ISSN: 1813-4890

Analysis of Portfolio Trade based on Wavelet Neural Network and Dynamic Programming

Yifan Yao

School of Economics and Management, North China Electric Power University (Baoding), Baoding, 071066, China

2532628017@qq.com

Abstract

In market transactions, traders tend to frequently buy and sell asset based on market trends in order to ensure that they can achieve the highest returns after withdrawing the funds from the market. Through the trading period, commission is usually charged with a certain proportion for each transaction. In this problem, the trading portfolio is Gold and Bitcoin. Our goal is to create a daily trading strategy based on daily market fluctuations in order to maximize estimated profit with an initial asset of $1000 cash. We propose a hypothesis of rational traders who inevitably pursue the maximum return and the minimum commission in the process of trading. For the first part we build a forecasting model based on wavelet neural network. We test the data with White Noise Verification and verify that the daily price of Bitcoin and Gold is predicted and not random; then, due to Gold is only traded on open market day, we do data cleaning and multiple imputation on the data in order to do further prediction and analysis; finally, according to the processed data, we establish the forecasting model of daily price trends of Bitcoin and Gold based on wavelet neural network.

Keywords

Portfolio Selection; Multiple Imputation; Wavelet Neural Network; Bi-objective Dynamic Programming; Test the Sensitivity of the Model; Asset Management.

1. Introduction

In the market, if volatile assets are traded with good strategies, a good return probably can be achieved. These assets such as Bitcoins and Gold are traded frequently by traders who want to maximize their return. However, it is hard to rely on intuition to accurately predict the prices of them, and commission is deducted as part of the cost of each transaction as well. Thus, having an effective strategy is the key to successful investing.

2. Models

2.1. Data Processing

In the process of Data Processing, we carry out two steps: Data Integration and Multiple Imputation.

2.1.1. Data Integrating

In provided data, the price of Bitcoin is continuous from 9/11/2016 to 9/11/2021, while the price of gold is not, that of them are separated into two files, and the date data also has some errors in the representation of the year and month. So, we recreate a data column in a new file called data, and integrate the price data of Bitcoin and Gold from BCHAIN-MKPRU and LBMA-GOLD respectively into it according to the corresponding date.

2.1.2. Multiple Imputation

In the provided data, because there are 569 days without gold price data, and the data will be greatly affected by deleting the row of lacking gold price. Therefore, we use the Multiple Imputation (MI) to impute these missing values. Multiple Imputation (MI) is a method of dealing with missing values based on repeated simulations. When facing complex missing value problems, MI is the most common used method, which will generate a complete set of data from a data set containing missing values. In each simulated data set, the missing data will be filled in by the Monte Carlo method. And the process of Multiple Imputation is separated in two steps:

• Build a statistical model

In this project, we expect to use a linear model of gold price affected by time and price of Bitcoin to achieve multiple imputation. From the meaning of the problem, time is related to the price of Gold, so we need to validate the price of Bitcoin is related to that of Gold.

In the validation, First, we calculate the Pearson Correlation Coefficient between these price data to judge their linear correlation degree. In addition, we design a hypothesis testing to verify the correctness of the coefficient and the correlation. In the test, we suppose H_0: the overall degree of correlation between the daily price of Bitcoin and the that of gold is zero, and H_1: the overall degree of correlation between the daily price of Bitcoin and the that of gold is not zero. Then we work out that

$$P < 2.2e^{-16}$$

So, finally we can conclude that the result refuse H_0 and accept H_1, and the overall correlation between the daily price of Bitcoin and that of gold is not zero, they are correlative.

• Use R Programming Language for multiple Imputation [1] In the process of imputation, the function, "mice ()", is firstly used on the incomplete data and then returns an object that contains multiple complete data sets. So that each complete data set is imputed by the original incomplete data. At the same time, because the imputation is random, each complete data set is slightly different. Then, apply the validated model above to each complete data set in turn with the "with ()" function and obtain the analysis results. Finally, "pool ()" function is used to combine these individual analysis results into a set of pooled results. In the final pooled results, both P values (P_t and P_B) accurately reflects the uncertainty due to missing values and multiple imputations. As the result of the operation, both P values are less than 0.025. Therefore, we can conclude that the regression coefficients of the two are significant. So, the multiple imputation is available, and we get data2 through it.

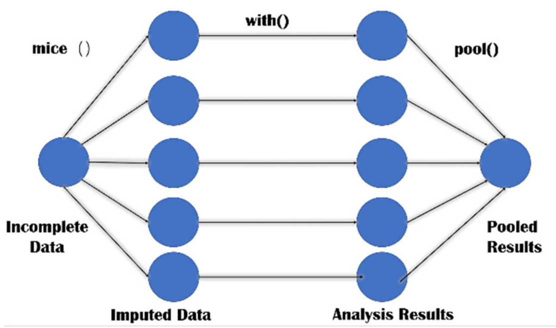

Figure 1. Multiple Imputation

2.2. White Noise Verification

Do Box-Lijung test on the daily price of Gold and Bitcoin, and the results we have come to are: $\chi G=1816.4$ and $\chi B=1253.9$, and both p values of them are less than 2.2×10^{-16}. As a result of that, the change of their prices with time is non-white noise and can be predicted by wavelet neural network.

2.3. Submodel 1: The Forecasting Model based on Wavelet Neural Network

In the project, we use Wavelet Neural Network (WNN) to predict the changes in the prices of Gold and Bitcoin. To predict the data, we first initialize the network parameters (Invisible Node: n = 9, Learning Probability: $Ir_1 = 0.01$, Iteration Number: num = 200). In addition, we train the network circularly, accumulate the error, and calculate the sum of error to adjust and update the weights [2]. Then, we use the trained network to make predictions, predicting the data for day 5 by that for the four days prior. And so on, we come to the final predicted data and draw the charts with it.

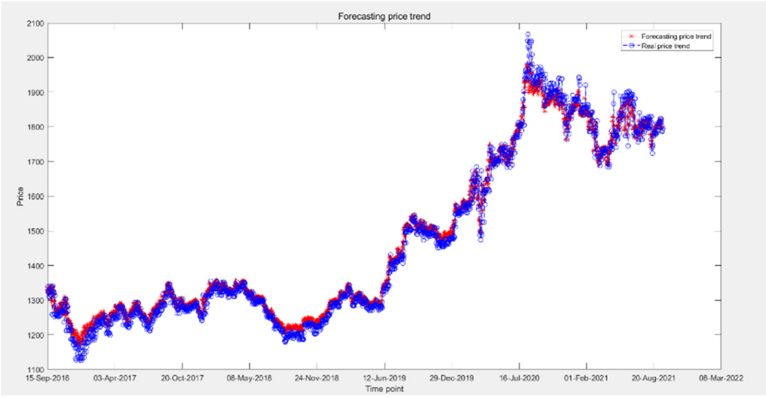

Figure 2. Essentially Coincident Forecasting- Price Trend and Real Price Trend of Gold

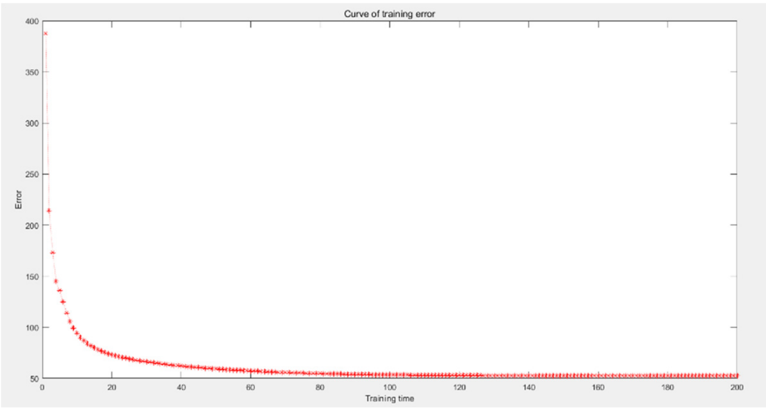

Figure 3. Training Error of Gold

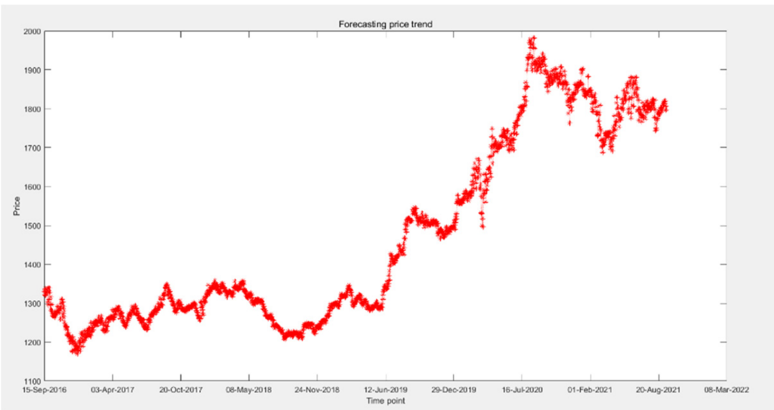

Figure 4. Forecasting Price Trend of Gold

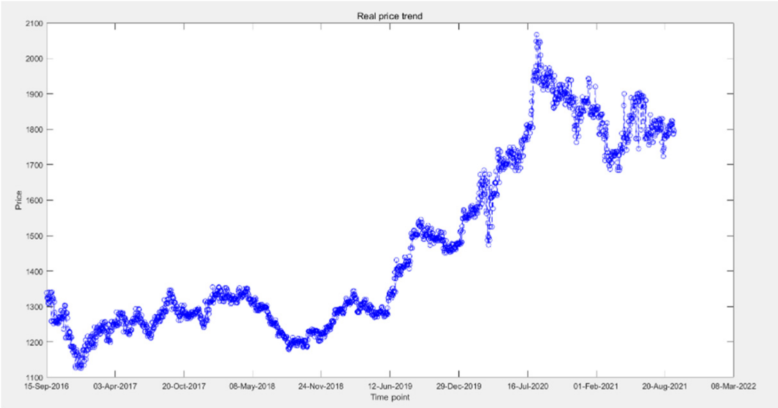

Figure 5. Real Price Trend of Gold

Figure 2, Figure 3, Figure 4, and Figure 5 are Essentially Coincident Forecasting Price Trend and Real Price Trend, Training Error, Forecasting Price Trend and Real Price Trend of Gold respectively. In Figure 2, it is clearly that the forecasting price trend and the real price trend basically coincide. At the same time, as can be seen from Figure 3, with the increase of the number of iterations, the error of the training network is getting smaller and smaller. So, the prediction of wavelet neural network meets expectation. And Figure 4 and Figure 5 show the prices trend of forecasting and real respectively.

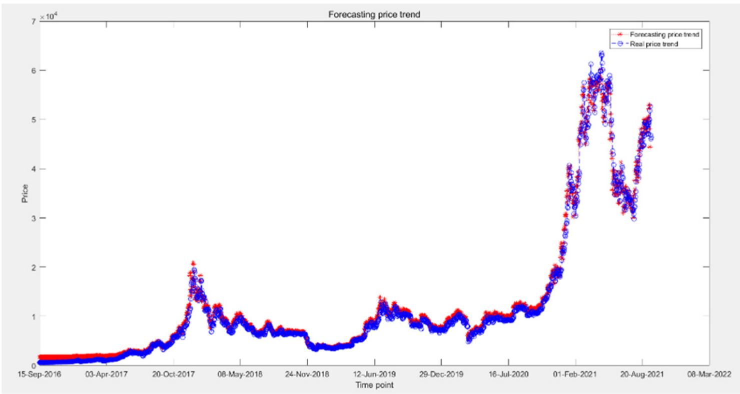

Figure 6. Essentially Coincident Fore- casting Price Trend and Real Price Trend of Bitcoin

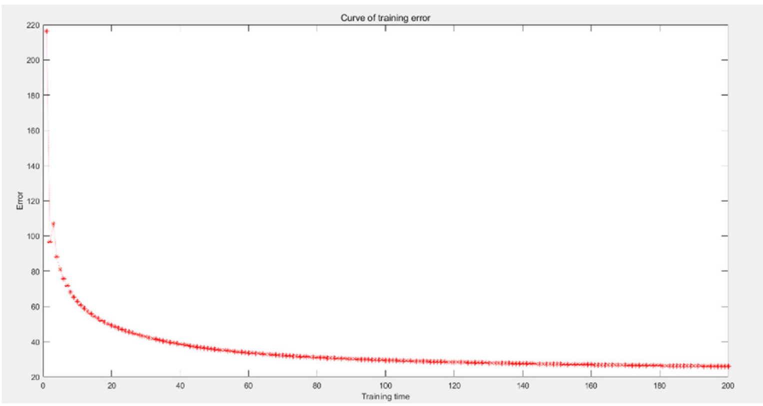

Figure 7. Training Error of Bitcoin

ISSN: 1813-4890

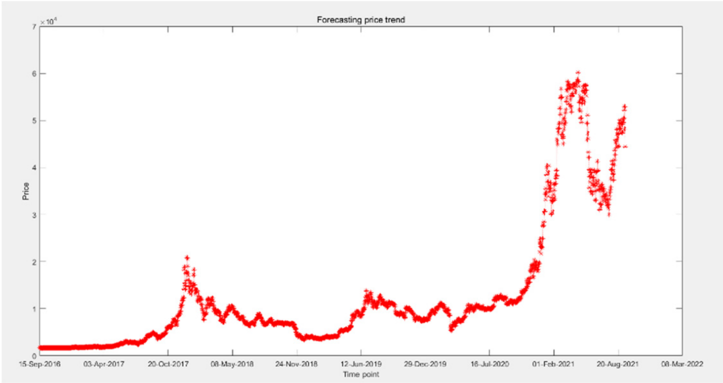

Figure 8. Forecasting Price Trend of Bitcoin

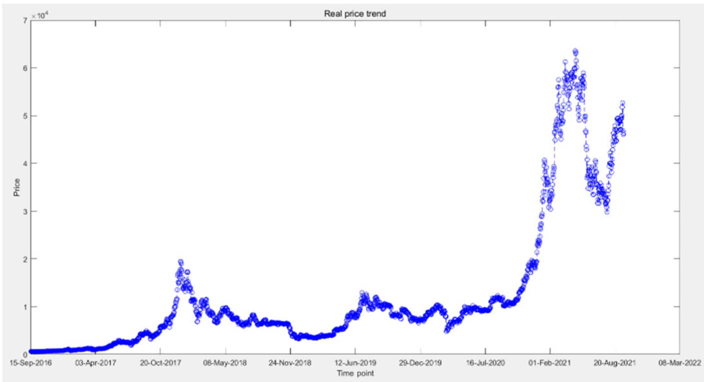

Figure 9. Real Price Trend of Bitcoin

Similarly, Figure 6, Figure 7, Figure 8, and Figure 9 are Essentially Coincident Forecasting Price Trend and Real Price Trend, Training Error, Forecasting Price Trend and Real Price Trend of Bitcoin respectively. In Figure 6, it is clearly that the forecasting price trend and the real price trend basically coincide. At the same time, as can be seen from Figure 7, with the increase of the number of iterations, the error of the training network is getting smaller and smaller. So, the prediction of wavelet neural network meets expectation. And Figure 8 and Figure 9 show the prices trend of forecasting and real respectively.

2.4. Submodel 2: The Optimal Trade Strategy Selection Model based on Bi-Objective Dynamic Optimization

2.4.1. Application of Submodel 2

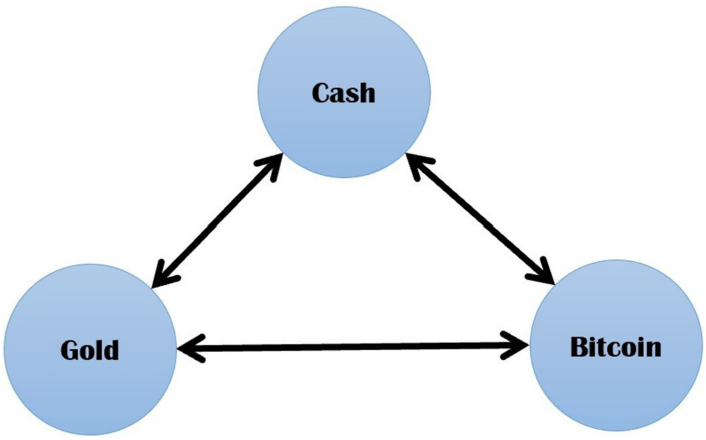

Figure 10. Transaction of Cash, Gold and Bitcoin

ISSN: 1813-4890

According to assumption 5 Investors choose the investment portfolio and investors in the market of all kinds of operations, and their purpose is to pursue in maximum returns when funds are withdrawn from the market.Therefore, in the case of commission in the trading process, investors definitely pursue in maximum returns and minimum commission through the trading process, and investors strictly follow the predicted results for the next step.

According to the analysis of the origins and whereabouts of Bitcoin and gold, there are six cases in total.

Table 1. Transactions of six cases

	Transaction Relation	Transaction Quantity	Commission incurred in the Transaction process
1	C → B	q1n	2%pbn q1n
2	B → C	q2n	2%pbn q2n
3	C → G	q3n	2%pgn q3n
4	G → C	q4n	1%pgn q4n
5	G → C → B	q5n	1%pgn q5n + 2%pbn q5n
6	B → C → G	q6n	1%pgn q6n + 2%pbn q6n

Investors' asset holding by day: Initial Holding: [c1, b1, g1] (from the problem:c1 = 1000, b1 = 0,g1 = 0)

...

The holding of the nth day: [cn, bn, gn]

The holding of the n + 1th day: [cn+1, bn+1, gn+1]

The relation between the holding of the nth day and the holding of the nth day:

bn+1 = bn + q1 + q5

Commissions from Transaction of the nth day:

2%Pbn [|q1 | + |q2 | + |q3 | + |q4 |] + 1%Pgn [|q3n | + |q4n | + |q5n | + |q6n |] Predicted commissions of the n + 1th day from the commissions of the nth day:

Cn+1 + pbn+1bn+1(1 − 1%) + pgn+1gn+1(1 − 2%)

Then, from the meaning of the problem:

$$q1n > 0 \cap q2n > 0 = \emptyset$$
$$q3n > 0 \cap q4n > 0 = \emptyset$$
$$q5n > 0 \cap q6n > 0 = \emptyset$$

Aim 1: Max Predicted − Value(n + 1) = Cn+1 + pbn+1bn+1(1 − 1%) + Ppgn+1gn+1(1 − 2%)

Aim2: MinCost(n) = $\sum_{i=1}^{n}$ 2%Pbn [|q1 |+|q2 |+|q3 |+|q4 |]+$\sum_{i=1}^{n}$ 1%Pgn [|q3n |+|q4n |+|q5n |+|q6n |]

S.T.:

Cn+1 = Cn + q2n + q4n

bn+1 = bn + q1n + q5n

gn+1 = gn + q3n + q6n

$q1n > 0 \cap q2n > 0 = \emptyset$
$\leftrightarrow q1nq2n = 0$
$q3n > 0 \cap q4n > 0 = \emptyset \leftrightarrow q3nq4n = 0$
$q5n > 0 \cap q6n > 0 = \emptyset \leftrightarrow q5nq6n = 0$
$qij \geq 0, \quad i = 1 \sim 6, \quad j = 1 \sim n, \quad i, j \in N+$
$C1 = 100$
$b1 = 0$
$g1 = 0$

When $\alpha g = 1\%$, $\alpha b = 2\%$ and the initial state is [1000,0,0], total asset changes as the trend of Figure 11 and the value of cash, Gold (transacted into US dollar) and Bitcoin (transacted into US dollar) change as Figure 12 after transacted by the the strategy provided from submodel 2:

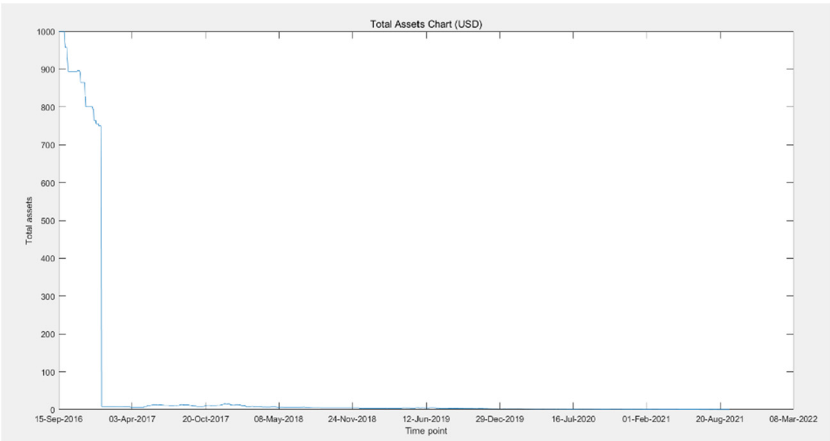

Figure 11. The trend of total asset

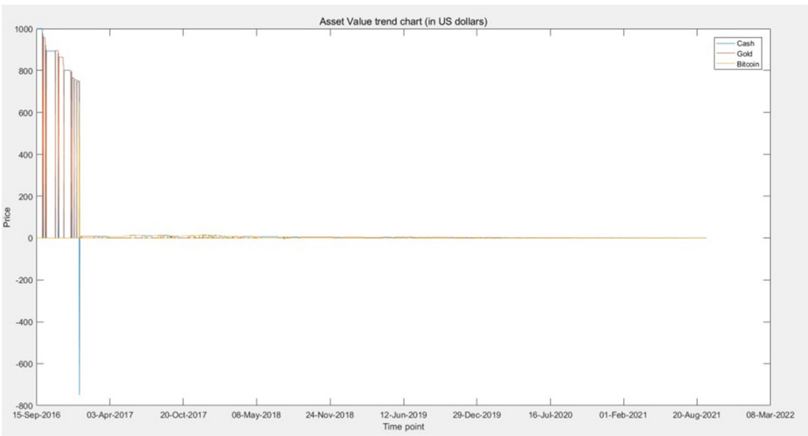

Figure 12. The trend of the value of cash, Gold (transacted into US dollar) and Bitcoin (transacted into US dollar

2.5. Analysis of the Result of Submodel 2

After substituting the imputed data in to submodel 2,the following result is obtained: a few months later, 1000 dollars are gone. We believe that this is due to the inconsistence between the prediction model and the actual situation (When the forecast result is that the next day will rise sharply and choose to buy all, but in fact the next day may fall, and it will lead to huge losses.). This operation is just like a gamble with excessive risk, so traders will not get a desired result. So, submodel 2 is unavailable and it should be improved.

2.6. Submodel 3: The Selecting Model of Optimal Investment Strategy with Asset Management

2.6.1. Introduction of Asset Management

In submodel 3, compared with submodel 2, we add the concept of asset management. In order to minimize the risk of choosing the wrong investment strategy due to the inconsistence between the prediction model and the actual situation, we decided to use the idea of building positions in stock and securities trading to avoid risks as much as possible.

In the process of opening a position, we first determine the asset we hold, and then sell/buy stocks and securities with the value of m when stocks and securities rise/fall. So that we can gradually reduce the cost when stocks and securities fall and make profits when they rise, which can help to reduce the risk.

After introducing the asset management, the submodel 2 is correspondingly improved to obtain the submodel 3 as follows:

2.6.2. Application of Submodel 3

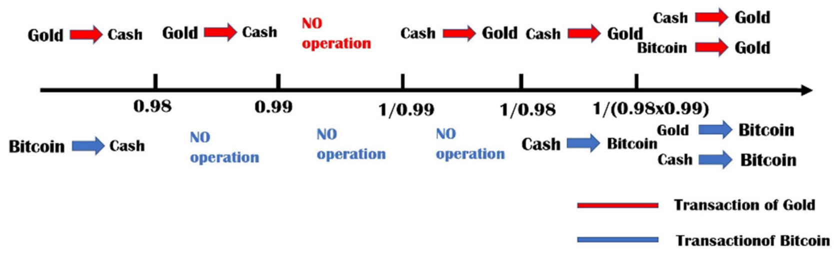

Figure 13. Transactions of Gold and Bitcoin

In the process of using Submodel 3 to invest, we divide the price prediction value of Gold and Bitcoin on the n + 1th day by the price prediction value of the nth day(pgn+1/pgn and pbn+1/pbn), and compare the calculation results with Figure to judge what operation the trader will perform on the nth day. Transaction of Gold

- pgn+1/pgn<0.98 : transact Gold to Cash
- 0.98<pgn+1/pgn<0.99 : transact Gold to Cash
- 0.99<pgn+1/pgn<1/0.99 : No operation
- 1/0.99<pgn+1/pgn<1/0.98 : transact Cash to Gold
- 1/0.98<pgn+1/pgn<1/(0.98×0.99) :
- 1/(0.98×0.99) <pgn+1/pgn :
  - if 1/(0.98×0.99) <pbn+1/pbn: calculate and compare the profits of transacting Cash and Bitcoin to Gold and transacting Cash and Gold to Bitcoin. Then, choose the transaction that makes more profit.
  - if pbn+1/pbn<1/(0.98×0.99): transacting Cash and Bitcoin to Gold

Transaction of Bitcoin

- pbn+1/pbn<0.98 : transact Bitcoin to Cash
- 0.98<pbn+1/pbn<0.99 : No Operation

- 0.99<pbn+1/pbn<1/0.99 : No operation
- 1/0.99<pbn+1/pbn<1/0.98 : No operation
- 1/0.98<pbn+1/pbn<1/(0.98×0.99) : transact Cash to Bitcoin
- 1/(0.98×0.99) <pbn+1/pbn :
  - if 1/(0.98×0.99) <pgn+1/pgn: calculate and compare the profits of transacting Cash and Gold to Bitcoin and transactting Cash and Bitcoin to Gold. Then, choose the transaction that makes more profit.
  - if pgn+1/pgn<1/(0.98×0.99): transacting Cash and Bitcoin to Bitcoin

When αg = 1%, αb = 2%,m=0.001 and the initial state is [1000,0,0], total asset changes as the trend of Figure 15 and the value of cash, Gold (transacted into US dollar) and Bitcoin (transacted into US dollar) change as Figure 16 after transacted by the the strategy provided from submodel 3:

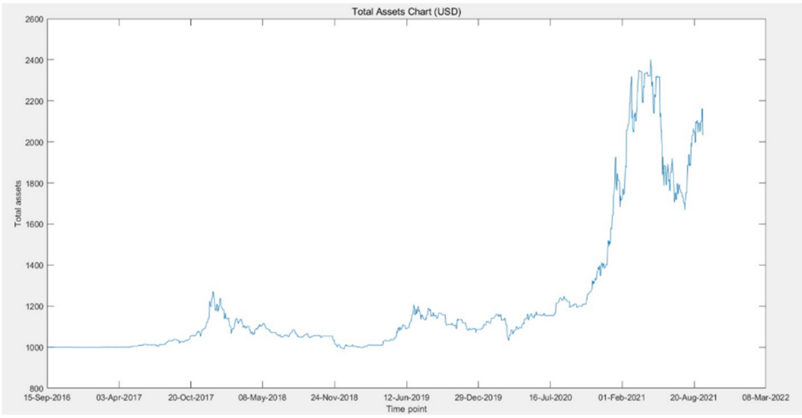

Figure 14. The trend of total asset

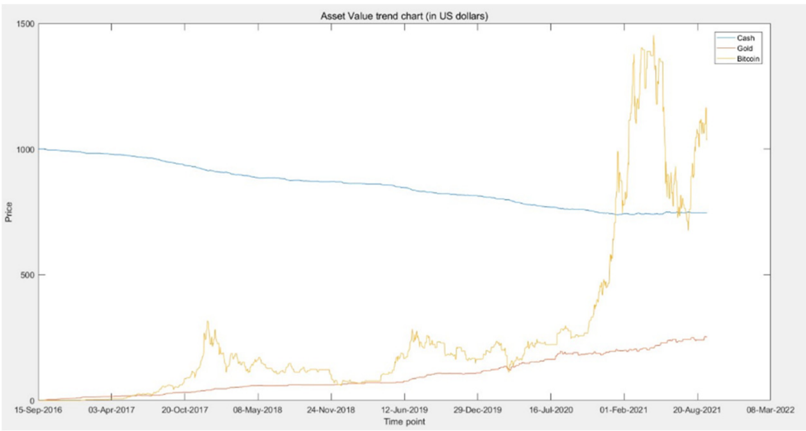

Figure 15. The trend of the value of cash, Gold (transacted into US dollar) and Bitcoin (transacted into US dollar)

When αg = 1%, αb = 2%,m=0.005 and the initial state is [1000,0,0], total asset changes as the trend of Figure 16 and the value of cash, Gold (transacted into US dollar) and Bitcoin (transacted into US dollar) change as Figure 17 after transacted by the the strategy provided from submodel 3:

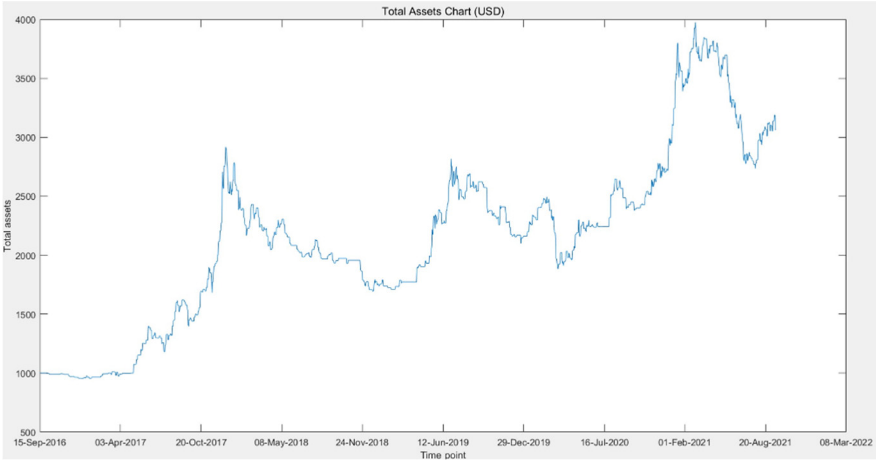

Figure 16. The trend of total asset

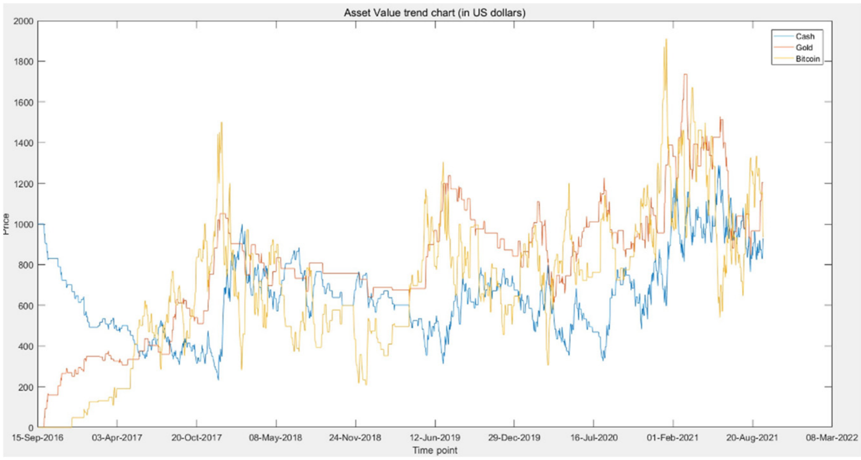

Figure 17. The trend of the value of cash, Gold (transacted into US dollar) and Bitcoin (transacted into US dollar)

When αg= 1%, αb= 2%, m=0.01 and the initial state is [1000,0,0], total asset changes as the trend of Figure 18 and the value of cash, Gold (transacted into US dollar) and Bitcoin (transacted into US dollar) change as Figure 19 after transacted by the the strategy provided from submodel 3:

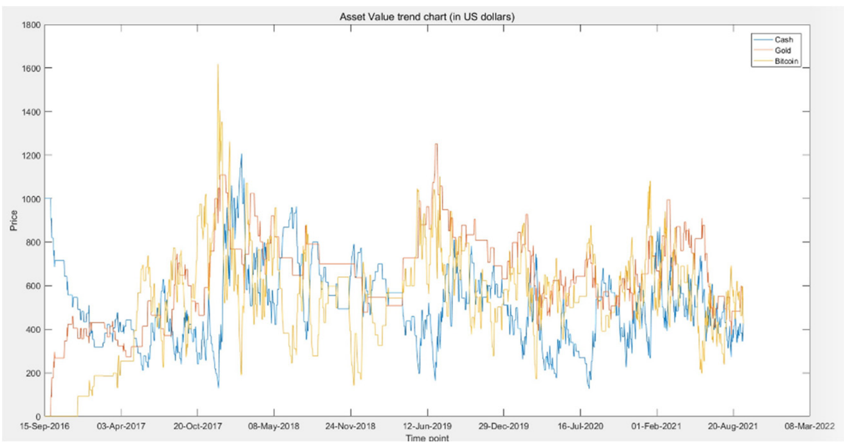

Figure 18. The trend of total asset

ISSN: 1813-4890

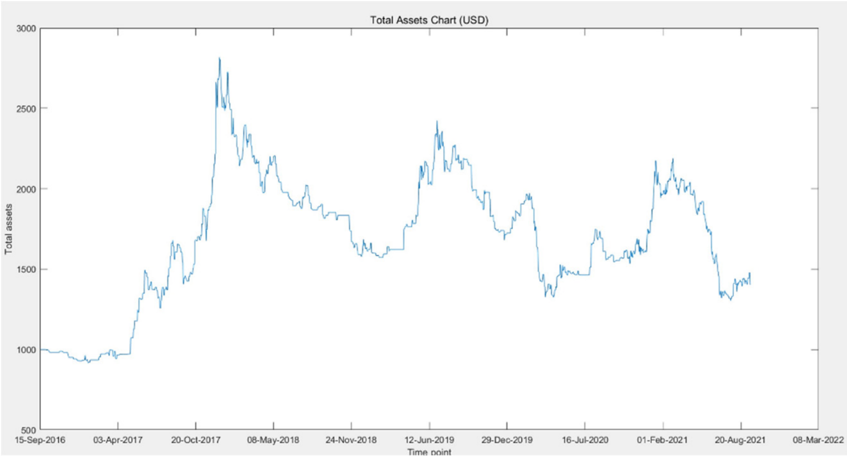

Figure 19. The trend of the value of cash, Gold (transacted into US dollar) and Bitcoin (transacted into US dollar)

When αg = 1%, αb = 2%,m=0.014 and the initial state is [1000,0,0], total asset changes as the trend of Figure 20 and the value of cash, Gold (transacted into US dollar) and Bitcoin (transacted into US dollar) change as Figure 21 after transacted by the the strategy provided from submodel 3:

Figure 20. The trend of total asset

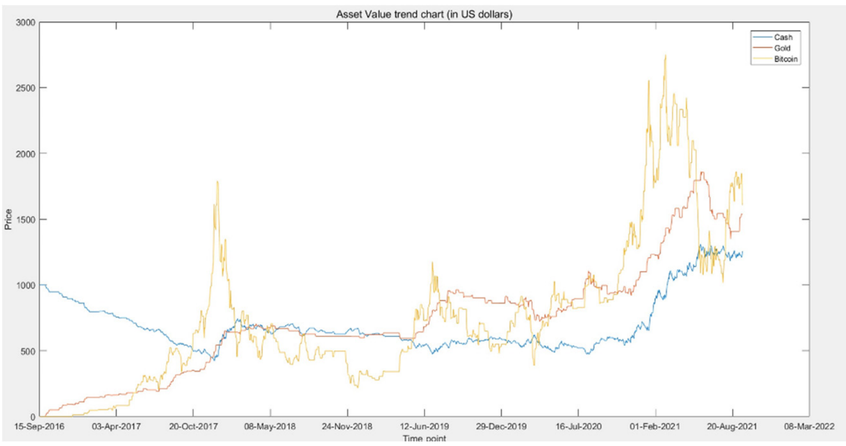

Figure 21. The trend of the value of cash, Gold (transacted into US dollar) and Bitcoin (transacted into US dollar

When αg = 1%, αb = 2%, and the initial state is [1000,0,0], as m increases, the value of the final asset increases first and then decreases, reaching a maximum at m = 0. 014.

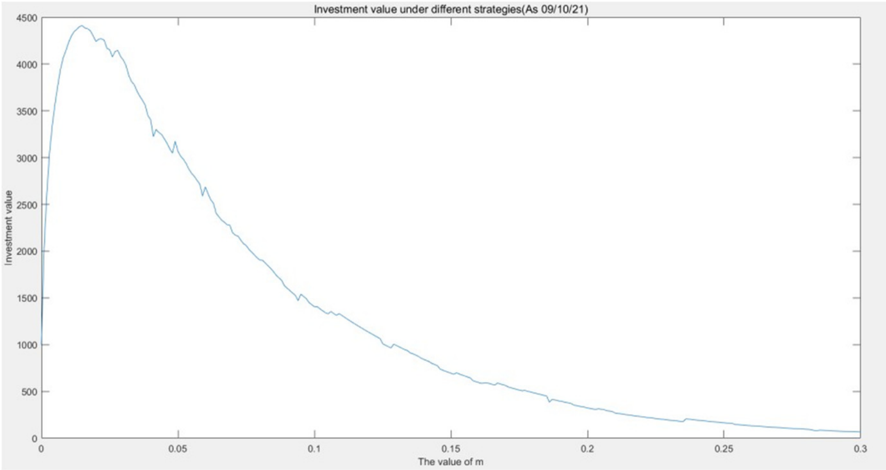

Figure 22. The trend of the value of final asset

2.6.3. Analysis of the Result of Submodel 3

• Figure 22 is a single peak curve, when m = 0.014, the final asset value reaches the maximum, and the data is the best investment choice. As m increases, the risk increases and the value of the asset begins to fall on the last day. So, when m=0.014, submodel 3 provides the best daily trading strategy, and the final asset worth 4403.92 dollars.

• Attention:However, if we choose to invest all the money in Bitcoin on September 11th, 2016, and liquidate all on September 10th , 21, the final capital will be 74607.71 dollars. We find that 4403.92 is much smaller than 74607.71, but we still think the model is correct for the following three reasons:

– It is not enough to determine the buying and selling strategy only from the index of daily price in the market. In reality, the determination of buying and selling strategy also needs to refer to other index.

– It is not feasible to predict the next 1000 days only by relying on a few data in the past few days, and the prediction accuracy is very low.

– Graphically, although buying Bitcoin can make money, it does not increase significantly between 2016 and 2020.

3. Conclusion

Submodel 3 is the most effective model among submodel 1, submodel 2 and submodel 3. It can acheive profitability. So we can proof that the model is available.

References

[1] Robert I. Kabacoff, R in Action.

[2] Jingyuan Wang,ăYang Zhang, Ke Tang, Alpha Stock: A Buying-Winners-and-Selling-Losers Investment Strategy using Interpretable Deep Reinforcement Attention Networks.

[3] Julius Adebayo, Justin Gilmer, Michael Muelly, Ian Goodfellow, Moritz Hardt, and Been Kim. 2018. Sanity checks for saliency maps. In NIPS'18. 9525–9536.

[4] Yingmei Chen, Zhongyu Wei, and Xuanjing Huang. 2018. Incorporating Corporation Rela-tionship via Graph Convolutional Neural Networks for Stock Price Prediction. In CIKM'18. ACM, 1655–1658.

[5] Xiao Ding, Yue Zhang, Ting Liu, and Junwen Duan. 2016. Knowledge-driven event embed-ding for stock prediction. In COLING'16. 2133–2142.

ISSN: 1813-4890

[6] Eugene F Fama and Kenneth R French. 1996. Multifactor explanations of asset pricing anomalies. J. Finance, Vol. 51, 1 (1996), 55–84.

[7] Hao Hu and Guo-Jun Qi. 2017. State-Frequency Memory Recurrent Neural Networks. In ICML' 17. 1568–1577.