

Research on Fixed Proportion Investment Strategy based on ARIMA and LSTM

Jiangqi Zhao, Yipei Liu, Yingying Zhu

Nanjing Tech University, Nanjing, Jiangsu, 211816, China

Abstract

With the enormous changes in investment in the global financial industry, quantitative investment is playing an increasingly significant role in traders' life. Investors are paying more and more attention to this problem, that is, how to avoid risks while excavating the investment value of financial product portfolio, to grasp the future price trend of financial products, and to formulate strategic market trading plan, so as to accurately maximize profits. In order to address the above-mentioned challenge, we first preprocess and analyze the data. Then, we innovatively employ the ARIMA algorithm by using BOX-COX, thus greatly increasing the accuracy of prediction. At the same time, we also tried the LSTM method to make predictions. Finally, we construct a model that utilizes the predicted data to select the corresponding threshold for proportional investment.

Keywords

Quantitative Investment; ARIMA; LSTM.

1. Introduction

Currently, there are the huge changes in global investment market, the wave of quantitative investment has swept over, attracting the extensive attention from a large number of investors and speculators. The diversification of financial investors as well as financial products has continuously promoted the prosperity of the financial trading market.[1]

In a variety of financial products, gold has been regarded as the income source of the stock market, while the birth of bitcoin is the innovation of monetary and financial field in the Internet information age. Nevertheless, the attribute of bitcoin is different from statutory currency or traditional financial assets (Gold). Gold has low yield and low volatility, while bitcoin has high yield and high volatility [2]. The high volatility of bitcoin may lead to greater potential financial risks in the trading market [3]. Markowitz utilized the mean-variance model to quantify the income and risk, and concluded that portfolio can effectively reduce the risk. Therefore, it is of great practical significance to study how to explore the investment value of financial product portfolio while avoiding risks, and formulate strategic scheme for market transactions in order to maximize total revenue.

2. Data Preprocessing

In order to determine the investment scheme based on gold and bitcoin prices, we will complete or delete the 40 missing values in different cases. For example, when predicting future gold prices, we will make up for missing prices on the basis of the previous trading day. Note that though the price vacancy caused by the weekend market closure of gold does not belong to the missing value, but we also utilize the above method to work with it. More specifically, this method can improve the authenticity of the prediction, and it is also in line with the actual situation. When conducting investment transactions, we think the missing data is due to gold's weekend and holiday, so we delete it without considering trading.

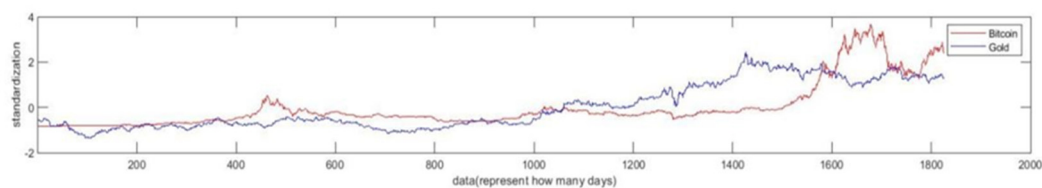


Figure 1. Can more intuitively reflect the price fluctuations of bitcoin and gold

3. Price Prediction Models Based on ARIMA and LSTM

We know that having more information in the trading market means that there may be greater profits. If we can predict future changes in financial commodity prices, we can know when to buy (at low prices) and when to sell (at high prices), in order to obtain profits. Shock earning is to point to the price when selling a shock minus the price buys multiplied by number of shares. In fact, the price fluctuation of financial products is a random process, and we can't predict its future trend completely. But that doesn't mean price movements are absolutely unpredictable. Its price change process is a set of time series data actually. It not only reflects the random price changes, but also contains operating rules of the system. Therefore, we first forecast the price in the next few days of the trading day. Since the prices of gold and bitcoin are affected and restricted by many factors, their changes are disorderly. However, the current popular investment theory is often based on a series of statistical characteristics. Previous studies have proved that this method has certain limitations, and the results may have large errors. The next, due to bitcoin has the prices with frequent fluctuations, many influencing factors and weak regularity, the real historical price data can be used to predict the future is also very few. In addition, the prediction of such financial products will directly affect the investment decisions of investors, related to the vital economic interests of investors, so the accuracy of the prediction is also important.

Based on this, we use ARIMA Forecast Model and LSTM (Long Short-Term Memory) Forecast Model respectively conduct short-term price forecasts and revise the results real-time.[4]

This result is used as an important basis for traders to make decisions on whether to purchase, hold or sell on the day of trading. At the same time, it is necessary to compare the predicted data with the actual data to verify the accuracy of the forecast model.

ARIMA model is also called Autoregressive Moving Average model, which is mainly applied in the short-term prediction of time series variables. Since the single time series value is unpredictable, but the overall time series value has certain disciplines, the ARIMA model is established to realize the short-term prediction of time series value. Based on the characteristics of this model, we select the price data on the day of the trading day and the past long period to predict the price of the next day. In this section, we combined with the actual situation, selected 30 days as a forecast sample.

Firstly, we perform box-cox transformation on the prediction samples. This method is effective for many actual data, and is suitable for the case that continuous response variables do not meet the normal distribution. It can reduce the unobservable error and the correlation of prediction variables to a certain extent, and significantly improve the normality, symmetry and variance equality of data. In the prediction sample, we carry out sliding window prediction and solve the trend function. Compared with the other method, the accuracy of these operations is higher. Subsequently, we perform the stationary ADF test. If the test results are not stable, it is necessary to repeatedly change through difference until the original sequence tends to be stable. Next, we calculate the partial autocorrelation coefficient PACF and autocorrelation coefficient ACF, and determine the order of the model by combining the AIC criterion. The least square method is used to estimate the parameters of the model, conduct the rationality test, and

restore the difference. The data is predicted by the generated model and compared with the actual data.

While n is equal to 5, the results are as follows:

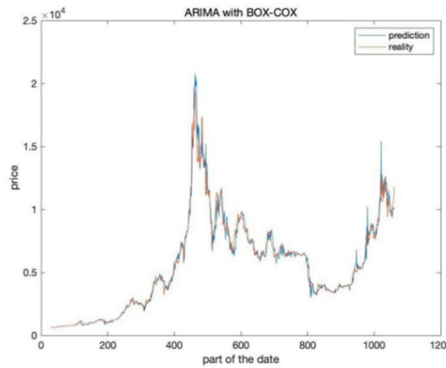


Figure 2. Forecast Bitcoin Price Using ARIMA with BOX-COX Method

The prediction accuracy of some bitcoin data in this topic reaches 95.19%. This is unprecedented.

Based on the above research and analysis, it can be found that the ARIMA model is ideal for the price prediction of gold and bitcoin, and has a good judgment for the prediction of price trend. Then we construct LSTM price forecasting model.[5] Compared with the RNN recurrent neural network, LSTM long-term and short-term memory network can avoid some errors such as the gradient value Nan in the calculation process and improve the computational efficiency. As a kind of time recurrent neural network, it can accept more extensive time series input, and is more suitable for processing nonlinear time series information. It has certain exploration and practical significance for market price prediction in the financial field. In the price prediction problem of gold and bitcoin, we set up a single dimension to single step prediction method, that is, using the price of the previous n days to predict the price of the $n+1$ day.

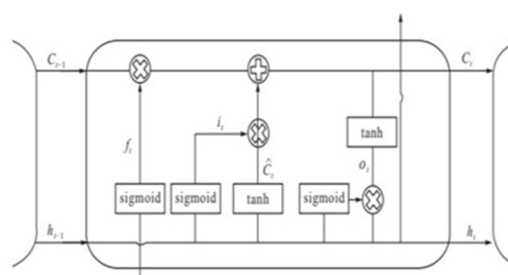


Figure 3. LSTM neuron structure

The results are as follows:

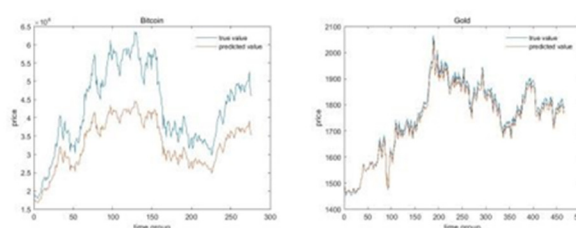


Figure 4. Forecast Bitcoin & Gold Price Using LSTM

4. Fixed Proportion Investment Strategy based on Threshold

By studying the literature on quantitative investment strategies, this paper improves and simulates some strategies of investors, such as “don't put all eggs in the same basket”, and so on. A fixed proportion investment strategy is constructed, which has high risk return and profitability.

Firstly, we set a trading threshold $\eta \geq 0$ to simulate the degree of investors' tolerance to the risk. Subsequently, all the assets are divided into a number of f parts, where we simply take $f = 10$ (some experts believe that). In fact, when evaluating and analyzing later, it can also show that the value here is appropriate. From September 11, 2016, on the first five days, we will not trade but only observe the change trend of various assets, and predict the market price on the sixth day according to the price on the first five days. Then will make decisions according to the rise and fall of the next day, and make continuous decisions by analogy.

Actually, a significant parameter in this investment strategy is the transaction threshold, which setting directly affects the transaction state, namely buying, selling and holding. If the threshold is too high, trading sensitivity is low, trading frequency is low, but this also means giving up a lot of arbitrage opportunities; on the contrary, if the threshold is set too low, the transaction sensitivity is high, and the transaction frequency is higher, it will bring a lot of uneconomical charges and unreasonable trading date. Therefore, setting a suitable trading threshold is very important to obtain a higher quantitative transaction yield. The above steps are explained in detail as follows:

Step 1. Set transaction threshold η

We can use historical price data to adjust the trading threshold in real time, so that the average interval days of transactions are in line with the actual investment in the trading market. The average trading interval days are 3 days, 5 to 7 days (one week) and 22 days (one month), respectively. Then according to the number of days, we calculate the threshold that historical trading days can trade according to the corresponding frequency. When all five years of trading are completed, we are surprised to figure out that the above threshold determination just makes trading days $1\sigma, 2\sigma, 3\sigma$ of the total number of days.

This also indirectly shows that the traditional quantitative investment idea using the mean-variance method for quantitative investment has certain value.

Table 1. Selection of threshold

PROPORTION OF TRADING DAYS	AVERAGE TRADING INTERVAL DAYS	G-THRESHOLD	B-THRESHOLD
0.6826	3	0.68%	2.87%
0.8571	7	1.14%	5.42%
0.9544	22	1.80%	9.60%

Step 2. Liquidation of assets

In the initial stage, our assets are \$ 1000. We put bitcoin and gold in dollars, so that we have a measure of value, and we can easily buy and sell our assets by a share, which is mean n dollars' worth of assets, both traded at most once a day, and simply consider n to be $1/10$ of our total assets.

Step 3. Trading according to the bullish or bearish prediction of bitcoin and gold.

The strategies: (we substitute F.P for the forecast price. B.F.P is the forecast price of Bitcoin, G.F.P is also the gold's.)

Note: The service charge for bitcoin is 2%, and the gold's is 1%.

A) Selling Bitcoin.

1. If the decline of B.F.P is less than the bitcoin threshold η_B , and there is enough bitcoin to sell, the number of bitcoins sold is determined according to the following:

$$\frac{r_i}{w_B} \times \frac{m}{RP_{t-1}} \tag{1}$$

And obtain $\frac{r_i}{w_B} \times m \times (1 - 2\%)$ dollars.

2. If the increase of G.F.P is more than twice the gold threshold η_G , then we will sell bitcoin as 1 and obtain the same dollars.

B) Purchasing Bitcoin

1. If the increase of B.F.P is more than the bitcoin threshold η_B , and there are enough dollars and do not sell bitcoins on the same day. The number of bitcoins bought is determined according to the following:

$$\frac{r_i}{w_B} \times \frac{m}{RP_{i-1} \times (1 + 2\%)} \tag{2}$$

and spend $\frac{r_i}{w_B} \times m$ dollars.

2. If the increase of B.F.P is more than twice the bitcoin threshold η_B , then we will buy bitcoin as 1 and spend the same dollar.

C) Selling Gold.

1. If the decline of G.F.P is less than the gold threshold η_G , and there are enough golds to sell, the number of golds sold is determined according to the following:

$$\frac{r_i}{w_G} \times \frac{m}{RP_{t-1}} \tag{3}$$

And obtain $\frac{r_i}{w_G} \times m \times (1 - 1\%)$ dollars.

2. If the increase of B.F.P is more than twice the bitcoin threshold η_B , then we will sell gold as 1 and obtain the same dollars.

D) Purchasing Gold

1. If the increase of G.F.P is more than the gold threshold η_G , and there are enough dollars and do not sell gold on the same day. The number of golds bought is determined according to the following:

$$\frac{r_i}{w_G} \times \frac{m}{RP_{i-1} \times (1 + 1\%)} \tag{4}$$

and spend $\frac{r_i}{w_G} \times m$ dollars.

2. If the increase of G.F.P is more than twice the gold threshold η_G , then we will buy gold as 1 and spend the same dollar.

The result is:

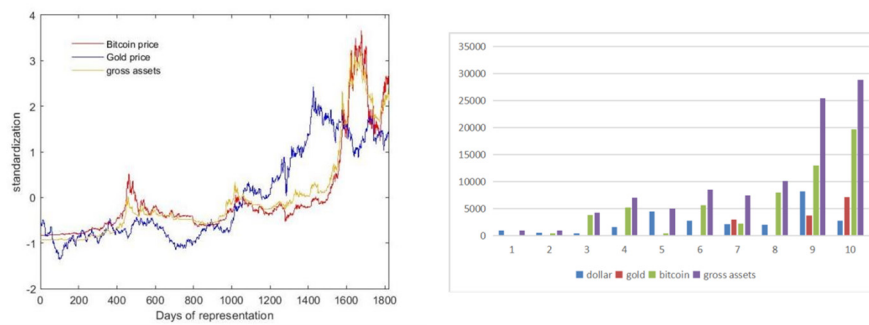


Figure 5. Relationship between total assets and financial commodity prices

In this strategy, the trend of total assets is similar to the bitcoin price, and the investment strategy is bitcoin-led. Among them, figure II represents the share of each asset, statistics per 200 days.

5. Conclusion

The construction of investment portfolio is a way to diversify investment. Investors allocate limited funds to different investment assets in a specific proportion on the basis of considering the amount of capital they have and the ability to bear risks, thus forming their investment portfolio, in order to avoid the loss caused by non-systematic risk while obtaining certain income. By considering the actual situation of financial market transactions and the limitations of classical models, this paper constructs double price prediction models based on ARIMA and LSTM. We have done a lot of comparative analysis and result analysis to consider how to maximize the value of the portfolio and come up with what we believe is the best daily investment strategy to provide viable solutions and recommendations for investors and investment institutions. After we have established those double price prediction models, with comparison and analysis, the prediction results of this are of high accuracy. At the same time, we propose a portfolio strategy: fixed proportion investment strategy, which makes different investment decisions by equalizing assets and setting reasonable thresholds on the basis of considering transaction cost constraints. The results show that our model has good robustness, and the investment based on our model and investment strategy will get considerable returns. In addition, our model can be widely applied to the investment of other investment products in the financial field, providing a reference for investors in risk management of investment products and making investment portfolio analysis decisions.

References

- [1] Cai Ziwei. Design of quantitative trading system based on Django framework [J]. Journal of Shanxi Datong University (Natural Science Edition), 2021,37(05):39-42.
- [2] Sun Liping. Cointegration and risk contagion of gold and bitcoin [D]. China University of Science and Technology, 2020.
- [3] Ganke. Analysis of the Influencing Factors of Bitcoin Price [D]. Central China Normal University, 2021.
- [4] Zhou Junmin, Zhang Bingyu. Prediction of my country's foreign direct investment trend based on seasonal product ARIMA model [J]. Journal of Hunan University of Finance and Economics, 2020, 36(06): 107-117.
- [5] Li Yafeng. LSTM Futures Investment Strategy Integrating Attention Mechanism [J]. Application of Computer System, 2021, 30(08): 22-30.