability (much less than 90%) of the bitcoin price falling for four consecutive days, it is not considered here. According to Table5, we get t_g and t_b is 4.

At the same time, we set the adding to positions for gold at the first fall day as S_g^1, then the adding to positions for gold S_g^2 at the second fall day satisfies Eq. (2)

$$\left[S_g^1(1-\alpha_g) - S_g^1(1-\alpha_g)\cdot\frac{Q_{0.1}^{GF}}{t_g} + S_g^2(1-\alpha_g)\right]Q_{0.1}^{GR} + S_g^2(1-\alpha_g)\cdot\frac{Q_{0.1}^{GF}}{t_g} + S_g^1(1-\alpha_g)\cdot\frac{Q_{0.1}^{GF}}{t_g} = 0 \tag{2}$$

The adding to positions S_g^3 at the third fall day satisfies Eq. (3) and the adding to positions S_g^4 at the fourth fall day satisfies Eq. (4):

$$-S_g^1(1-\alpha_g)\cdot\left(\frac{Q_{0.1}^{GF}}{t_g}\right)^2 + S_g^2(1-\alpha_g)\cdot\left(\frac{Q_{0.1}^{GF}}{t_g}\right) + \left\{S_g^1(1-\alpha_g)\cdot\left[1-\left(\frac{Q_{0.1}^{GF}}{t_g}\right)^2\right] + S_g^1(1-\alpha_g)\cdot\left[1-\left(\frac{Q_{0.1}^{GF}}{t_g}\right)\right] + S_g^3(1-\alpha_g)\right\}\cdot Q_{0.5}^{GR} = 0 \tag{3}$$

$$S_g^1(1-\alpha_g)\cdot\left(\frac{Q_{0.1}^{GF}}{t_g}\right)^3+S_g^2(1-\alpha_g)\cdot\left(\frac{Q_{0.1}^{GF}}{t_g}\right)^2+S_g^3(1-\alpha_g)\cdot\left(\frac{Q_{0.1}^{GF}}{t_g}\right)$$
$$+\left\{S_g^1(1-\alpha_g)\cdot\left[1-\left(\frac{Q_{0.1}^{GF}}{t_g}\right)^3\right]+S_g^2(1-\alpha_g)\cdot\left[1-\left(\frac{Q_{0.1}^{GF}}{t_g}\right)^2\right]\right.$$
$$\left.+S_g^3(1-\alpha_g)\cdot\left[1-\left(\frac{Q_{0.1}^{GF}}{t_g}\right)\right]+S_g^4(1-\alpha_g)\right\}\cdot Q_{0.5}^{GR}=0 \tag{4}$$

The same is true when set the adding to positions for bitcoin. When S_b^1 is determined, we can get S_b^2, S_b^3, S_b^4 according to Eq. (5), Eq. (6) and Eq. (7).

$$\left[S_b^1(1-\alpha_b)-S_b^1(1-\alpha_b)\cdot\frac{Q_{0.1}^{BF}}{t_b}+S_b^2(1-\alpha_b)\right]Q_{0.1}^{BR}+$$
$$S_b^2(1-\alpha_b)\cdot\frac{Q_{0.1}^{BF}}{t_b}+S_b^1(1-\alpha_b)\cdot\frac{Q_{0.1}^{BF}}{t_b}=0 \tag{5}$$

$$-S_b^1(1-\alpha_b)\cdot\left(\frac{Q_{0.1}^{BF}}{t_b}\right)^2+S_b^2(1-\alpha_b)\cdot\left(\frac{Q_{0.1}^{BF}}{t_b}\right)+\left\{S_b^1(1-\alpha_b)\cdot\left[1-\left(\frac{Q_{0.1}^{BF}}{t_b}\right)^2\right]+\right.$$
$$\left.S_b^1(1-\alpha_b)\cdot\left[1-\left(\frac{Q_{0.1}^{BF}}{t_b}\right)\right]+S_b^3(1-\alpha_b)\right\}\cdot Q_{0.5}^{BR}=0 \tag{6}$$

$$S_b^1(1-\alpha_b)\cdot\left(\frac{Q_{0.1}^{BF}}{t_b}\right)^3+S_b^2(1-\alpha_b)\cdot\left(\frac{Q_{0.1}^{BF}}{t_b}\right)^2+S_b^3(1-\alpha_b)\cdot\left(\frac{Q_{0.1}^{BF}}{t_b}\right)$$
$$+\left\{S_b^1(1-\alpha_b)\cdot\left[1-\left(\frac{Q_{0.1}^{BF}}{t_b}\right)^3\right]+S_b^2(1-\alpha_b)\cdot\left[1-\left(\frac{Q_{0.1}^{BF}}{t_b}\right)^2\right]\right.$$
$$\left.+S_b^3(1-\alpha_b)\cdot\left[1-\left(\frac{Q_{0.1}^{BF}}{t_b}\right)\right]+S_b^4(1-\alpha_b)\right\}\cdot Q_{0.5}^{BR}=0 \tag{7}$$

3.3 Calculation of Decision Income

We set S_g^1=100 and S_b^1=100 for the initial increase, we can calculate the adding to positions for the second, third, and fourth of gold and bitcoin respectively in Table 5.

Based on the above model, we substitute the predicated data to get the decisions made every day, and according to the data, calculate the total income of each day. We select the first five days and the last five days for display in Table 6.

Table 5: The Adding to Positions for the Second, Third, and Fourth of Gold and Bitcoin

Transaction Item	S_g^1/S_b^1	S_g^2/S_b^2	S_g^3/S_b^3	S_g^4/S_b^4
Gold	100	118.73	141.62	192.57
Bitcoin	100	111.09	136.56	176.87

Table 6: Daily Holdings and Total Income

Transaction date	U.S. dollar holdings on the day	Gold holdings on the day	Bitcoin holdings on the day	Total Income
9/11/2016	945.54346	0	0.0876	1378.26619
9/12/2016	923.1468	0.016908244	0.0876	1398.269402
9/13/2016	923.1468	0.016908244	0.0876	1392.303189
…	…	…	…	…
9/9/2021	3297.896106	0.0158	0.0082	3751.183678
9/10/2021	3303.884706	0.0162	0.0083	3770.301956

Title: Plot and observe the original data, it can be seen that the image is very close to the exponential function image.

$$y = a \cdot e^{bx} \tag{8}$$

We use the form as Eq. (8) to fit the data. The fitted function is Eq. (9):

$$y = 1278 \cdot e^{0.0006085x} \tag{9}$$

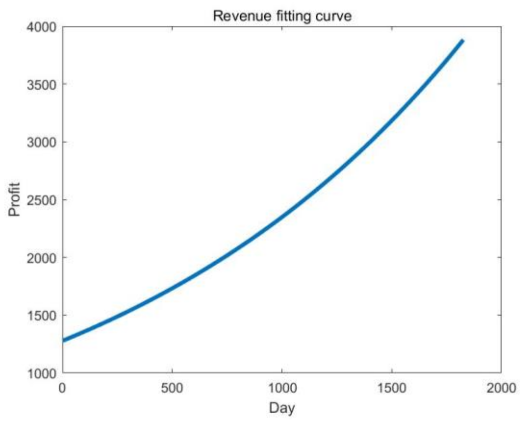


Figure 4 Revenue Fitting Curve

According to calculation and Revenue Fitting Curve which is shown in Figure 4, the fitted R-square is about 0.9778, indicating that the fitted curve can roughly reflect the relationship between our decision benefit and the number of days.

4 Conclusion

Since the 1980s, financial investment has developed rapidly, and gold bitcoin has gradually become an important investment in the wave of financial investment. Market traders often buy and sell unstable assets, so it is very important to specify reasonable strategies to obtain higher returns.

This paper designs a model based on ARIMA to predict the development trend of gold and currency over a period of time, and compares the predicted results with the actual results. It is found that the relative error of gold is smaller and more stable, while as a new investment product, currency fluctuates due to the intervention of a large number of speculators. Therefore, the relative error of bitcoin is larger and the stability is poor. Based on the predicted data, this paper designs a decision-making model to calculate the amount of gold and bitcoin that should be bought every day, and predict the total return. After fitting the curve, the results show that the fitted curve has high accuracy.

This paper makes a better decision by predicting the unknown data. Through the prediction model, we find that the number of days of continuous rise or fall does not exceed 4 days, so we take 4 days as the investment cycle. If the number of days of continuous fall, we need to increase our positions to adjust our position price, but our capital is limited. Therefore, when the number of days of decline exceeds 4 days, we choose to slow down the investment to avoid deepening. Through continuous decision-making, we have achieved a total annual return of about 30%, which is a better strategy for real investment.

References

[1] Shona Chayy Bilqisth, Khabib Mustofa. Determination of Temporal Association Rules Pattern Using Apriori Algorithm [J]. IJCCS (Indonesian Journal of Computing and Cybernetics Systems), 2020, 14(2):

[2] Slocombe Geoff. Night fighting equipment: Research and technology developments for the ADF[J]. Asia-Pacific Defence Reporter (2002),2019,45(10):

[3] Mandrikova Oksana, Fetisova Nadezhda, Polozov Yuriy. Hybrid Model for Time Series of Complex Structure with ARIMA Components [J]. Mathematics,2021,9(10):

[4] Ye R, Suganthan P N, Srikanth N, et al. A hybrid ARIMA-DENFIS method for wind speed forecasting[C]// IEEE International Conference on Fuzzy Systems. IEEE, 2013.

[5] Zhang W H. Algorithm improvement and simulation of ARIMA model for time series pre-diction [J]. Information & Computer (Theory Edition), 201,33(05):53-56.

[6] Velankar S, Valecha S, Maji S. Bitcoin price prediction using machine learning[C]// 2018 20th International Conference on Advanced Communication Technology (ICACT). 2018.

[7] Yao Haixiang, Li Junwei, Xia Shenghao, Chen Shumin. Fuzzy Transaction Decision Based on Apriori Algorithm and Neural Network [J]. Journal of Systems Science and Mathematics, 201,41(10):2868-2891.