

Sensibility and rationality in investment

Hongjian Fan*, Tianqi Wang

Southeast University Chengxian College, Nanjing, Jiangsu, 210088

*Corresponding author: 541237901@qq.com

Abstract. Feeling and impulse are the first level of thinking, and investing only by feeling and impulse is an unreliable investment behavior. Unfortunately, some people either take their own course by feeling, or following other people blindly. Such behavior is very risky. The purpose of investment itself is to grab the biz. In order to achieve the purpose of profit, we should keep a sensitive mind and formulate concrete scientific investment methods. In an effort to make investors come with joy and back with satisfaction, we start our modeling effort. Finding the result is also an important part of the paper. Based on two models of gray prediction and trading strategy, our team solved the profit amount set out on October 9,2021. The final solution result is \$242,584.835. What's more, we also give 5 reasons to show that our trading strategy is optimal. Among them, our team thinks that the most important thing is that our model can not only achieve profit purposes in the long term, but also better regulate the market with frequent short-term fluctuations. Sensitivity analysis is also an important part of this paper. Our team visualizes the obtained results and then analyzed from macro and micro perspectives, respectively, to speculate on the possible causes. Finally, we communicated our trading strategy to traders as a memo.

Key words: Gray Prediction; Data Mining; Trading Strategy; Analysis Sensitivity.

1. Introduction

“If we choose the suitable time to buy stock, even in the period of weak earnings and no matter how much the stock price toss and turn, the opportunity to make money is great.” said by an investor [1-3]. In recent years, the stock market scale in American has witnessed a tremendous growth, triggering more and more people abandon themselves to the stock market. Big as the gain is, risk always comes with it like shadow. As a result, investor needs a scientific strategy to deal with the ever-changing market. In a gesture to give investor a helping hand, our team will followed some methodologies to model these requirements [4-5]. Thanks to the accessories including “LBMA-GOLD.csv”, “BCHAIN-MKPRU.csv”, we can preliminary obtain the information and start our work[6].

In order to avoid complicated description, intuitively reflect our work process, the flow chart is shown in Figure 1.

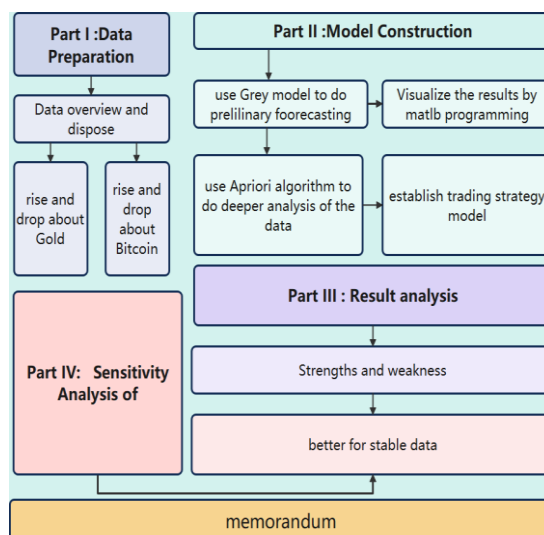


Figure1. The overall framework of our research

2. Assumptions and Justifications

To simplify the problem, we make the following basic assumptions, each of which is properly justified.

Assumption1: Suppose that all the data given by the topic are real.

Justification: If the data is not real, we have no practical significance to solve this problem, and we can we analyze this problem combined with life experience. The model can only be built for this problem without extensive applicability.

Assumption2: For subsequent grey prediction models, we assume negligible error after continuity of discrete variables.

Justification: The essence of the grey prediction is to serialize the discrete variables into a differential equation, where the predicted value is approximately represented by the solution of the differential equation, with the specific details discussed in the establishment of the subsequent model.

Assumption3: In a model of trading strategy, we assume that gold and bitcoin have the same gains and declines.

Justification: There are too many factors to consider when analyzing the short-term laws of gold and Bitcoin, and doing so can make our model more simplified.

3. Notations

The key mathematical notations used in this paper are listed in Table 1.

Table 1. Notations used in this paper

Symbol	Description	Unit
p_i	The rise or fall rate	×
J_i	The number of successive rises in p_i	×
j_k	The number of consecutive drops in p_i	×
$h_{0.1}$	The lower 10 quantiles in the sequence $p(i, u(type))$	\$
$h_{0.5}$	The median drop in the sequence p_i	\$
$H_{0.5}$	The median rise in the sequence p_i	\$
$H_{0.9}$	The higher 90 quantiles in the sequence $p(i, u(type))$	\$
$u(type)$	Maximum number of consecutive purchases or sales	×

4. Grey modeling

Then the bitcoin data is preprocessed, the format in the date column is reset, and then sorted from year to large, the date from 1 by day coding, a total of 1826 days. Similarly, gold data can be preprocessed, but unlike Bitcoin, gold data has several missing values. Due to the small number of missing values, our team chose to delete these dates with missing values and then coded by day for 1255 days[7].

Gray prediction is a prediction of a system containing both known information and uncertain information. The known information in this question is the date and corresponding value before a date to be predicted, and the unknown information is the value of the date to be forecast. And the dates to be predicted are also constantly changing, so this prediction process can be seen as dynamic.

The entire model establishment can be simplified to the following flow chart. As shown in Figure 2.

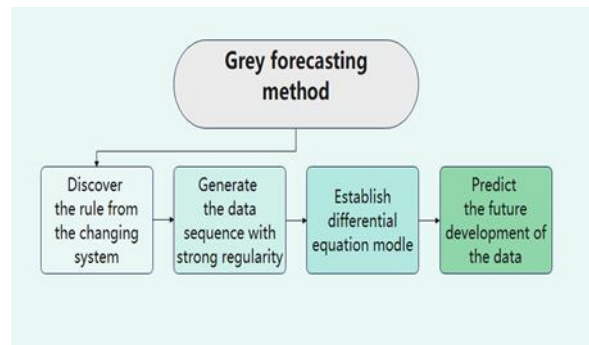


Figure 2. The process of gray prediction model

Gray prediction is using the original discrete non-negative data column, to generate a more regular new discrete data column weakening the random randomness, and then to model the differential equation to obtain the approximate estimate of the original data, so as to predict the subsequent development of the original data.

This process may be abstract, so our team took the bitcoin value of January and February 2001 as an example, visualize this process, took the date sorting and the number as the abscissate, and the value of Bitcoin as the Figure 3 (a) as the ordinate:

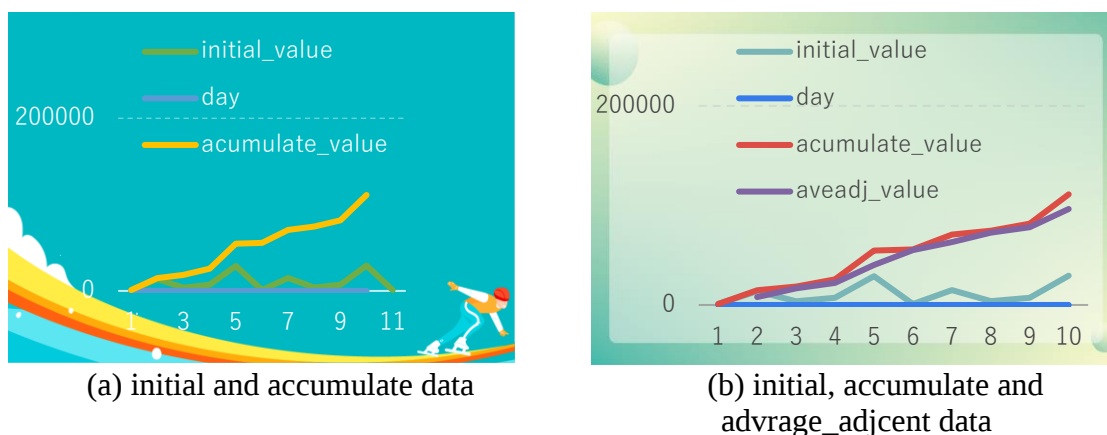


Figure 3. Visualization of gray prediction model data patterns

Next, we describe our process in mathematical language. Let $x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$ be the original non-negative data column, which we add once to get the new generated data column $x^{(1)}$ (1-AGO sequence of $x^{(0)}$):

$$x^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)) \quad (1)$$

$$x^{(1)}(m) = \sum_{i=1}^m x^{(0)}(i), m = 1, 2, \dots, n \quad (2)$$

$$z^{(1)}(m) = \delta x^{(1)}(m) + (1 - \delta)x^{(1)}(m - 1), m = 2, 3, \dots, n \text{ and } \delta = 0.5 \quad (3)$$

Next, the gray prediction model may be established. By consulting the relevant literature, the GM (1,1) model, the first "1" means that the differential equation is of first order, the second "1" means only one variable, and the GM (1,1) model satisfies the following gray differential equation:

$$x^{(0)}(k) + az^{(1)}(k) = b, k = 2, 3, \dots, n \quad (4)$$

Where, b indicates the gray action amount, and $-a$ indicates the development coefficient.

Then, we can define the matrix as follows:

$$u = (a, b)^T, Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}, B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix} \quad (5)$$

Moreover, we can represent the model in the matrix form as follows:

$$Y = Bu \quad (6)$$

It is quite clear that the area of the trapezoid is approximately equal to that of the curved edge trapezoid.

Combine the formulas (11) and (12) to obtain the following formula:

$$\int_{k-1}^k \frac{dx^{(1)}(t)}{dt} dt \approx -\hat{a} \int_{k-1}^k x^{(1)}(t) dt + \int_{k-1}^k \hat{b} dt = \int_{k-1}^k [-\hat{a} x^{(1)}(t) + \hat{b}] dt \quad (7)$$

After removing the integral number, the following differential equations are obtained:

$$\frac{dx^{(1)}(t)}{dt} = -\hat{a} x^{(1)}(t) + \hat{b} \quad (8)$$

$$\hat{x}^{(1)}(t) \Big|_{t=1} = x^{(0)}(1) \quad (9)$$

The result of solving the differential equation is:

$$\hat{x}^{(1)}(t) = \left[x^{(0)}(1) - \frac{\hat{b}}{\hat{a}} \right] e^{-\hat{a}(t-1)} + \frac{\hat{b}}{\hat{a}} \quad (10)$$

Order $t = m + 1$, the variable replacement after getting the following formula:

$$\hat{x}^{(1)}(m+1) = \left[x^{(0)}(1) - \frac{\hat{b}}{\hat{a}} \right] e^{-\hat{a}m} + \frac{\hat{b}}{\hat{a}}, \quad m = 1, 2, \dots, n-1 \quad (11)$$

By subtracting the accumulated data, the predicted value of the original data is obtained, with the formula as follows:

$$\hat{x}^{(0)}(m+1) = \hat{x}^{(1)}(m+1) - \hat{x}^{(1)}(m) = (1 - e^{-\hat{a}}) \left[x^{(0)}(1) - \frac{\hat{b}}{\hat{a}} \right] e^{-\hat{a}m}, \quad m = 1, 2, \dots, n-1 \quad (12)$$

Similarly, we can also get the predicted values of gold from 9 November 2016 to 2021 through this process. Visualization of our obtained results will give us a deeper understanding of the results, and the Figure 4 are as follows.

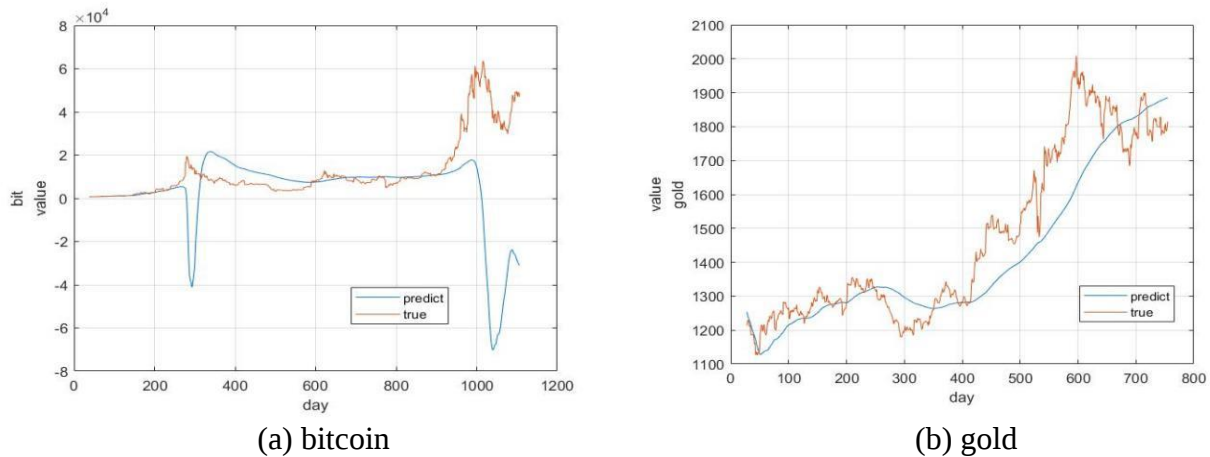


Figure 4. The original gray prediction model

Since the sample size of our original solution is increasing, as mentioned earlier, the data has a deadline. So, the status of the early data in the later predictions is gradually declining. Therefore, we can make some improvements to our model, so that the sample size remains constant, and we are able to remove the effects of the earlier data. Taking Bitcoin as an example, the mathematical formula for the original sequence improvement is as follows:

$$x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(36)) \quad (13)$$

It can be intuitively seen from the image, the prediction effect is still relatively good after some improvement of the model. The improved model not only predicts close values, but also predicts long-term trends in bitcoin and gold. More importantly, this model also has good predictions of short-term trends, indicating that the ability to adjust the market is better. After the model improvement, we think our prediction is feasible, so we can further advance the formulation of the trading strategy.

5. Data analysis and mining

Although we find out the rise and fall rate of gold and bitcoin, it is not enough to analyze our problem. We will make further mathematical statistics on these data. After consulting the relevant information, our team gave a feasible strategy. First, the rise and fall rate data is divided into rise and fall, respectively to find the median of these two parts. What's more, based on the association rule algorithm, the number of continuous rises and falls are counted. Then we find out the total number of times, and then calculate the cumulative contribution rate of over 90%, which is set as the maximum number. Finally, with the maximum number of times as the interval, the distribution law of rise and fall is counted again. Because the text description is not easy to understand, our team gave a flow chart of this mathematical statistics, and redefined the mathematical symbols to represent the process:

The results of this part lay the foundation for the establishment of the later trading strategy model, and the results are also more satisfactory, making the model more simplified.

6. Trading strategy model

Based on the above data processing and mining, we can build a trading strategy model. Initially we had \$1,000, which we want to reasonably allocate, and the distribution principles are as follows:

$$f_initial(bitcoin) = \frac{1}{2} * \frac{f_initial(all)}{u(bitcoin)} \quad (14)$$

$$f_initial(gold) = \frac{1}{2} * \frac{f_initial(all)}{u(gold)} \quad (15)$$

Of these cases are $f_initial(all) = 1,000$, $u(bitcoin) = 4$, and $u(gold) = 3$. This allocation principle is more conservative when the number of consecutive rises or falls is too large, avoiding the occurrence of risk to some extent.

Here's what we write about is the critical conditions for losses and earnings after consecutive gold or bitcoin $n(n \leq u(type))$ days. The cost of each transaction is also considered here, under the following conditions:

$$\begin{cases} [f_2 + f_1 \frac{h_{0.1}}{u(type)}] * h_{0.5} * (1 - \alpha) = 0 \\ [f_3 + f_2 \frac{h_{0.1}}{u(type)} + f_1 * (\frac{h_{0.1}}{u(type)})^2] * h_{0.5} * (1 - \alpha) = 0 \\ \dots \\ [f_n + f_{n-1} \frac{h_{0.1}}{u(type)} + f_{n-2} * (\frac{h_{0.1}}{u(type)})^2 + \dots + f_1 * (\frac{h_{0.1}}{u(type)})^{n-1}] * h_{0.5} * (1 - \alpha) = 0 \end{cases} \quad (16)$$

Let the fitting function be the following function:

$$y_i = a * b^{X_i} + u_i, X_i = \{1, 2, \dots, u(type)\} \quad (17)$$

The fitting, even with the minimum sum of square residues we obtain, is converted into the following mathematical formula:

$$\hat{a}, \hat{b} = \arg \min_{a, b} \hat{u}_i^2 = \arg \min_{a, b} \sum_{i=1}^n (y_i - a * b^{X_i})^2 \quad (18)$$

The result of accumulating the rise and fall rate of the interval $u(type)$ times is converted into the following mathematical formula:

$$w(X) = \sum_{k=1}^X p(i, u(type)) \quad (19)$$

With the previous assumption, if the number of consecutive rises or falls is X , we can get the buy and sell functions:

$$y_j = \hat{a} * \hat{b}^{\wedge} (\frac{w(X)}{h_{0.1}} * u(type)) \quad y_j = \hat{a} * \hat{b}^{\wedge} (\frac{w(X)}{H_{0.9}} * u(type)) \quad (20)$$

Where j means buying gold and bitcoin, and J means selling gold and bitcoin.

Based on the buy and sell function we seek, we can get the dollars we have daily after we buy and sell gold and bitcoin. Let the gain after day x be $own(x)$. Since $u(bitcoin) = 4$, $u(gold) = 3$, substituting this result into the model above, we can get that $own(0, bitcoin) = 125$, $own(0, bitcoin) = 166.67$.

7. The Result of Our Question

Since the question requires us to give the results of 9 October 2021, the data in the question were only given until 31 August 2021. We can therefore make predictions with the improved grey model. Since there were no data after 31 August 2021, we could only substitute the predicted results into the model for reprediction. After data analysis and mining results, we calculate into the trading strategy model. Our team got a final result of \$242,584.835, achieving a profit from \$1,000 on November 9, 2016 to \$242,584.835 on October 9, 2021 based on the trading strategy model and prediction model we established.

Through the above procedure, we can prove that our model is the best trading strategy, with the evidence as follows: We built a gray prediction model, predicted the future market trends. Therefore, we have a preliminary plan for the future trading strategies. Our model does not blindly pursue high returns, and also considers the risks that should be assumed. Our model fully excavates the intrinsic patterns of the data to maximize the value of it.

8. Sensitive Analysis

In order to test the stability of our model, the sensitivity analysis is inevitable. Order the transaction cost α change up and down for the sake of further analyze the model's sensibility.

The final results were obtained by Matlab programming, and to make the results more intuitive, we will do the visual processing.

Although, we change α , from a macros perspective, the figure almost coincides. We extracted the part of the figure for our further study, and analyzed it from a microscopic perspective. According to the information of the figure and the chart, $\hat{a}$ occur small variation, however, $\hat{b}$ doesn't change anymore. When the small changes of $\hat{a}$ occur, the figure only translating the figure up and down, without changing the overall trend of the figure, which coincides with the form of the fitting function. The figure reflects that changing a simply pans the figure up and down without changing the overall trend of the figure, which coincides with the form of the fitting function $y_i = a * b^{x_i} + u_i$.

But further observe the figure of bitcoin and gold, we still found some differences. The rising or down about bitcoin or even in table situation has very little effect on the sensitiveness change of α . We speculate that the reason may be the overall rise of bitcoin is stable. Moreover, the tip in the figure, in the other words, the invisible point is less, as a result, the adaptability to this model is strong.

As for gold, in the rising and falling situation, the sensitivity of α does not change much, but in the stable case, α has greater sensitivity. We think the possible reason is that the gold figure has more extremum, and these extreme value and can not derivation, gold figure stability is worse than bitcoin, so our model seems more suitable for buying and selling stability strategy. That means more suitable for buying and selling bitcoin strategy.

9. Model Evaluation

Strengths: We made further improvements on the original gray prediction model to make the model prediction effect better. The trading strategy model we built has better adaptability in both the long and short term.

Weaknesses: Our trading strategy model has too-idealized assumptions and therefore cannot be widely used in life. The gray prediction model we built generally uses about 10 data to predict. Obviously, it is not the most appropriate one in this question.

References

- [1] Computer and Digital Engineering, 2012 (1): 40-41,62.
- [2] HAO Ying, LIU Xing, LIN Chao-nan. An Empirical Research on the General Manager O 1. Conrad, Jennifer, and Gautam L. Cowle."Transaction strategy analysis."" Review of Financial Research, Volume 11, Issue 3, [Oxford University Press, Society for Financial Research], 1998, p. 489-519, [http: / / www.jstor.org/stable/2646008](http://www.jstor.org/stable/2646008).
- [3] Empirical study with investment decisions. Chinese Journal of Management Science, 2005, (5): 142-148 Physica, Heidelberg.
- [4] Verconfidence and Investment Decision for the Listed Companies[J].Chinese Journal of Management Science, 2005, (5): 142-148. Zhao Ying, Liu Xing, Lin Chaonan. Overconfidence about the general managers of listed companies.
- [5] Wang Zhixiang, Li Jiange, Chen Chupeng. Gray prediction of asphalt pavement use performance based on variable weight evaluation [J]. Journal of Chongqing Jiaotong University: Natural Science Edition, 2021 (5): 95-101.
- [6] Si Xicai, Chen Yukun, Li Zhigang. Study of the data association algorithms [J]. Journal of Harbin Engineering University, 2007 (7): 813-817.
- [7] Borgelt, C., & Kruse, R .(2002).Induction of association rules: Apriori implementation.In Compstat (pp.395-400).