Trading Strategy: An Asset Decision Model Based on Past Daily Trading Price

Xiaoyu Miao[1,*], Fanqi Xu[2], Yunpeng Xie[1]

[1]School of water conservancy and civil engineering NEAU, Harbin, Heilongjiang, 150006

[2]School of Economics and Management, Northeast Agricultural University, Harbin, Heilongjiang, 150006

*Corresponding author. Email: miaoxiaoyu0302@163.com

Keywords: quantitative trading; ARIMA model; dynamic programming; price forecasting.

Abstract: There are considerable risks in the entire financial trading market. However, in the investment capital market line, the risk and return of financial transactions are usually proportional. To this end, market traders often buy and sell volatile assets to maximize total returns. In this paper, a decision-making model for a portfolio of three assets, cash, gold, and bitcoin, is constructed based on the daily trading prices so far, and the investment value of the assets is successfully determined, based only on the price of the day. The data gives the optimal daily trading strategy, which is very sensitive to transaction costs based on calculations.

1. Introduction

With the continuous development and improvement of financial technology and the progress of modern information technology, the field of financial transactions has been mainly expanded. However, due to the instability of the financial market and the spread of the global novel coronavirus epidemic, there are considerable risks in the entire financial trading market. In past research, the prediction of the dynamic price and volatility of the stock market has always been a challenge for researchers. With the continuous development of the financial market, it has gradually expanded from predicting stock prices to time series prediction of bond prices and derivatives prices. Bitcoin is the first decentralized digital currency not regulated by any central bank or authority. Originated in 2009, it became a trend in 2017. Much research has been done on Bitcoin price prediction in recent years. There are various factors affecting the price of Bitcoin.

Similarly, the gold market is also fluctuating all the time. After hitting a high point in 2011, the price of gold fell and gradually stabilized and rose at the end of 2015. After that, it fell into a period of consolidation that lasted for more than two years and began to rebound strongly in the second half of 2018. Finally, due to the spread and panic of the new crown epidemic, the price of gold hit a high record of $2,075 in August 2020 and has fallen since then. However, due to investments' high-risk and high-return properties, traders often buy and sell volatile assets to maximize total returns.

2. Date Processing

The data includes the daily Bitcoin settlement price from April 28, 2016, to August 14, 2018, which is 290 days. A total of 1265 data are used. About 75% are used for training, and the others are used for validation. Then the data from September 11, 2018, to September 10, 2021, is used for testing.

Prices have different dimensions and units, which will affect the results of data analysis. In order to eliminate the dimensional influence between prices, data standardization processing is required to solve the comparability between data. Then, for the prices of gold and bitcoin, we use a ratio between today's price and yesterday's price instead of today's price for subsequent predictions, which can make the data more impressive.

3. AnalysiS Establishment and solution of the model s and Modelling

3.1 Analytic Hierarchy Process (AHP)

Calculation of weight of each indicator based on Analytic Hierarchy Process (AHP)

Build a gradient model

Target layer: It is a bull market if the mean value is larger than the bull market evaluation index, and it is a bear market if it is smaller than the bull market evaluation index.

Criterion layer: The initial value of all indicators is set to 0. If the current calculation is a bull market, the previous quarter's value will be increased by 1; if it is a bear market, it will be decreased by one. The result is greater than 0 for a bull market, and less than 0 for a bear market

Scheme layer: The first 90-day average of gold's increase and the first 90-day average of gold's 15-day deviation rate

Construct judgment matrix

Firstly, a judge matrix $O = (A_{ij})_{n\times n}$ is created for the criterion layer. Secondly, judge matrices of

$$A_1 = (C_{ij})_{m\times m}$$

$$A_2 = (C_{ij})_{m\times m}$$

$$A_3 = (C_{ij})_{m\times m}$$

$$A_n = (C_{ij})_{m\times m}$$

Are created for the scheme layer.

Hierarchical single sorting and consistency check

First, each column vector of the data matrix is normalized, and then the rows are summed up to obtain an n-dimensional column vector w. After normalizing the vector w, the vector is used as the eigenvalue of the approximate eigenvector, and the eigenvalue is used as the weight of the indicator. When CR is smaller than 0.10, it means the consistency of the matrix can be accepted. If CR is no smaller than 0.10, it means that a modification to matrix elements should be made.

Hierarchical total ordering and consistency check

It is assumed that criterion layer B has n elements $A_1, A_2, \ldots, A_n$. Their weights are a_1， $a_2, \ldots, a_n$. Scheme layer P has m elements $C_1, C_2 \ldots C_m$. Their weights are based on a single criterion of $A_j$are $c_{1j}, c_{2j} \ldots c_{mj}$. The combined weight for the target is $\omega_1, \omega_2, \ldots, \omega_j$.

The combined weight of each element to the target can be calculated as

$$\omega_j = {}^{n}{}_{j=1}\, a_j c_{ij} \quad i=1，2，\ldots，m$$

Although the weights under every single criterion are consistent since the weights accumulate the combined weights under a single criterion, their inconsistencies may accumulate, so the combined weights should also be checked for consistency. Based on the consistency index CI(j) obtained above and its corresponding average random consistency index RI(j) (j=1, 2, ..., n), the consistency ratio CR is obtained[1].

$$CR = \frac{{}^{n}{}_{j=1}\, CI(j) a_j}{{}^{n}{}_{j=1}\, RI(j) a_j}$$

When CR<0.10, the consistency of hierarchical total ranking results is acceptable.

The calculated CI and CR are 0.1187 and 0.0957, respectively. The consistency ratio CR = 0.0957 < 0.1 meets the consistency requirements so that the previous weight vector can be used. Therefore, the weight of the first 90-day average of gold's increase is 0.674, and the first 90-day average of gold's 15-day deviation rate is 0.326.

In the same way, the weights of the first 90-day average of Bitcoin's increase and the first 90-day average of Bitcoin's 5-day deviation rate are calculated below. The gold purchase risk and the weight of gold's deviation rate are calculated similarly.

3.2 Bull and bear market model

Bull market evaluation indicator = the average value of the first 90 days of gold's increase × 0.674 + the average value of the first 90 days of gold's 15-day deviation rate × 0.326

The mean value greater than the bull market evaluation index is a bull market, and the mean value less than the bull market evaluation index is a bear market. Therefore, we can obtain the gold bull market indicator map and the gold bull market distribution map.

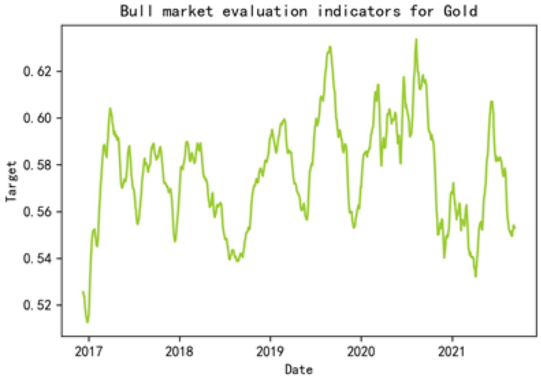

Figure 1 Gold bull market indicator map

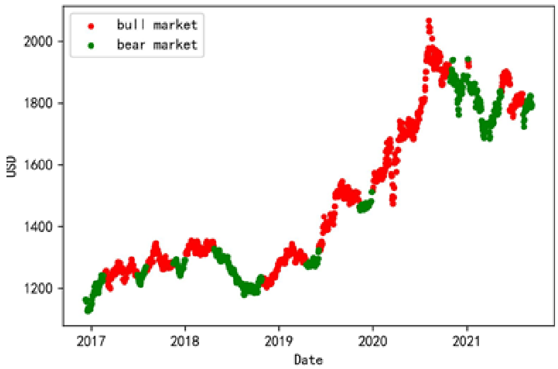

Figure 2 Gold bull market distribution map

According to the indicators, the time of bull and bear markets is determined by the voting method. In order to reduce the error, the initial value of all indicators is set to 0. If the current calculation is a bull market, the value of the previous quarter will be increased by 1, and if it is a bear market, it will be decreased by 1. The final result is greater than 0 for a bull market, less than 0 for a bear market, and the final distribution map of the gold bull market is obtained.

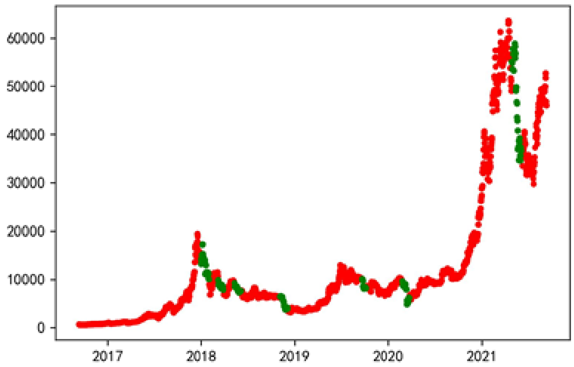

Figure 3 Final bitcoin bull market distribution map

3.3 Model of Investment risk

When the model is established, we add the 5-day moving average as the evaluation index of the investment strategy. Due to the large increase of Bitcoin and the small fluctuation of gold, we adopt the 5-day moving average as an essential trading evaluation indicator for Bitcoin and the 15-day moving average as an important transaction evaluation indicator for gold[2].

The gold purchase risk and the gold deviation rate are defined to be related to the gold bull market. The weight is determined by the AHP method, the weight of the gold bull market is 0.641, and the weight of the gold deviation rate is 0.359.

Gold buying risk = gold bull market × 0.641 + gold 15-day deviation rate × 0.359 (5).

The gold buying risk map can be calculated.

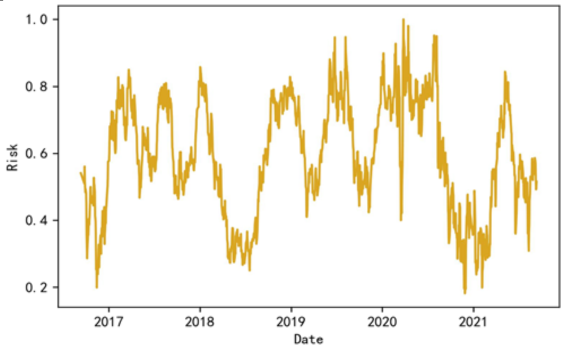

Figure 4 Gold buying risk map

In the same way, the Bitcoin purchase risk and the Bitcoin deviation rate are defined to be related to the Bitcoin bull market, and the Bitcoin purchase risk map is calculated.

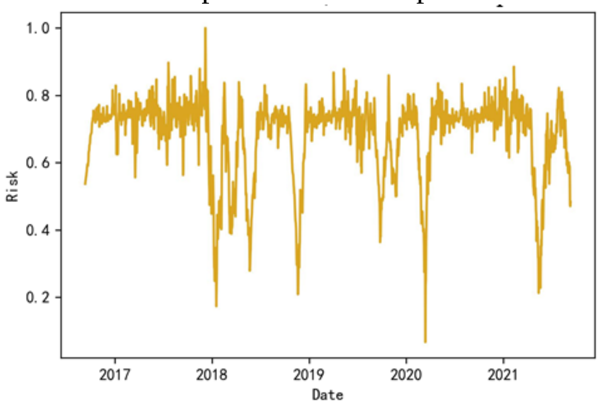

Figure 5 Bitcoin buying risk map

3.4 Time sequence prediction

For time sequence prediction, the traditional ARIMA model is used. The first step is to select the gold prediction parameters and check the gold trend chart and gold difference [3].

Table 1 Selection of gold prediction parameters

Date	USD (PM)	Value	Deal Day	Gold Rise	Bitcoin rise	Gold Average price of 15 days	…	Gold 15-day deviation rate	Bitcoin Average price of 5 days	Bitcoin 5-day deviation rate
2016/9/11	0	621.65	0	0.566144708	0.510080379	0		0.571622156	0	0.550392334
2016/9/12	1324.6	609.67	1	1324.6	−11.98	0		0.571622156	0	0.550392334
2016/9/13	1323.65	610.92	1	−0.95	1.25	0		0.571622156	0	0.550392334
2016/9/14	1321.75	608.82	1	−1.9	−2.1	0	…	0.571622156	0	0.550392334
2016/9/15	1310.8	610.38	1	−10.95	1.56	0		0.571622156	0.009833764	0.54535617
2016/9/16	1308.35	609.11	1	−2.45	−1.27	0		0.571622156	0.009793483	0.548616596
…					…					

ARIMA (p,d,q) model can be denoted as:

$$\left(1-\sum_{i=1}^{p}\emptyset_i L^i\right)(1-L)^d X_t = \left(1+\sum_{i=1}^{p}\theta_i L^i\right)\varepsilon_t$$

Where L is delay operator, $d \epsilon Z$, $d > 0$.

Differentiate the non-stationary time sequence as follows:

Where ∇ is the differential operator

$$\nabla^2 y_t = \nabla(y_t - y_{t-1}) = y_t - 2y_{t-1} + y_{t-2}$$

For delay operator B,

$$y_{t-p} = B^p y_t,\ \forall p \geq 1$$

Then we can get

$$\nabla^k = (1 - B)^k$$

Assuming that y_t is a non-stationary time sequence of order d, then $\nabla^d y_t$ is a stationary time sequence, which can be set as an ARMA p, q model as follows:

$$\lambda(B)(\nabla^d y_t) = \theta(B)\varepsilon_t$$

In which

$$\lambda(B)=1-\lambda_1 B-\lambda_2 B^2-\cdots-\lambda_p B^p$$

$$\theta(B)=1-\theta_1 B-\theta_2 B^2-\cdots-\theta_p B^p$$

Denotes Autoregressive Coefficient Polynomial and Moving Average Coefficient Polynomial. ε_t is a Zero-mean white noise sequence. The above autoregressive summation moving average model is ARIMA (p,d,q)[4].

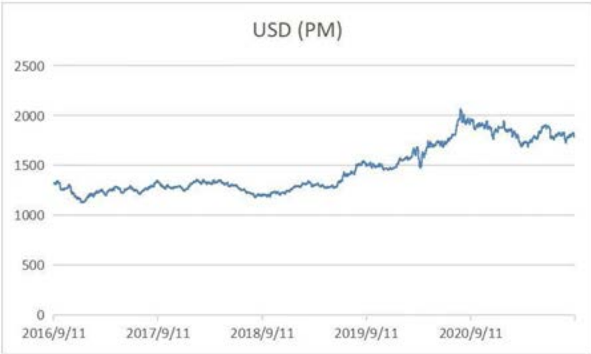

Figure 6 Gold trend chart

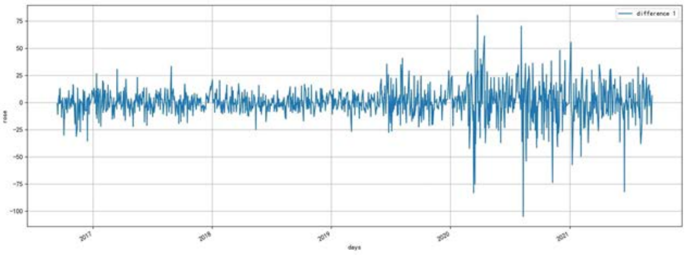

Figure 7 Gold differential chart of order 1

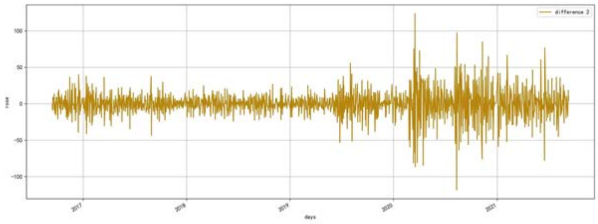

Figure 8 Gold differential chart of order 2

After that, we use the adfuller unit root to test the data stationarity, and the formula is as follows:

$$r_t = P_{i=1} \alpha_i r_{t-i} + \omega_t$$

$\alpha_1, \cdots, \alpha_p$ are all smaller than 1, which means the sequence is stationary. If any of α_i, is larger than 1, this sequence is non-stationary. To determine whether $\alpha_1, \cdots, \alpha_p$ are all smaller than 1, the eigenfunction of the AR model is applied as follows:

$$1-\alpha_1 x-\alpha_2 x^2-\cdots-\alpha_p x^p=0$$

It is tested whether there is a root greater than or equal to 1 tests whether the AR series is stationary. The adfuller test results are the borderline ADF values at 99%, 95%, and 90% confidence intervals. It can be observed that the null hypothesis is that it is not a stationary time series, and the P-value of the original data is > 0.05, so it does not meet the requirements of stationarity; the P-value of the first-order difference is < 0.05, and the T value is less than 1%, 5%, and 10%. The original hypothesis can be rejected significantly, indicating that the data is stationary. The first-order difference data has been stable, so there is no need to continue to make the second-order difference.

Based on the p-statistic of the chi-square distribution, the p-values of the first-order differences are very small, so the data reject the original hypothesis. That is, the data is not purely random. The best p and q can be selected by the exhaustive method.

max_ar and max_ma are the maximum possible parameters, and the value is 5. When training the ARIMA model, the three values in the order parameter are (p, d-order difference, q), respectively. When the result obtained by the exhaustive method is 0, 0, and the drawn graphic effect is poor, it can observe it according to personal observation.

4. Advantages and disadvantages of the model

4.1 Advantage

Reasonable assumptions: This article has established a series of scientific assumptions by reading much literature, ignoring the factors that have little influence on the results, and obtaining a good modeling effect under the condition of greatly simplifying the model and algorithm.

The scientific nature of modeling: The model is built with classical economic formulas, rigorous thinking, and high accuracy. It has certain advantages in sequence modeling problems and is simple to implement.

In the problem analysis process, the meaning of the problem is corrected and analyzed, and the particle swarm algorithm is used to optimize the data.

The accuracy of the model: This paper uses the dynamic programming method based on the established model, which has a high accuracy after testing.

4.2 Shortcomings

The change of w in the basic particle swarm algorithm used in the second question of the article is linearly related to the number of iterations, which will make the algorithm unable to solve complex and nonlinear optimization problems well.

This paper ignores factors other than the given conditions. The problem is single in constructing the mathematical model, which is not comprehensive enough. There may be deviations in the data statistics, so it brings errors to the data.

5. Improvements to the model

The more traditional ARIMA model is used in time series forecasting, but in essence, it can only capture the linear relationship, not the nonlinear relationship. If the Apriori algorithm and the neural network coupling model solve the problem, the results are more accurate, and the algorithm principle is simple and easy to implement. It is suitable for mining association rules of transaction databases and is more practical.

6. Promotion of the model

In the context of the increasingly rapid development of financial markets and commodity markets, the application of modern financial derivatives has become increasingly prominent. Our model is the best daily trading strategy model developed only based on the day's price data.

The application of machine learning in asset price forecasting and allocation can mine and utilize the information contained in financial data related to asset price changes. The information is used to make accurate future asset prices or profitable asset allocation decisions. It can provide better risk control and decision-making basis for investors who participate in China's financial derivatives market in the future. The combination of machine learning algorithms and financial theory will be an important direction for future research.

References

[1] Wang Xin. Based on the prediction effect of the stock option price on the future price of the underlying asset [D]. East China Normal University, 2013.

[2] https://mp.weixin.qq.com/s?src=11&timestamp=1645428872&ver=3633&signature=TLW0BccnQEWTO2lTIhc9ehNJmroWRMcBDdU10tYDOAsfqvvpk1RiL1EkXkuFPawUwd*-3EEeNbhFmKOQoqgoJIOdS7JjEvFzD0QYadjjdxwDtCx2g*1fVTs5C9DP5Tmz&new=1

[3] Zhao Qi, Xu Weijun, Ji Yucheng, Liu Guifang, Zhang Weiguo. Application of Machine Learning in Financial Asset price prediction and allocation [J]. Journal of Management, 2020,17 (11): 1716-1728.

[4] Ji Shouling, Du Tianyu, Deng Shiguang, Cheng Peng, Shi Jie, Yang Min, Li Bo. Review of the robustness studies of deep learning models [J]. Journal of Computer, 2022,45 (01): 190-206.

[5] Li Huiping, Qing Weiping, Wang Hong, Dong Qiaoyun, Chen Ruo. Financial Risk analysis of enterprise investment projects [J]. Journal of Guizhou University of Technology (Social Science edition), 2004 (04): 30-33 + 41.

[6] https://xuewen.cnki.net/CMFD-1019850606.nh.html