Quantitative Portfolio of Gold and Bitcoin with Synthesized Prediction Models

Xuan Wu[1], Zhinan Xie[2]
wuxuan249@stu.shmtu.edu.cn[1], 201110006@cupl.edu.cn[2]

Shanghai Maritime University[1], China University of Political Science and Law[2]

Abstract. Quantitative portfolio of gold and bitcoin investment can be determined by synthesized quantitative model, with the help of various quantifying indicators. Previous prediction models labor to disperse risk of the investment portfolio as well as maximize the return. To handle it, this essay digs into a synthetic quantitative model of diversified regressive and prediction model, based on an enhanced ARIMA model, and this essay also integrates Analytic Hierarchy Process weight analysis and Monte Carlo model into it. In stage one, the daily value of the portfolio in a given time is predicted by enhanced ARIMA model, then the AHP method helps calculate the weight of the price rise and fall (%) of it, which leads to the establishment of two investment evaluation models on risk and income, respectively. In stage two, the judgment vector of whether to buy or sell was quantified through the ratio relationship. Eventually, the quantity of the two assets is determined by iterations, the accuracy and sensitivity of judgment of bitcoin and gold trading are quantified precisely, and decision-making is made with a slighter influence from inter-group conflicts.

Keywords. Time-series analysis; Autoregressive prediction; Statistical simulation

1 Introduction

Quantitative analysis greatly reduces the effect of investor sentiment volatility [1], and avoid making irrational investment decisions under the condition of the feverish or pessimistic market. By integrating statistical and predicted model into quantitative strategies, prediction of investment portfolio of gold and bitcoin can be achieved with high accuracy and wide applicability.

When facing portfolio of assets with discrepant nature, accessible and applicable data lacks. Market traders always adopt obtain excess return on investment strategies from a large number of historical data with the help of computers when buying and selling volatile assets [2]. Established models like ARIMA mainly predicts the future unknown data through the existing historical data. Worse still, the time spent using all historical data for time series prediction [3] will be getting longer. Time series data has difficulty in capturing nonlinear relations and inter-group relations [4]. Consequently, it is hard for a single prediction model to disperse risk of the whole investment portfolio and maximize the investment return.

We predicted the decision-making indicators of gold and bitcoin investment. By comparing the predicted time with the real time, the annual average absolute percentage error [5] is very low, which shows the accuracy of time series model. Second, we compare the differences between the four indicators we chose, and made a detailed analysis of the rationality of the indicators.

MSIEID 2022, December 09-11, Chongqing, People's Republic of China

DOI 10.4108/eai.9-12-2022.2327582

Finally, this paper compares four derivative models respectively with the decision model, concluding that the mathematical model can propose the best trade strategy. Last but not least, we evaluate the sensitivity of the synthetic model by comparing the amount of final total assets under different trade costs.

Our detailed contributions are as follows:

(1) Using the mixture of sliding window and non-sliding window in ARIMA, curtailing the average relative error between the predicted price and the actual value.

(2) Consideration of inter-group relations of regressive data, wiping out the conflicts between linear algorithm and non-linear data.

(3) Great sensitivity and accuracy, adapting itself to assets of different nature.

2 Proposed approach

2.1 Assumptions & Nomenclature

We put the symbols that we use in the model and explanations in Table 1.

Table 1. Symbol and explanations

Symbol	Explanation
αgold	Gold Trading Commission Cost Percentage.
αbitcoin	Bitcoin Trading Commission Cost Percentage.
Si(x)	The *i*-th spline function.
AIC	Akaike Information Guidelines.
k	Number of model parameters.
L	Likelihood function.
Rb (i)	Bitcoin share held on day i.
Rg(i)	Gold shares held on day i.
Rc(i)	Cash share held on day i.
Sb(i)	Bitcoin share sold on day i.
Sg(i)	Gold shares sold on day i.
Bb(i)	Bitcoin shares bought on day i.
Bg(i)	Gold shares bought on day i.
a	Trade cost of bitcoin.
b	Trade cost of gold
d	Judgment vector for whether to buy or sell

2.2 Spline interpolation

Spline interpolation is a method of approximating data using multiple low-order polynomials (Functions simulated using data) that pass through all data points.

The interpolation conditions of cubic spline interpolation are:

(1) The function value of the simulated function at the known point is the same as the real value;

(2) The proposed piecewise function is second-order continuous, that is, the derivative and the second derivative are equal at the junction of segments;

(3) We need to know the second-order derivative between the endpoints a and b of the interval, or the variation law of the second-order derivative at these n+1 points.

Denote the spline function of the i−th segment as Si(x), and the corresponding function value at the end point is yi. The spline function and its differential form are as in Eq. (1), Eq. (2) and Eq. (3).

$$S_i(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3 \tag{1}$$

$$S'_i(x) = b_i + 2c_i(x - x_i) + 3d_i(x - x_i)^2 \tag{2}$$

$$S''_i(x) = 2c_i(x - x_i) + 6d_i(x - x_i) \tag{3}$$

The coefficients ai, bi, ci, and di are solved by constructing matrix equations through data points and end-point conditions.

2.3 Baseline models

We decided to use the mixed ARIMA model of sliding window and non-sliding window to solve the problem. To implement the steps followed by Auto ARIMA, we implement the 3 to 7 steps of the original steps with the method of AIC automatic ordering.

The AIC (Akaike Information Criterion) value is an estimator used to compare models and is a measure of how well a statistical model fits. It builds on the concept of entropy and provides a measure of the trade-off between the complexity of the estimated model and the goodness of fitting the data. Typically, AIC is defined as in Eq. (4):

$$AIC = 2k - 2ln(L) \tag{4}$$

where k is the number of model parameters and L is the likelihood function. When choosing the best model from a set of available models, the model with the smallest AIC is usually chosen. The order is automatically determined using AIC.

Then, on the basis of the AHP model, we deduce the judgment vector of whether to buy or sell. After making price predictions for Bitcoin and Gold, we derived four original sets of indicators based on the only price indicators:

Gold, Bitcoin prices rise and fall (%);

The rate of change of gold and bitcoin prices (rate of change);

15-day average price of gold, 5-day average price of Bitcoin;

Due to the large fluctuations of Bitcoin, the five-day average price is best used as an important transaction evaluation indicator for Bitcoin, and the 15-day average price is best as an important transaction evaluation indicator for gold.

Gold 15-day deviation rate, Bitcoin 5-day deviation rate (BIAS);

We establish AHP investment risk evaluation model and return evaluation model for gold and bitcoin respectively.

And take the vector as the main basis for decision-making, combined with Monte Carlo method to solve the problem, by which way can the optimal trading strategy be made. According to the law of large numbers, when the sample size is large enough, the occurrence frequency of an event is its probability.

2.4 Price forecasting indicators

To predict the price of gold, we include four critical indicators:

(1) Margin of fluctuating price

$$\text{Margin of fluctuating price} = \frac{today's\ price - yesterday's\ price}{yesterday's\ price} \tag{5}$$

(2) Fluctuating rate

$$\text{Fluctuating rate} = \frac{today's\ margin - yesterday's\ margin}{|yesterday's\ margin|} \tag{6}$$

(3) Average price on 5 or 15 days

(4) 5 or 15 BIAS

Regarding the deviation degree relative to the most commonly used 5-day average, 5-day divergence rate (5BIAS), its formula is:

$$5BIAS = \frac{today's\ price - Average\ price\ on\ 5\ days}{Average\ price\ on\ 5\ days * 100} \tag{7}$$

The selection of indicators is based on economic theory and established under the support of statistical law.

2.5 Model variants

In order to better verify the accuracy of the synthesized model, we will compare the decision-making model with the final share of total assets under the conditions of "all buy and all sell", "increase of original data" and "random judgment vector".

The original situation refers to the use of the rise and fall of real data for judgment.

The following figures are the flow of Rc cash assets in these three cases and the flow of Rc cash assets of our model.

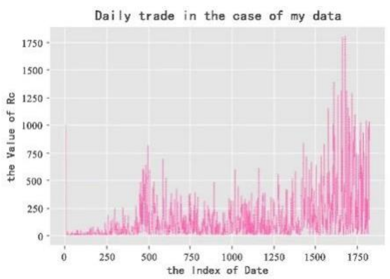

Fig. 2. Our model

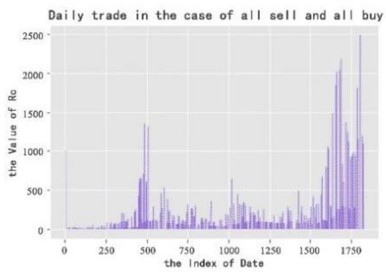

Fig. 3. 'Buy All, Sell All' Situation

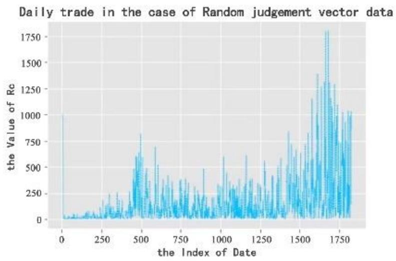

Fig. 4. The ‘random vector’ case

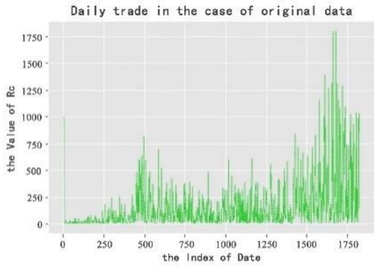

Fig. 5. The ‘real data’ situation

The figure below shows the column chart of total assets in the four cases. It can be shown that our decision-making model has a good effect.

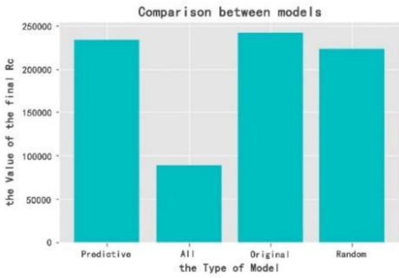

Fig. 6. Comparison between models

3 Experiments

The experiments section includes an introduction of the experimental scenario and datasets used for implementation of our quantitative model, an experimental setup, containing inductive learning settings, some baseline models, and the analysis of the experiment results.

3.1 Datasets

The experiment was conducted under a real quantitative trading scenario: A investor starts with \$1000 from 9/11/2016 to 9/10/2021. On each trading day, the trader will have a portfolio consisting of cash, gold, and bitcoin [C, G, B] in U.S. dollars, troy ounces, and bitcoins, respectively. The commission for each transaction (purchase or sale) costs: $\alpha_{gold} = 1\%$ and $\alpha_{bitcoin} = 2\%$. The following table shows the statistics on the datasets.

Table 2: Statistics on the datasets

Datasets	Starting Date	Ending Date	Max Value	Min Value	#raw	#prediction
Gold	9/13/16	8/31/21	2067.15$	1125.7$	1265	500
Bitcoin	9/13/16	8/31/21	59964.87	607.18	1826	720

3.2 Implementation and results

The price prediction results of Bitcoin on the two days are shown in Table 3.

Table 3: Bitcoin price on 11/9/2016 and 11/10/2016

Results	11/9/2016	11/10/2016
Actual Price	720.93	721.5
Predicted Price	719.01	720.73

Figure 7 shows the trend of predicted bitcoin value in the next two days.

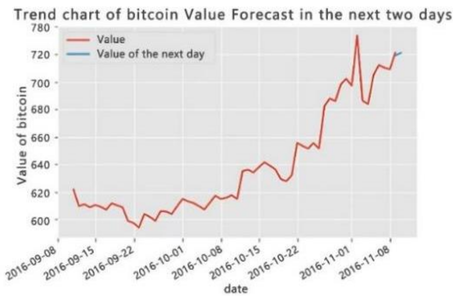

Fig. 7. The forecast trend chart

For the implementation of the entire prediction model, the non-sliding window part shows the price prediction results of gold and bitcoin are shown in Figure 8 and Figure 9.

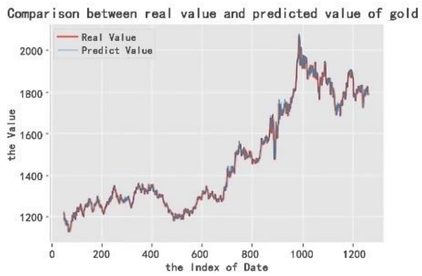

Fig. 8. Comparison in gold

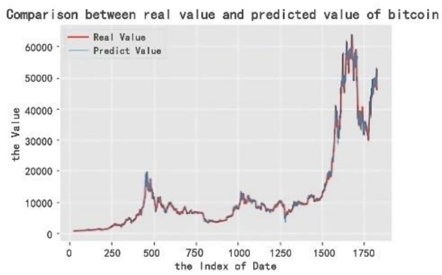

Fig. 9. Comparison in bitcoin

It can be roughly seen that the prediction model has reached a good level.

On the basis of the Monte Carlo model, we list constraint equations.

The objective function can achieve the optimal effect, after calculating, the highest value is 233055.59$.

3.3 Sensitivity analysis

To test the sensitivity of the model, we first dissect the sensitivity of the model by setting the costs of bitcoin and gold as a and b, respectively, and by changing the values of a and b to observe the range of changes in the model. As can be seen from the figure, the final asset is inversely proportional to the cost of the two.

Second, we examine the sensitivity of the model to buying and selling. After parameter adjustment, it is found that when the selling value is set between 0 and Rb(i), i = 1, 2, ..., 1826, the total assets can only reach a maximum of 4.3248 dollars. Therefore, it can be seen that the sensitivity of the model is high. When there are more sales and less purchases, the total assets show an upward trend.

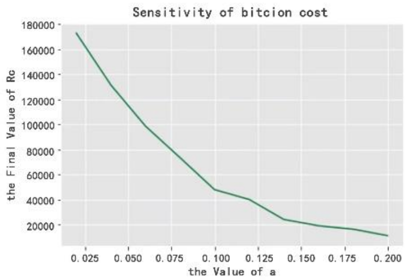

Fig. 10. The 'random vector' case

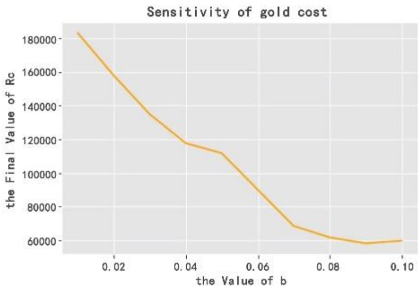

Fig. 11. The 'real data' situation

4 Conclusion

This essay proposed a synthesized prediction model of quantitative strategy on portfolio of gold and bitcoins. The model enhanced the traditional ARIMA, with the integration of AHP method, and was eventually determined by Monte Carlo algorithm. By using little historic data, the model supports more flexible investment portfolio, and reaches quicker convergence with less iteration times. It shows wider applicability and generalizability of quantitative models and improves reliable reference value for trade decision making.

References

[1] LI Wen (2020) The impact of investor sentiment on Stock market Returns: An empirical analysis based on the "Stock Market Crash" in 2015. MSc thesis, Harbin Institute of Technology

[2] Michael Di, Ke Xu (2022) COVID-19 vaccine and post-pandemic recovery: Evidence from Bitcoin cross-asset implied volatility spillover. In: Finance Research Letters, Volume: 50

[3] Yang Wang, Lixin Han (2021) Adaptive time series prediction and recommendation. In: Information Processing & Management, Volume 58, Issue 3

[4] Zbigniew Karpinski (2017) Advances in the attraction model for inter-group relations. In: Mathematical Social Sciences, Volume 89, pp: 109-118

[5] Arnaud de Myttenaere, Boris Golden, Benedicte Le Grand, Fabrice Rossi (2016) Mean Absolute Percentage Error for regression models. In: Neurocomputing, Volume 192, pp: 38-48