ISSN: 2688-9323

Research on Gold and Bitcoin Portfolio Investments

Nan Hu*

School of Management and Economics, Beijing Institute of Technology, Beijing, 102400, China

Abstract

In this paper, we establish an optimal model for trading gold and bitcoin markets. Firstly, we consider the fluctuation difference between gold and bitcoin, and use autoregressive integrated moving average (ARIMA) model and LSTM neural network to predict the market price of gold and bitcoin. In addition, we employ the dynamic programming method with the highest sharpe ratio to determine the optimal investment strategy, further replacing the forecasting algorithm in the model and giving a small range perturbation to the portfolio. Finally, the results of our analysis are verified by Monte Carlo simulation, and it is concluded that the trading strategies given by the models we consider can generate huge gains.

Keywords

Autoregressive Integrated Moving Average (ARIMA); LSTM Neural Networks; Investment Risk; Dynamic Programming.

1. Introduction

Investment, as one of the troikas of economic development, involves the accumulation of property for future benefits. One of the great advantages of gold as a wealth management product is that as a rare precious metal, it has its own value, and this value is very stable. For thousands of years, the value of gold has been affirmed by people. Unlike previous currencies, bitcoin relies entirely on the internet to create and trade [1]. The unique technical attributes of bitcoin derive its economic and cultural attributes beyond the current currency.

Bitcoin has a strong similarity to the gold market. Compared with gold, which has a long history and a relatively stable price trend, the market price of bitcoin is extremely volatile. At present, the international research on quantitative investment has achieved remarkable results. At the theory level, the time-series analysis method represented by the autoregressive integrated moving average model (ARIMA) plays an important role in econometrics [2]. On the other hand, with the continuous progress of artificial intelligence, market price forecasting has also experienced the evolution from the application of back propagation (BP) neural network and support vector machine (SVM) to random forest and deep learning [3-4]. At the application level, the ARIMA model requires the assumption of stationarity and linearity, so it is suitable for predicting price curves with small fluctuations and strong data correlation. The deep learning algorithm is good at solving nonlinear problem, but the neural network model trained on a small amount of data often has the problem of low prediction accuracy.

Our paper is designed to help traders in gold and bitcoin portfolios make decisions to maximize returns on their investments. In order to provide traders with trading strategies more reasonably and comprehensively, we establish a series of models to predict the prices of two volatile assets, quantify the relationship between returns and risks, and establish an investment risk assessment model to assist traders in making optimal decisions.

ISSN: 2688-9323

2. Price Forecast Model

2.1. ARIMA Model

The ARIMA model has three parameters, specifically p, the autoregressive coefficient, d, the differences, and q, moving average coefficient. The regressive formula of the model can be expressed as:

$$y_t = \mu + \sum_{i=1}^{p} \gamma_i y_{t-i} + \epsilon_t + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} \tag{1}$$

where γ_i *and* θ_i is the coefficients, and ϵ_t is the error.

2.1.1. Model Establishment

a) Time series stationary. The difference order is obtained by making the data stationary by difference processing. Then, as shown in the figure 1, through the ADF test, d=2 is calculated.
b) Determine the parameters p, q. The ACF and PACF tests were used to obtain the number of autoregressive items and the number of moving average items that the data fell into the confidence interval. As shown in the figure 2 and 3, p=13 and q=2 is judged by ACF and PACF test.
c) Use the ARIMA (13, 2, 2) model to obtain time series analysis results.

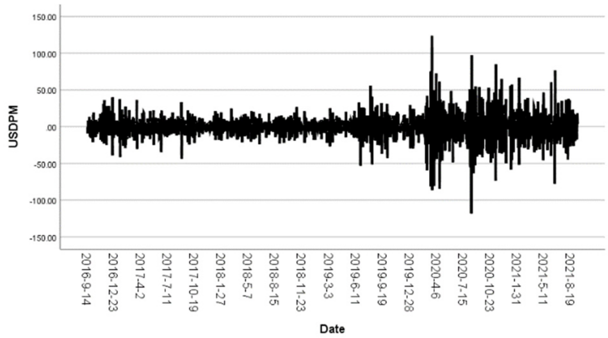


Figure 1. 2-times differences

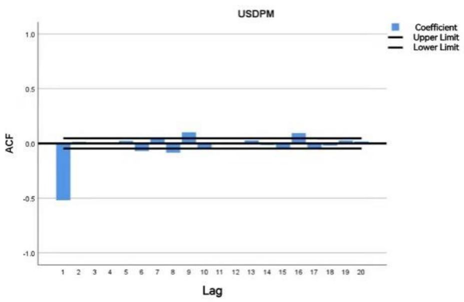


Figure 2. ACF test

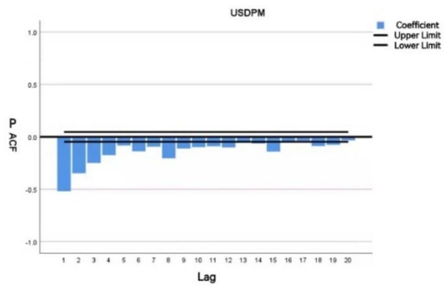


Figure 3. PACF test

2.2. LSTM Neural Network

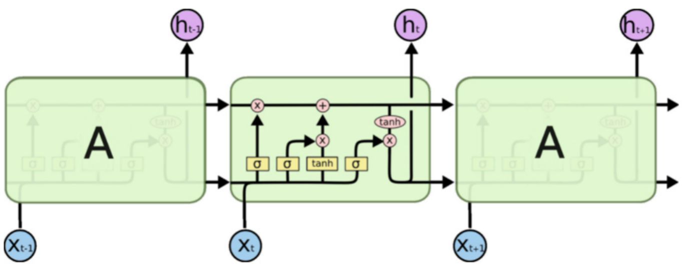


Figure 4. The architecture of LSTM units

ISSN: 2688-9323

LSTM is a recurrent neural network suitable for extracting temporal features from time series [5], with the ability to learn long-term time series dependencies. Its structure is shown in the figure 4.

There are three stages inside LSTM, namely:

1.Forgetting stage: Forgetting stage: Selective forgetting of the input c^{t-1} from the previous node as a forgetting gate z^f .

2.Selective memory stage: Selective memory of the input x^t by selecting memory gate z^i.

$$c^t = z^f \odot c^{t-1} + z^i \odot z \tag{2}$$

Add this part to the first part to get the new output resultc^t

3.Output stage: Through output gating z^0, the tanh activation function is c^t used to scale to obtain the output h^t. Converted y^t to h^t.

$$h^t = z^0 \odot tanh(c^t) \tag{3}$$

$$y^t = \sigma(W'h^t) \tag{4}$$

To evaluate the accuracy of the LSTM model, the commonly used indicators are Root Mean Squared Error(RMSE) and Mean Absolute Percentage Error (MAPE), the formula is as follows.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i' - y_i)^2} \tag{5}$$

$$MAPE = \frac{100\%}{n}\sum_{1}^{n}\left|\frac{y_i' - y_i}{y_i}\right| \tag{6}$$

where $y_i^{'}$ is the output result of the model at the date i, which means the predicted value of the price, and y_i is the actual value of the price at the date i.

For LSTM neural network training, we employ the widely used stochastic gradient descent optimizer (SGD). The adaptive moment estimation optimizer (adam, adaptive moment estimation) is also a more common optimizer. We use the above optimizers to predict the model and calculate the corresponding RMSE and MAPE. It can be obtained from the above data that when the adam optimizer is used, the model achieves the smallest RMSE and MAPE values.

3. Investment Risk Assessment Model

3.1. Model Establishment

When assessing the risk of volatile assets between gold and bitcoin, in order to make full use of the data, we calculated the volatility, dispersion coefficient, deviation rate, and maximum drawdown rate of gold prices and bitcoin prices respectively. calculate the weights of the four indicators, and the weighted summation will get the investment risk of gold and Bitcoin. The flow chart is shown in the figure 5:

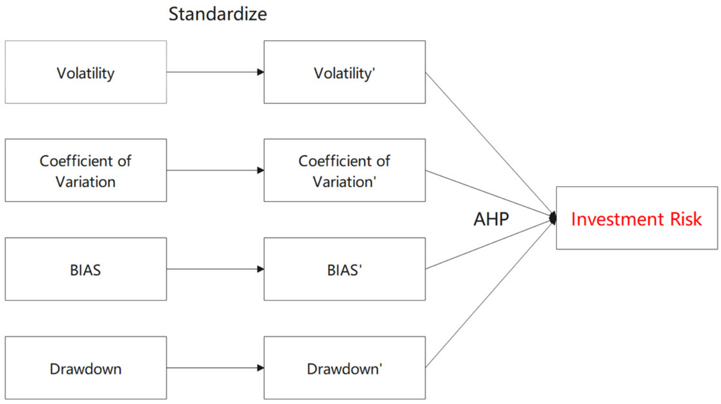


Figure 5. Investment risk calculation flow chart

In our model, we use parkinson's historical volatility (HL) method to calculate volatility . In order to reflect the volatility of gold and Bitcoin more accurately, when calculating the volatility of gold, we choose 15 days as a cycle, and when calculating the volatility of Bitcoin, we choose 5 days as a cycle.

The formula for calculating volatility is

$$Volatility_{Parkinson} = \sigma_P = \sqrt{\frac{F}{N}}\sqrt{\frac{1}{4Ln(2)}\sum_{i=1}^{N}(Ln(\frac{h_i}{l_i}))^2} \tag{7}$$

The formula for calculating the dispersion coefficient is

$$\text{Coefficient of Variation}=\frac{\text{Standard Deviation}}{\text{Average Value}} \tag{8}$$

The formula for calculating the deviation rate is

$$BIAS = \frac{P-AP}{AP} \times 100\% \tag{9}$$

where P is the closing price of the day, and AP is the average price of N days .

The formula for calculating the maximum drawdown rate is

$$Drawdown = \frac{max(D_i-D_j)}{D_i} \tag{10}$$

where D is the net value i of a certain day, is a certain day, and j is i a certain day after.

The data normalization formula is expressed as:

$$X_i = \frac{x_i}{\sum_{i=1}^{n} x_i} \tag{11}$$

The data is normalized so that the influence of dimensions is eliminated after processing, making the value-at-risk added by weights more meaningful.

We use the analytic hierarchy process to find weights. The decision-making problem is decomposed into three layers, the top layer is the target layer O, which calculates the respective investment risks of gold and bitcoin, the middle layer is the criterion layer, including four influencing factors C1 (volatility'), C2 (discrete coefficient) '), C3 (deviation rate'), C4 (maximum drawdown rate'), the bottom layer is the solution layer, namely P1 (gold investment risk), P2 (bitcoin investment risk).

Constructing the judgment matrix OC, comparing the four elements in the criterion layer C in pairs to obtain a pairwise comparison matrix, which is shown in table 1.

Table 1. Comparison matrix

O	C1	C 2	C3	C4
C1	1.0000	0.2857	0.2222	0.2500
C 2	3.5000	1.0000	0.6667	1.0000
C3	4.5000	1.5000	1.0000	2.0000
C4	4.0000	1.0000	0.5000	1.0000

Solving the eigenvalue of OC, which is easy to solve $\lambda_{\max} = 4.0286$. The weights calculated according to the arithmetic mean method, the geometric mean method and the eigenvalue method respectively take the arithmetic mean value to obtain the weight vector, which is $\omega_i = (0.0758, 0.2621, 0.4084, 0.2537)^T$ calculated $CR = 0.0107 < 0.1$ according to $CR = \frac{CI}{RI}$, $CI = \frac{\lambda_{\max} - n}{n-1}$ and passed the consistency test. We have obtained the investment risk assessment formula:

$$Investment\ \ Risk = 0.0758 \times C1 + 0.2621 \times C2 + 0.4084 \times C3 + 0.2537 \times C4 \tag{12}$$

3.2. Model Analysis

We use the existing data to calculate the investment risk of gold and bitcoin respectively, and draw the distribution map as figure 6 and 7.

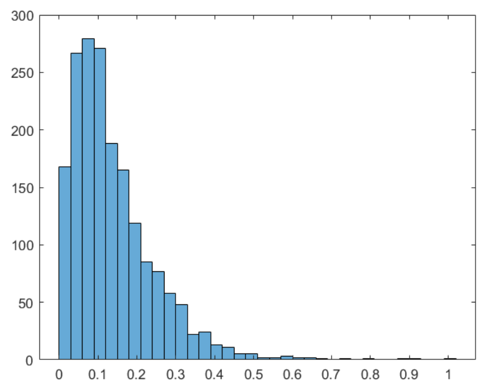


Figure 6. Gold investment risk statistic

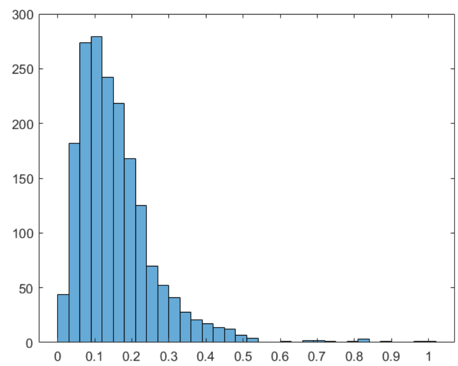


Figure 7. Bitcoin investment risk statistics chart

The results show that the investment risk of gold and bitcoin is similar to the chi-square distribution diagram. Combined with the knowledge of mathematical statistics, we give the investment risk assessment table as table 2.

ISSN: 2688-9323

Table 2. Investment risk assessment

Investment Risk	Prudent	Intermediate	Aggressive
Gold	<0.05	0.05~0.23	>0.23
Bitcoin	<0.03	0.03~0.27	>0.27

For gold, when the model solution result is less than 0.05, we recommend an active or aggressive investment method. When the model solution result is greater than 0.23, we recommend a conservative or prudent investment method. When the result is between 0.05 and 0.23, we take a balanced investment approach.

For Bitcoin, when the model solution result is less than 0.03, it is recommended to adopt an active or aggressive investment method. When the model solution result is greater than 0.27, it is recommended to adopt a conservative or stable investment method. When the result is between 0.03 and 0.27,we take a balanced investment approach.

In addition, we use the given data to calculate the rate of return, and map the rate of return to the value at risk one-to-one as figure 8 and 9.

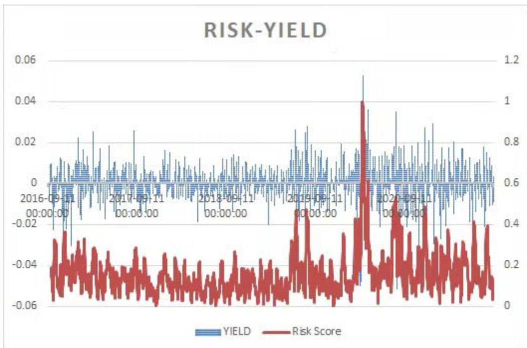


Figure 8. Gold risk-return

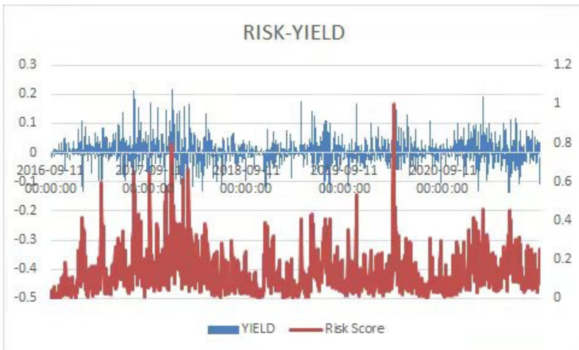


Figure 9. Bitcoin risk-return

4. Quantitative Trading Strategy based on Dynamic Programming

4.1. Model Establishment

In our model, we set the objective function $f(i)$ to be the Sharpe ratio on day i+3 after adjusting for gold and bitcoin holdings on day *i*. we think when $f(i)$ reaches the maximum, the maximum value of the Sharpe ratio can be achieved, thereby achieving the maximum rate of return per unit of risk taken .

After the transaction fee is settled, the proportion of cash, gold and bitcoin in the total assets on that day is:

$$[c_i-(1\pm0.01)\Delta g_i-(1\pm0.02)\Delta b_i,(1+\overline{G_i})g_i+\Delta g_i,(1+\overline{B_i})b_i+\Delta b_i] \tag{13}$$

Similarly, we normalize the assets, and multiply the sum of the three proportions by the total amount of assets yesterday, so as to obtain today's total amount of assets and rate of return and today's proportion.

In order to achieve the maximum profit, we can get the following constraints:

$$\begin{cases} c_i+(1\pm0.01)\Delta g_i+(1\pm0.02)\Delta b_i\geq0 \\ (1+\overline{G_i})g_i-\Delta g_i\geq0 \\ (1+\overline{B_i})b_i-\Delta b_i\geq0 \end{cases} \tag{14}$$

ISSN: 2688-9323

4.2. Model Analysis

Trading according to the strategy given by our model, by September 11, 2021, \$1,000 can appreciate to \$7,354.627. The data is shown in figure 10.

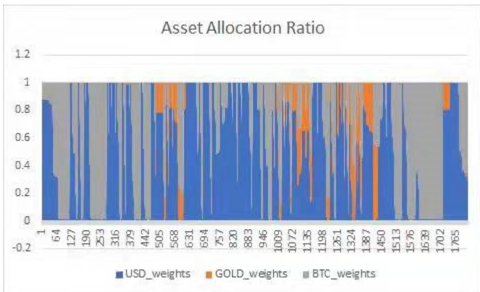


Figure 10. Asset allocation ratio

Therefore, we use intelligent algorithms to help speed up the convergence, and continue to plan the model under the condition of initial values.

Monte Carlo simulation can directly deal with the uncertainty of risk factors: express the uncertainty in the form of probability distribution, establish a stochastic model for risk decision-making, conduct sampling experiments on random variables, and analyze the simulation results, not only can the decision-making target output be obtained, and other statistics can also be given as a probability distribution. Stochastic sensitivity analysis of target outputs quantifies the magnitude and priority of the impact of risk inputs on target outputs.

According to the relevant knowledge of probability theory, if the expected μ and variance of a risk variable X σ^2 exist, and $\sigma \neq 0$ (otherwise X is a non-risk variable), the error of the Monte Carlo simulation is

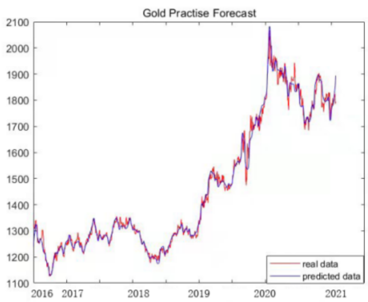
$$\varepsilon = \frac{\lambda_\alpha \sigma}{\sqrt{N}} \tag{15}$$

Among them $\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\lambda_\alpha} e^{\frac{1}{2}t^2}\, dt = 1 - \frac{\alpha}{2}$, λ_α is the normal difference, α is the significance level, N is the number of samples, σ is the standard deviation. In practice, it is usually replaced by the sample standard deviation δ_s.

The error in the above formula is ε, only related to the standard deviation σ and the sample size N, but has nothing to do with the space where the sample elements are located, that is, the convergence speed and probability of the Monte Carlo method have nothing to do with the dimension of the problem, which is very meaningful for solving multidimensional problems.

To provide that our model provides the best strategy, we modified the trading strategy/our model in the following two ways respectively [6].

Method 1: Replacing the price prediction model



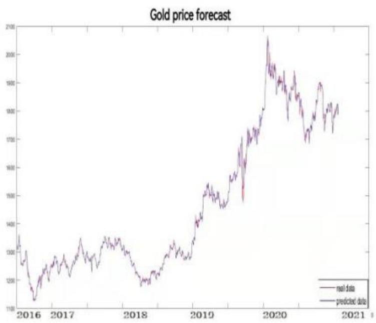


Figure 11. LSTM-based price forecast model results **Figure 12.** Arima price forecast model

ISSN: 2688-9323

Firstly, we replaced the arima model used to predict the price of gold with the model used to predict the price of bitcoin -- an LSTM-based price prediction model. We plotted the results predicted by the two models and compared the two graphs together as figure 11 and 12.

Next, we replaced the model used to predict the price of bitcoin, the LSTM-based price prediction model, with the arima model used to predict the price of gold. We plotted the results predicted by the two models and compared the two graphs together as figure 13 and 14.

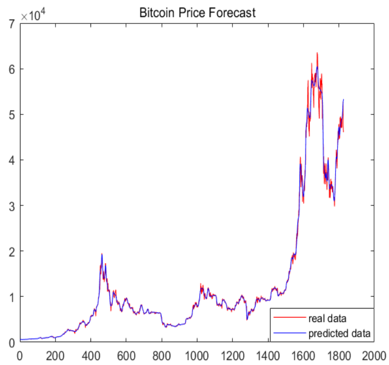


Figure 13. Prediction result of arima price prediction model

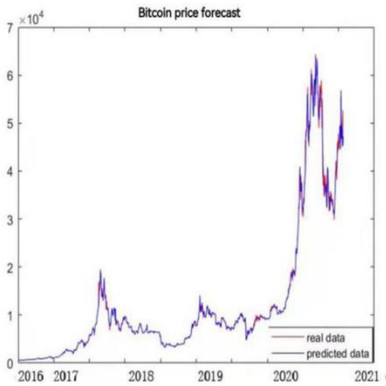


Figure 14. Prediction result of price prediction model based on LSTM

So far, we have concluded that using the Arima model to predict the gold price is better than using the LSTM-based price prediction model, and using the LSTM-based price prediction model to predict the Bitcoin price is better than the Arima model.

5. Conclusion

In this paper, we have successfully established a risk assessment model and a price prediction model through AHP, LSTM neural network algorithm and ARIMA model. Then, guided by these two models and the Monte Carlo method, the trading strategies are successfully simulated, proving the high efficiency of our model. Most meaningfully, this model has important implications for traders. When traders face a complex market environment, a guided quantitative investment model will significantly improve their work efficiency.

References

[1] Binjing Li, Keli Xie. (2020). LSTM-Based Quantitative Trading Using Dynamic K-Top and Kelly Criterion? International Joint Conference on Neural Networks (IJCNN). 4-6.

[2] Yulin Du. (2018).Application and analysis of forecasting stock price index based on combination of ARIMA model and BP neural network? 2018 Chinese Control And Decision Conference (CCDC). 12-13.

[3] Yujie Fang, Juan Chen, Zhengxuan Xue. (2019). Research on Quantitative Investment Strategies Based on Deep Learning? Algorithms 2019. 12(2). 28.

[4] Ezat Ahmadzadeh, Hyunil Kim, Ongee Jeong, Namki Kim, Inkyu Moon.(2021). A Deep Bidirectional LSTM-GRU Network Model for Automated Ciphertext Classification. IEEE Access 10: 3228.

[5] Mei-Li Shen, Cheng-Feng Lee, Hsiou-Hsiang Liu, Po-Yin Chang, Cheng-Hong Yang. (2021). Effective multinational trade forecasting using LSTM recurrent neural network. Expert Syst. Appl. 182: 138.

[6] Hans Weytjens, Enrico Lohmann, Martin Kleinsteuber. (2021). Cash flow prediction: MLP and LSTM compared to ARIMA and Prophet. Electron. Commer. Res. 21(2): 374-376.