

A Study of BOS Investment Strategy Based on BP Neural Network Prediction

Jialei Liu[#], Shuhan Yang[#], Qianjie Wang[#]

Business College, Xi'an International Studies University, Xi'an, 710000, China

[#]These authors contributed equally.

Abstract: As a new type of trading method, quantitative trading has been developing rapidly in recent years in developed countries such as the United States. By building BP neural network models and BOS models, we analyze the rise and fall of gold and bitcoin prices, develop the best trading strategies, and guide investors in their investment decisions. The future trading day prices of gold and bitcoin were predicted using the BP neural network model; based on the known historical data, the distribution law of the corresponding rise and fall and the distribution law of the number of consecutive rises and falls were obtained using the Apriori algorithm, respectively, to develop the BOS (buy or sell) system. The resulting return curves were found to have an annualized return of roughly 5.43% for gold and 56.86% for bitcoin, with a relatively smooth fluctuation for gold and a relatively large fluctuation for bitcoin. Then, the impact of the change in transaction cost on the BOS model and total return is investigated by varying the transaction cost several times. The model is discussed in three aspects to provide the best strategy.

Keywords: BP neural network predicting model; The Apriori algorithm; BOS (buy-or- sell) system

1. Introduction

For purpose of maximizing total return, market traders tend to buy and sell volatile assets continually, often with a commission for each transaction. Here, this paper get two examples (charts are omitted). This involves quantitative trading. Quantitative trading refers to the use of advanced mathematical models instead of subjective human judgment, using computer technology to select a variety of "probable" events that can bring excess returns from a large amount of historical data in order to develop strategies, greatly reducing the impact of investor sentiment fluctuations to avoid making irrational decisions in the case of extreme market frenzy or pessimism. This paper need to build a mathematical model which uses only the past stream of daily prices to date to determine each day whether traders should buy, hold, or sell their assets. The main work of this paper is shown in Figure 1.

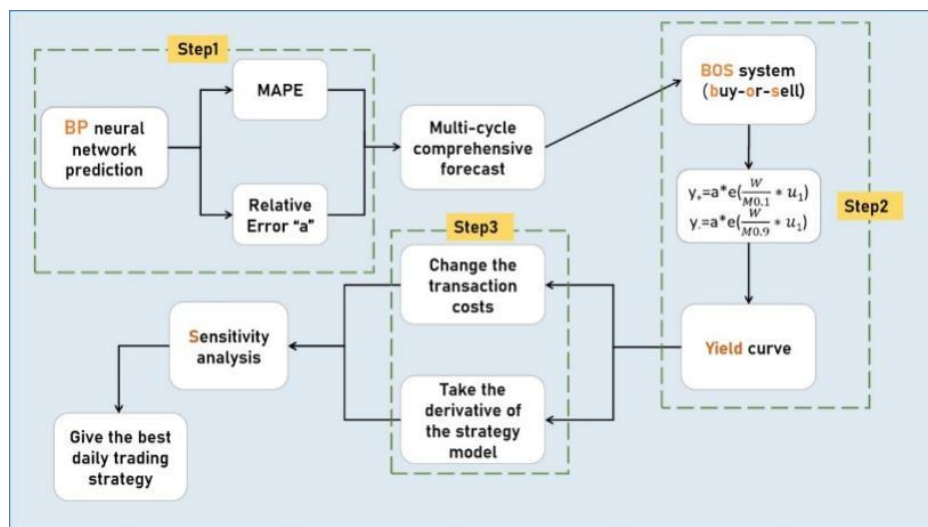


Figure 1: The main work of this paper

2. Assumptions and notations

2.1 Assumptions

- 1) Disregard the policy aspect.
- 2) Disregard the devaluation of financial products in the long term.
- 3) Assume that bitcoin and gold have the same investment risk and profit potential.
- 4) Investors must invest.

2.2 Notations

The primary notations used in this paper are listed as Table 1.

Table 1: Notations

Symbol	Description	Unit
CB CG	Initial investment amount of bitcoin Initial investment amount of gold	Dollar
$C_n C_i$	amount of the n^{th} buy in	\$
C_{t+1}	The predicted value of the next period	The predicted price of the working day
$C_{t+1} C_T$	The actual price of the working day	\$
$M_{0,x}$	Total initial investment amount of bitcoin and gold $0.x$'s percentile	\$
$n t_n T U_1$	The number of data groups Historical data with an interval of $(n-1)$ days	\$
$U_2 V_1 V_2$	period	\$
x	The number of times bitcoin price rose above 90% percentile	The number of
\hat{y}	times gold price rose above 90% percentile	The number of times bitcoin price
y_i	fell beyond 90% percentile	The number of times gold price fell beyond 90%
$y_+ y_-$	percentile The times of consecutive increases or decreases in the price	The
	predicted value	/
	The observed value	/
	Buy in strategy	/
	Sell out strategy	/
		/
		\$
		\$
		/
		/

3. Model construction and solving

3.1 The Establishment and Solution of BP Neutral Network

For the nonlinear problem, the traditional method is to linearize it as much as possible, and then analyze it with the theory and method of linear model. For many specific problems, this is an effective approach [1]. However, nonlinear systems have some special properties that are different from linear systems, such as chaos, bifurcation, and sensitive dependence on parameters and initial values, so linear methods will inevitably lead to the loss of these properties. In recent years, with the deepening of the people in the study of nonlinear problem, the neural network theory has been rapid development, since it has the self-organizing, self-learning and for input data and rules of robust fault-tolerant ability, has been increasingly applied in the stock market forecasts, economic forecasts, macroeconomic early warning research, population prediction field, etc.

Next, this paper formally introduce the neural network model. The Back-Propagation neural network composed of nonlinear change elements, is called B-P network for short. Take the two-layer network as an example, there are l inputs, m outputs, and two intermediate layers. Input node, intermediate node and output node are represented by subscripts i, j and h respectively [2]. The weight of h node input to the middle layer is represented by W_{ih} . The weights from middle layer node h to output layer node j are represented by W_{hj} [3]. 2-Layer Network Diagram is shown in Figure 2.

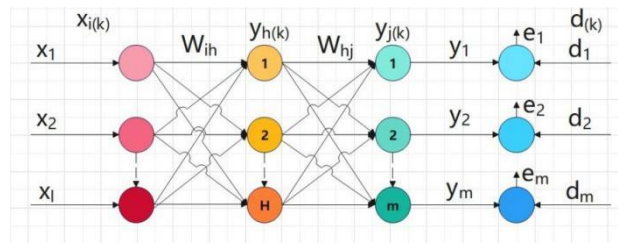


Figure 2: 2-Layer Network Diagram

Neural network has high requirements on data, and it should have several times of undetermined parameters to achieve high precision approximation. Here use the method of normalization and inverse normalization method.

In order to make neural network to get enough learning, for gold there will be 44 sets of data as a test set and 1,200 sets of data as a training set; for bitcoin there will be 30 sets of data as a test set and 1,767 sets of data as a training set.

Through carefully adjusting the number of hidden layer neurons, number of training and training error limit, obtained satisfactory results. The trained network was substituted into the calibration set to verify the value fluctuation from 9/11/2016 to 9/10/2021, as shown in Figure 4. Although there is a certain deviation between the sample value and the check value in the figure, the value fluctuation state is consistent every year [4]. Therefore, it can be considered that the neural network investment strategy model has acquired the dynamic behavior characteristics of value fluctuation state through the training of sample sets. In this way, Use it to predict the value fluctuation in the future. Comparison of Predicted and Actual Gold Price Fluctuations is shown in Figure 3.

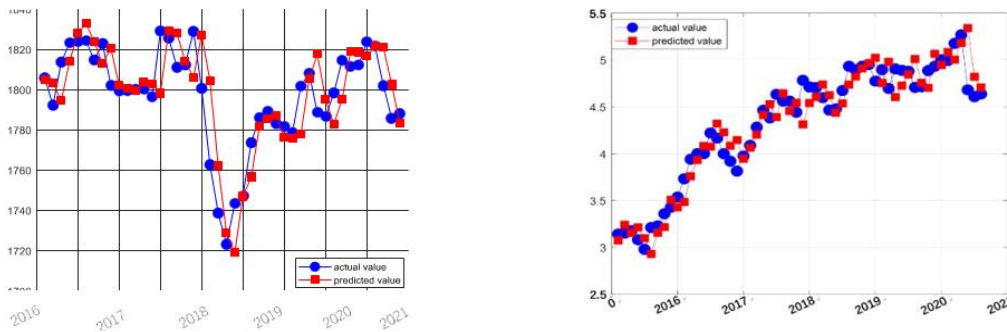


Figure 3: Comparison of Predicted and Actual Gold Price Fluctuations

To measure the prediction accuracy, this paper introduce the Mean Absolute Percentage Error, or MAPE. In the formula below, y_i represents the observed value and represents the predicted value. There are n groups of data in total. It is generally considered that when MAPE value is less than 10, the prediction accuracy is higher. Using the formula, this paper figured out the MAPE value, 0.035943, which means the relatively high accuracy.

$$MAPE = \frac{\sum_i^n \frac{|y_i - \hat{y}_i|}{y_i}}{n} \tag{1}$$

We also used another measure of prediction accuracy, which is the relative error rate a_t 's value. The price of the next working day is forecasted to be C_{t+1} , And the actual price of the next working day is C_{t+1} . The formula is presented as follows.

$$a_t = \frac{|\bar{C}_{t+1} - C_{t+1}|}{C_{t+1}} \times 100\% \tag{2}$$

We input the predicted value obtained by the BP neural network model into the formula with the actual value given by the question, and the relative error rate obtained is distributed around 0.03, which also means that the prediction result of the BP neural network model has a small error compared with the actual value[5]. Relative Error Rate Distribution is shown in Figure 4.

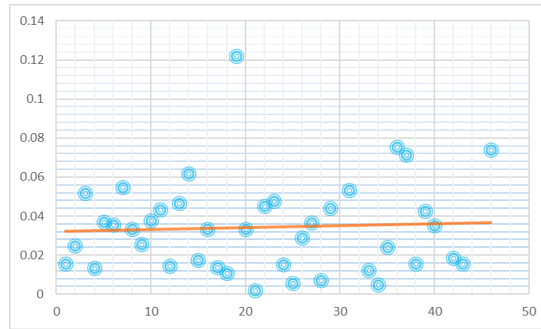


Figure 4: Relative Error Rate Distribution

The prediction of the future price of financial products is actually based on the historical price change rule, so the rule of price change has a great relationship with the way of data statistics. In order to make the prediction result more accurate, BP neural network is used to change the period to obtain the prediction result, and the specific operation is as follows.

Step 1, prediction of previous historical data selection. The corresponding data were selected as the original data set with the period T.

Step 2, the price of the next cycle is predicted through the prediction model established above, and the predicted value of the next period is obtained, which is $C_{t1}, C_{t2}, \dots, C_{ti}$

Step 3, average the predicted value for each period.

$$C_{T'} = \frac{\sum_{i=1}^n C_{ti}}{n} \tag{3}$$

According to the above model, the already known data of the problem are processed to verify the accuracy. First, price forecast after the 5th trading day are made. Because historical data is too small to be predictable, t_1, t_2, t_3 are corrected to 1, 2, 3, 4, 5, 6 respectively. For instance

$t_1 = 1$: to predict using everyday historical data as the original data set.

$t_2 = 2$: to predict using historical data with the interval of 1 day as the original data set. $t_n = n$: to predict using historical data with the interval of $(n-1)$ day data as the original data set.

Thus, this paper predict gold and currency's value respectively in advance to the next trading day price test using the historical prices. Take the price forecast on one of the days as an example. The Predicted Value from Period t_1 to t_6 is shown in Table 2.

Table 2: The Predicted Value from Period t_1 to t_6

Period	$t_1=1$	$t_2=2$	$t_3=3$	$t_4=4$	$t_5=5$	$t_6=6$	Average Value
Predicted Value(Gold)	1216.98	1212.43	1218.95	1215.16	1211.09	1218.76	1215.56
Predicted Value(Bitcoin)	682.46	680.09	677.31	676.47	687.04	686.7	681.68

This will be the model to predict the price on each trading day for the future. If the relative error between the predicted value and the actual value is small, the model established above is effective. Therefore, before making investment strategy, the price of the next trading day can be obtained through the above prediction model.

3.2 Justify Our Model Provides the Most Suitable Strategy

3.2.1 The Preparation of BOS Model

According to the above prediction model, Get the price of a particular trading day in the future, but the price can only determine the trading strategy of a particular trading day in the future, not a comprehensive trading strategy. Therefore, if want to develop a scientific and reasonable strategy, Need to use known historical data to derive the distribution of continuous price increases and decreases.

3.2.2 The Establishment and Solution of BOS Model

According to the investment rules of the financial market, when the value of financial products in the financial market falls, it needs to be appropriately bought, and when the value of financial products rises, it needs to be appropriately sold. In order to define the specific trading amount, it is necessary to study the investment rules of each financial product, so as to formulate corresponding strategies.

Used MATLAB software to process the data given by the question. Specific operations are as follows:

Excluding null values on non-trading days, for example, gold is only traded when market is open.

The actual increase or decrease range in the price of the financial product is calculated for each trading day. By processing data find that the maximum consecutive increase or decrease range of gold price does not exceed 4%.

Sort all increase ranges of gold and bitcoin respectively according to the size of the range, and get the median increase, M0.5, of gold and bitcoin.

Sort all decrease ranges of gold and bitcoin respectively according to the size of the range, and get the median decrease, M0.5, of gold and bitcoin. Figures are shown in the following Table 3.

Table 3: The Median Increase or Decrease Range of Gold and Bitcoin

M0.5	Gold	Bitcoin
Increase Range	0.0041	0.0158
Decrease Range	-0.0044	-0.0150

The association rule algorithm was introduced to calculate the number of subsets with two to eight consecutive upturns or downturns, as shown in the figure below. U1 represents the continuous increase of bitcoin and U2 represents the continuous increase of gold. V1 represents the sequential decline of bitcoin, V2 represents the sequential decline of gold.

Association rule algorithm is an important algorithm in data mining. Its core is a recursive algorithm based on the idea of two-stage frequent set. The association rules belong to single-dimensional, single-layer and Boolean association rules. The typical algorithm is Apriori algorithm. Apriori algorithm for mining association rules is a kind of frequent item-sets algorithm, its core idea is through the plot of the candidate set generated and closed down two stages for mining frequent item sets. The process of discovering association rules is divided into two steps by Apriori algorithm. The first step is to retrieve all frequent item sets in transaction database 1 through iteration, that is, item sets whose support is not lower than the threshold set by users. The second step uses frequent item sets to construct rules that meet the minimum trust of users. Among them, mining or identifying all frequent item sets is the core of the algorithm, accounting for the majority of the whole calculation. The algorithm has been widely used in business, network security and other fields.

According to the statistics of each u rise and fall in the historical data, Can obtain M0.9 of all rises and M0.1 of allfalls.

The investment model is constructed so that the number of consecutive rises or falls is X, and the maximum number of consecutive rises/falls of the product is U1, the maximum cumulative rise is M0.9, and the maximum decline is M0.1.

Add warehouse is the financial product prices when buying behavior, underweight is the financial product price rises to sell. In order to calculate the accuracy of the value, the commission cost is deducted in a lump sum after adding positions.

The amount of first addition is C1.

The amount of second addition is C2.

$$C_2 \times (1 - 1\%) \times M_{0.5} + C_1 \times (1 - 1\%) \times \left(\frac{C_{0.1}}{u_1}\right) \times M_{0.5} = 0 \tag{4}$$

The amount of second addition is C1.

$$C_3 \times (1 - 1\%) \times M_{0.5} + C_2 \times (1 - 1\%) \times \left(\frac{M_{0.1}}{u_1}\right) \times M_{0.5} + C_1 \times \left(\frac{M_{0.1}}{u_1}\right)^2 \times M_{0.5} = 0 \tag{5}$$

The amount of nth addition is C1.

$$(1 - 1\%) \times \left[C_n \times M_{0.5} + C_{n-1} \times \left(\frac{M_{0.1}}{u_1} \right) \times M_{0.5} + \dots + C_1 \times \left(\frac{M_{0.1}}{u_1} \right)^{n-1} \times M_{0.5} \right] = 0 \quad (6)$$

According to the above, the relationship between C1 and C2, C3, C4, C5 can be presented as follows.

The linear regression model was used to fit its parameters $y = a \times e^t$. An Exponential Function of The Amount of Gold and Bitcoin Being Added to and Removed from Positions is shown in Figure 5.

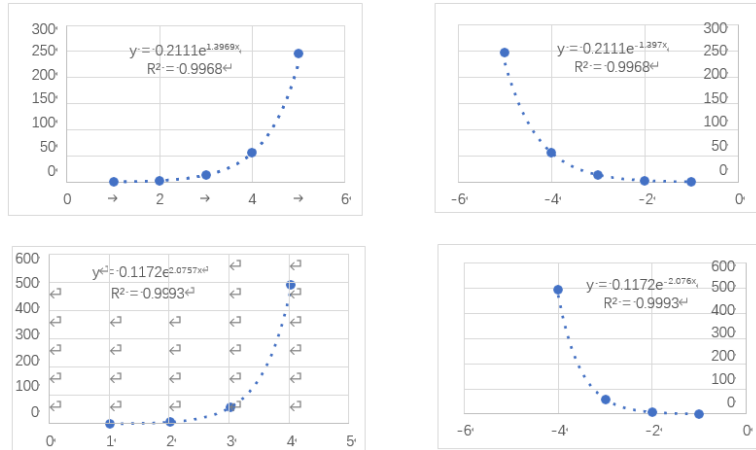


Figure 5: An Exponential Function of the Amount of Gold and Bitcoin Being Added to and Removed from Positions

On this basis to get under normal circumstances the positions of management model if the increase or decrease in x times in a row, total price for W, then.

$$y + a \times e^{\left(\frac{W}{M_{0.1}} \right)} \quad (7)$$

$$y = a \times eM_{0.9}^{\left(\frac{W}{x \times u} \right)} \quad (8)$$

In accumulating strategy y+, and underweight strategy y-. The combination of trend prediction mentioned in the first part makes the investment strategy more stable.

Here's the size of the initial investment in gold and bitcoin.

$$C_B = C_G = \frac{C_T}{2} \quad (9)$$

We split our initial \$1,000 investment 50/50 because we see gold and Bitcoin investing with roughly the same risk and return. Using historical data to simulate the above strategy, the return curve is shown in the Figure 6.

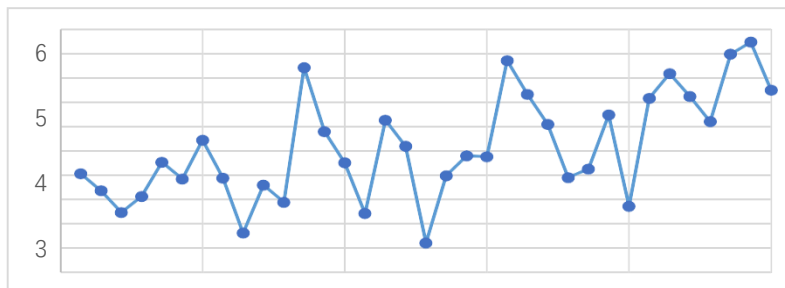


Figure 6: Daily yield curve from 2016 to 2021

3.3 Strategy's Sensitivity to Transaction Cost

After analysis, it can be found that BOS has strong stability and can be the optimal strategy that matches the short-term financial market on a large scale in the transaction strategy, thus avoiding certain

transaction risks. When the transaction cost is changed, the annual rate of return based on the exponential strategy expression is also different. The influence of transaction cost on gold annualized interest rate is shown in Table 4. The impact of transaction costs on bitcoin's annualized interest rate is shown in Table 5.

Table 4: The influence of transaction cost on gold annualized interest rate

Transaction cost(gold)	Exponential strategy	Annual yield
0.50%	$y=0.3094e^{0.9852x}$	5.63%
0.80%	$y=0.2149e^{1.3524x}$	5.76%
1.00%	$y=0.2111e^{1.3969x}$	5.43%
1.20%	$y=0.1882e^{1.5396x}$	5.98%
1.80%	$y=0.1697e^{1.6999x}$	5.35%

Table 5: The impact of transaction costs on bitcoin's annualized interest rate

Transaction cost(bitcoin)	Exponential strategy	Annual yield
1.50%	$y=0.1207e^{2.0508x}$	57.87%
1.80%	$y=0.1192e^{2.0583x}$	56.98%
2.00%	$y=0.1172e^{2.0757x}$	56.86%
2.20%	$y=0.1147e^{2.0972x}$	55.92%
2.50%	$y=0.1126e^{2.1178x}$	55.03%

4. Conclusion

Analyze and evaluate the results of the model to help improve the policy. The need and advantages of implementing this policy are explained in the context of the actual situation. The analysis of the points shows that the investment risk is mainly uncertain when it rises or falls. Therefore, a BOS investment strategy model is developed in this paper in order to gain profit in the investment market. First, the BP neural network is used to predict the daily rise and fall, and the BOS model is made by analyzing the rise and fall and continuous rise and fall to get the annual return and total return of gold and bitcoin financial products. Then the impact on the model is detected by changing the reaction cost and the results are analyzed to get the BOS model, which can avoid the risk to the maximum extent and obtain stable returns. Finally, the model has good stability.

References

- [1] Aldridge I. *High-Frequency Trading: A Practical Guide to Algorithmic Strategies and Trading Systems* [J]. April, 2013, 27(8):2267-2306.
- [2] Bao W, Yue J, Rao Y. *A deep learning framework for financial time series using stacked autoencoders and long-short term memory* [J]. Plos One, 2017, 12(7):1-24.
- [3] Baillie R T, Myers R J. *Bivariate garch estimation of the optimal commodity futures Hedge* [J]. Journal of Applied Econometrics, 2010, 6(2):109-124.
- [4] Chen D, Pan H, Dai Y, et al. *Investment strategy of colleges based on BP neural network and optimization program*[C]// International Conference on Advances in Mechanical Engineering & Industrial Informatics. 2016.
- [5] Lin Chunyan. *Research on price Prediction of Stock based on Elman Network* [J]. Computer applications, 2006, 26(2):477-484.