A Comparative Study of the Functions of Gold and Bit coin in the Financial Industry

Hongyu Guo

School of finance, Xiamen University, Xiamen, China

ghy15847200190@163.com

Abstract. With the development of society, gold and "digital gold" Bit coin have been incorporated into the portfolio by investors. This paper will help traders predict the price of them, and plan the optimal portfolio strategy, so as to maximize the return and minimize the risk cost We have developed a model that uses only the past stream of daily prices to date to determine each day if the trader should buy, hold, or sell their assets in their portfolio. We will start with $1000 on 9/11/2016. We will use the five-year trading period, 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 initial state is [1000, 0, 0]. The commission for each transaction (purchase or sale) costs α% of the amount traded. Assume α gold = 1% and α bitcoin = 2%. There is no cost to hold an asset. Note that bitcoin can be traded every day, but gold is only traded on days the market is open, as reflected in the pricing data files LBMA-GOLD.csv and BCHAIN-MKPRU.csv. Our model accounts for this trading schedule. To develop our model, we may only use the data in the two spreadsheets: LBMA-GOLD.csv and BCHAIN-MKPRU.csv.

Keywords: Asset trading strategy, AHP, ARIMA, correlation analysis.

1. Introduction

Problem 1: We develop a model that gives the best daily trading strategy based only on price data up to that day. And our model will solve the problem of "How much is the initial $1000 investment worth on 9/10/2021 using our model and strategy?"
Problem 2: We will present evidence that our model provides the best strategy.
Problem 3: We determine how sensitive the strategy is to transaction costs. And our model will solve the problem of "How do transaction costs affect the strategy and results?"

2. Problem 1

2.1 The Models

2.1.1 Model 1: Bull and bear market judgment model

Select relevant indicators

The bull market and bear market analysis mainly by a certain period of derivatives price trend (Fluctuation range and BIAS).

here, By initially counting the daily gains in gold and Bit coin, It can be observed and found that, Gold's gains have changed slightly, Suitable for choosing longer-term estimates, Bit coin's gains have changed widely, For the applicable short-term estimation indicators, Gold we chose to calculate the estimate with a time span of a quarter, Bit coin we chose to use data with a time span of one month to calculate the specific performance as: first, For the price increase change factors, We calculated the daily gains for gold and Bit coin, respectively, And calculate the average of the first 90 / 30 days; second, For the consideration of the good and departure rate, We finally chose the gold 15th departure rate and the Bit coin 5-day departure rate, The mean of the first 90 / 30 days was also calculated, The corresponding index data were obtained separately, Normalization was also performed.

- Calculation formula 1: n-day BIAS = (current price-n daily average price) / n daily average price
- Calculation formula 2: Normalization = (Current Value-Min.) / (Maximum-Min.)

Establish a judgment model

Based on the two sets of data selected above, we use the judgment matrix in the hierarchical analysis method to calculate the weights, taking gold as an example:

Table 1. Example of Gold

	Gold rose on the first 90-day average	Gold 15-day separation rate in the first 90-day average
Gold rose on the first 90-day average	1	0.5
Gold 15-day BIAS in the first 90-day average	0.5	1

The corresponding weight can be obtained by using the normalization method, and the corresponding calculation model of gold cattle bear market index is established and drawn:(Gold bull bear market indicator =Gold rose by the first 90-day average of 0.666 +Gold 15-day separation rate in the first 90-day average×0.333):

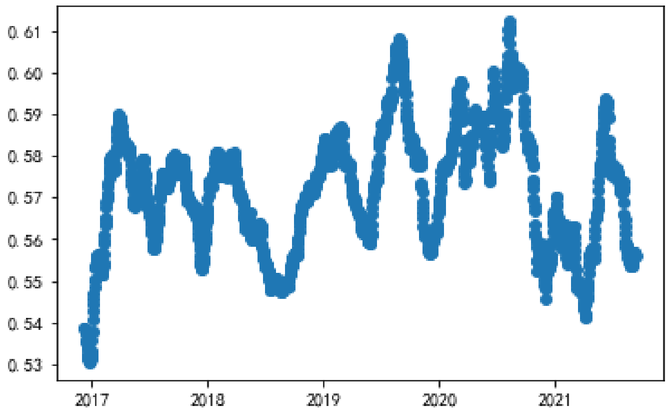


Fig. 1 Golden Bull and Bear Market Index

With the average value of 0.57 as the threshold, the gold bull market when the gold bull bear market index is greater than 0.57, when the gold bull bear market index is less than 0.57 is judged, and the distribution map is drawn:

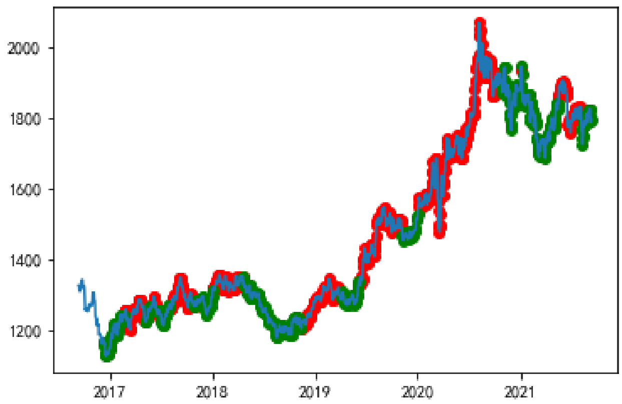
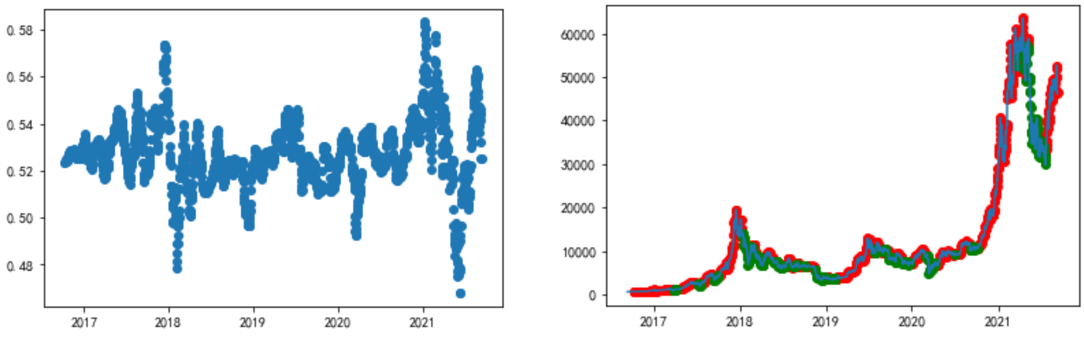


Fig. 2 Distribution of Gold Cattle Bear Market

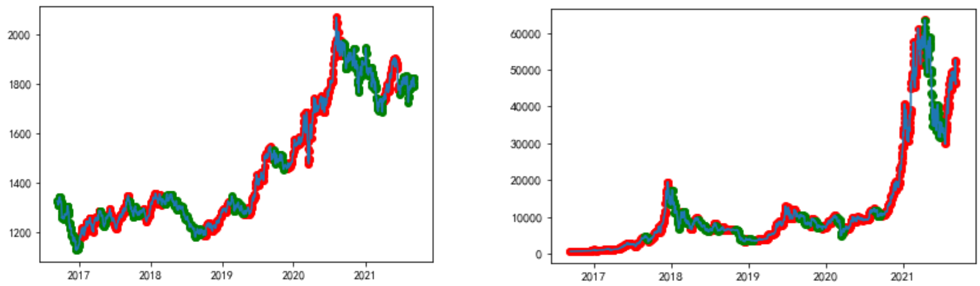
Similarly, the Bit coin bull and bear market index map and the bull and bear market distribution map can also be obtained:



(a) Image on the left (b) Image on the right

Fig. 3 Limitations and Correction

Due to the limited data, the gold / Bit coin bull bear market index in the quarter be- fore 2017 can not be effectively calculated and evaluated, so the voting method to bear judgment further correction, the initial index of all time to 0, if the current calculation for bull market, add 1,1 for bear market, the final result greater than 0 is bull market, less than 0 for bear market, and get the final gold Bit coin bull market distribution:



(a) Gold (b) Bitcoin

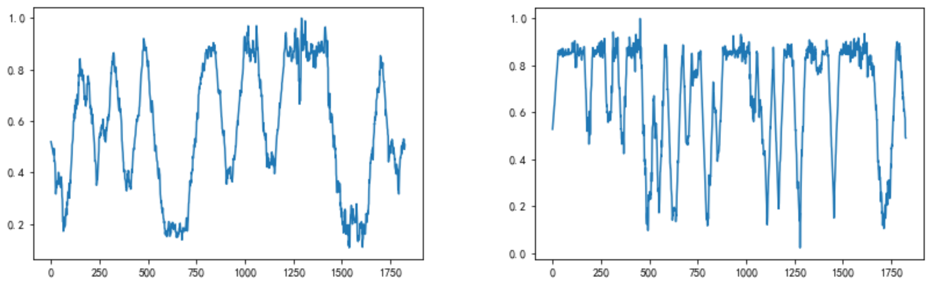
Fig. 4 Get The final Gold Bit Coin Bull Market Distribution

2.1.2 Model 2: Investment risk disclosure model

Based on the bull bear market analysis and existing data, we can simply measure bull bear market index and good separation rate to quantify gold and Bit coin in the current purchase risk, also take the hierarchical analysis of judgment matrix to calculate weight, calculate the bull bear market index of 0.666, BIAS of 0.333, so there are:

Gold investment risk index = gold bull bear market index of 0.666 + gold 15-day good departure rate of 0.333

Bit coin investment risk index = Bit coin bull bear market index 0.666 + Bit coin 5-day good departure rate 0.333



(a) Gold Investment Risk (b) Bit coin Investment Risk

Fig. 5 The calculated drawing yields the corresponding risk indicator map (the abscissa sorts the date label):

Through the comparison of the two figures, we can see that the risk of Bit coin is large, the risk index fluctuates frequently and violently, while the gold investment risk in a certain year is high, but the overall risk is low, especially in the mid-year of 2018 and the end of 2020 to early 2021, the low risk is suitable for increasing the proportion of asset investment.

2.1.3 Model 3: Time Series Prediction Model

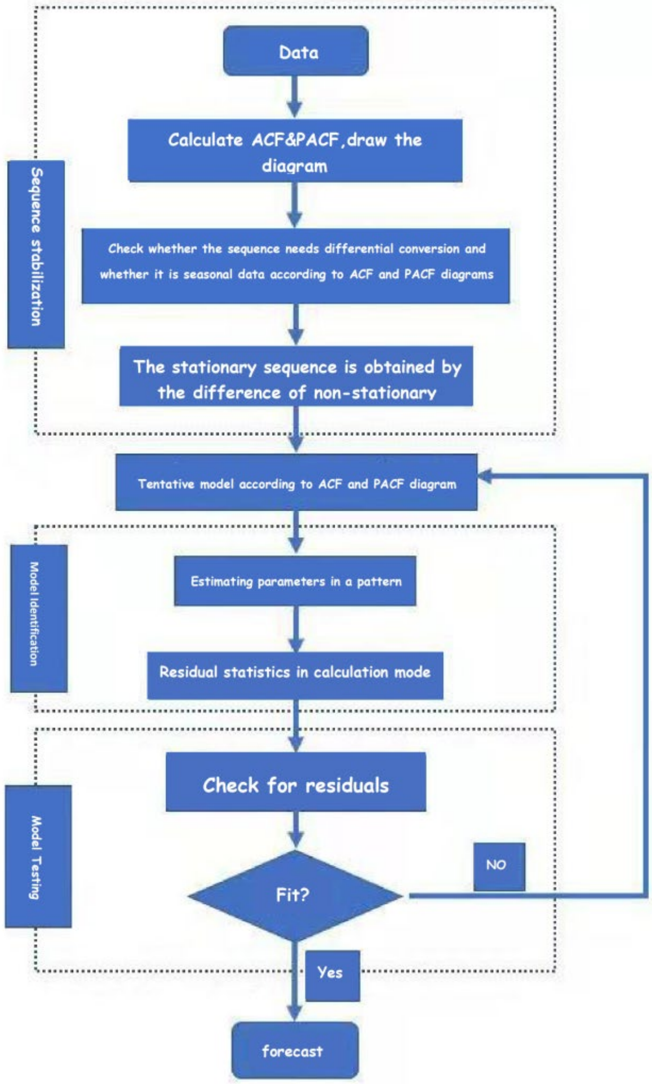


Fig. 6 Time Series Prediction Model

we used the ARIMA model to describe the relationship between the current value and the historical value, and predict ourselves with the historical time data of the variables itself, namely the auto-regressive model AR and the mobile average model MA model, and we obtained the auto-regressive mobile average model ARMA (p, q). The calculation formula is as follows:

$$y_t = \mu + \sum_{i=1}^{p} \gamma_i y_{t-1} + \varepsilon_t + \sum_{i=1}^{p} \theta_i \varepsilon_{t-i}$$

Get the trend of gold chart:

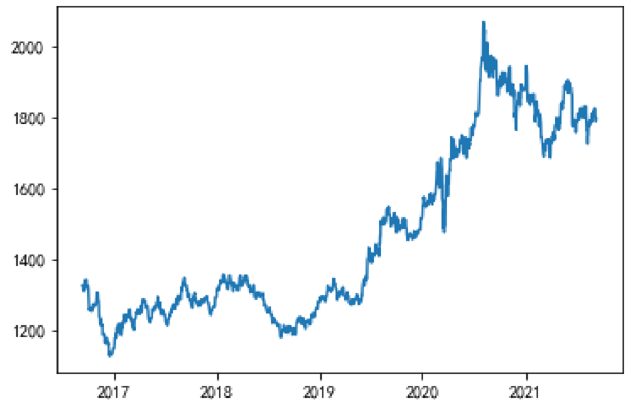


Fig. 7 Trend of gold

And gold difference comparison map, and from the comparison of several order difference, naked eye observation of data stability:

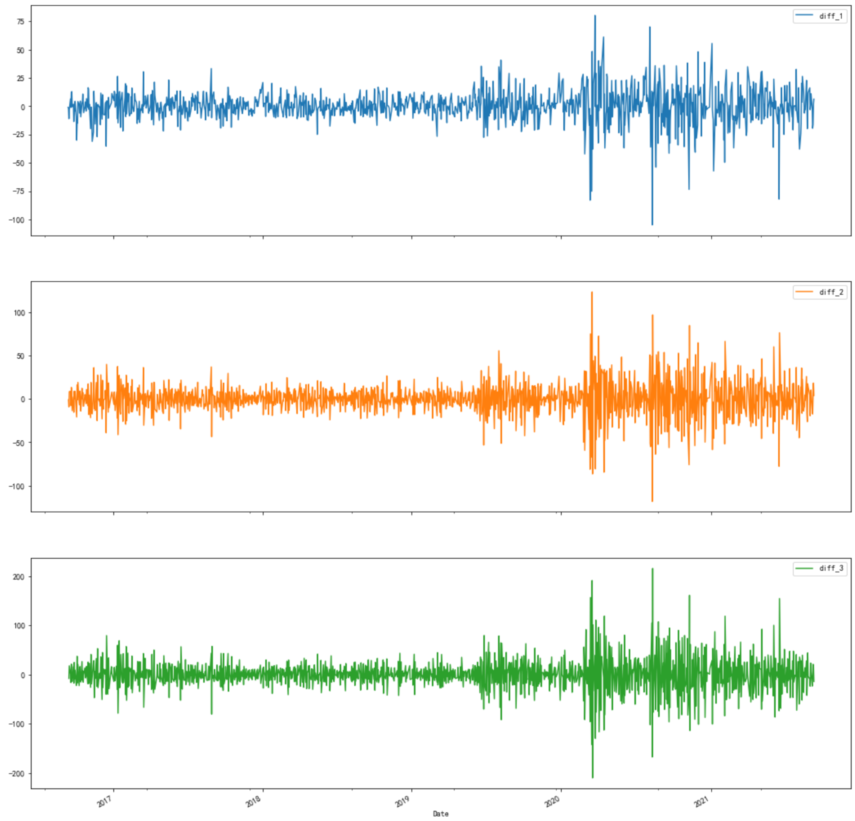


Fig. 8 Data Stability

We tested the data stationarity with the adfuller unit root.Output results:

The first is the result of the adt-test, simply the T values, representing the t-statistic.

The second is referred to as the p-value, representing the probability value corresponding to the t-statistic.

The third one represents the delay.

The fourth represents the number of tests. The fifth, viewed with the first, is the value of the critical ADF test at 99%, 95%, 90% confidence intervals.

• the comparison of 1%, 1%, 5% and 10% and ADF Test result, which is simultaneously less than 1%, 5% and 10% indicates that the hypothesis is rejected very well.

• p-value requires less than a given significant level and p-value less than 0.05, equal to 0 is the best.The null hypothesis of the ADF test is the existence of unit root, and as long as this statistic is a number at less than 1% level, the null hypothesis can be significantly rejected that the data is stable.Note that the ADF value is generally negative and positive, but it can only be considered as less than 1% and its significant rejection null hypothesis (the null hypothesis is not a stationary time series).

The following are the output results:

(-0.43415247222844744, 0.9042384812941653, 23, 1231,{'1%': -3.435673305025808, '5%': -2.863890744031555, '10%': -2.56802156936202}, 9957.827619599739)

(-8.158681091879238, 9.269711421536524e-13, 22, 1231,{'1%': -3.435673305025808, '5%': -2.863890744031555, '10%': -2.56802156936202}, 9948.53365765105)

(-12.876776886967274, 4.750971956332491e-24, 23, 1229,{'1%': -3.4356819860439574, '5%': -2.863894574114006, '10%': -2.568023609111511}, 9993.297129597462)

(-15.443541928994083,2.8443744636896075e-28,23,1228,{'1%':-3.4356863371792095,'5%':-2.8638964938393667,'10%':-2.568024631481501}, 10131.546643110885)

It can be observed that the P-value of the original data is bigger than 0.05, so the stability requirement is not met. The P-value of the first order difference is smaller than 0.05, and the T-value is less than 1%, 5%, and 10% can significantly reject the null hypothesis, indicating that the data is stationary.The first order differential data has been stable without no need to continue the second-order difference. Gold first order difference chart:

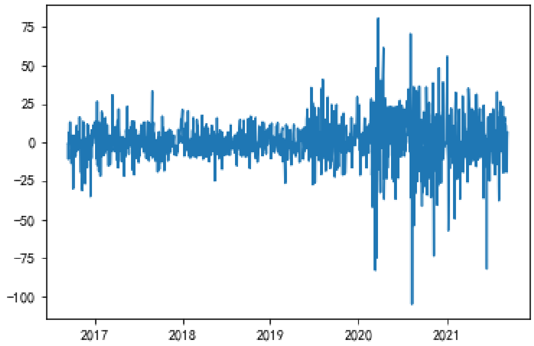


Fig. 9 Gold first order difference chart

The second term of the acorr_ljungbox output is the p-statistic based on the chi-square distribution, and the P-values of the first order difference are all small, so the data reject the null hypothesis that the data is not purely random and can do time series analysis. Then choose p and q, which share two methods: the first is the observation method, in which the partial auto-correlation function PACF is used, which describes the linear cor- relation between the time series observations expected in the past given the intermediate observation value. As well as the auto-correlation function ACF, it describes the linear cor- relation between the time-series observations and their past observations.The calculation formula is as follows:

$$ACF(k) = \rho_k = \frac{Cov(y_t, y_{t-k})}{Var(y_t)}$$

The PACF plots and ACF plots were obtained:

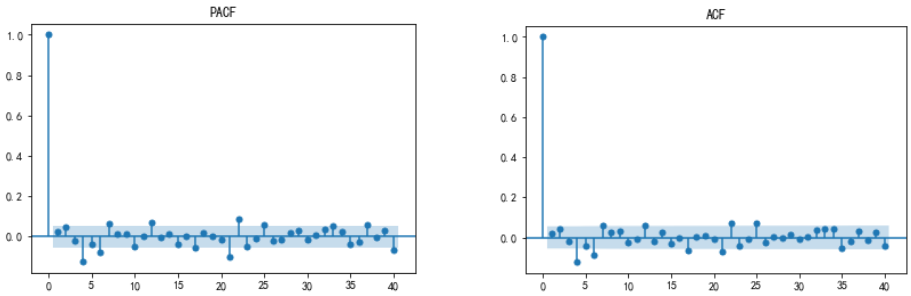


Fig. 10 PACF plots and ACF plots

The second is the exhaustive method, max_ar and max_ma are the maximum desirable parameters. When the value is large, so 5 and the result is (p, q) = (0,0). Next is the residual test: If the residuals are a white noise sequence, it shows that the useful information in the time series has been extracted, and the rest is all random perturbations, and cannot be predicted and used.In the qq diagram: if it is a normal distribution, it is a straight line, namely the red line.The results roughly correspond to the white noise.A QQ plot of the gold ar-ma residuals is plotted here.

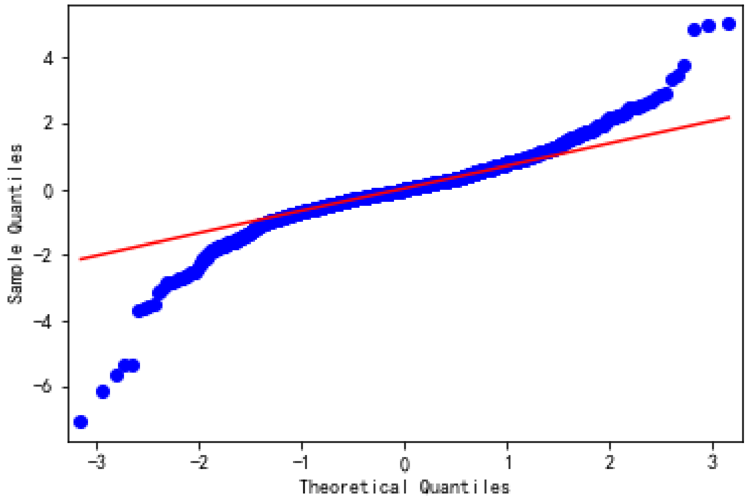


Fig. 11 Residuals

Durbin-Watson test was performed to test statistics:

$$d = \frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^2}$$

The resulting statistic, d-value is 2.01. Then start to predict the predict (start time, termination time), the predicted result is also the first order difference, that is, the gold rise, to get the comparison map of the gold increase (the actual and forecast comparison).

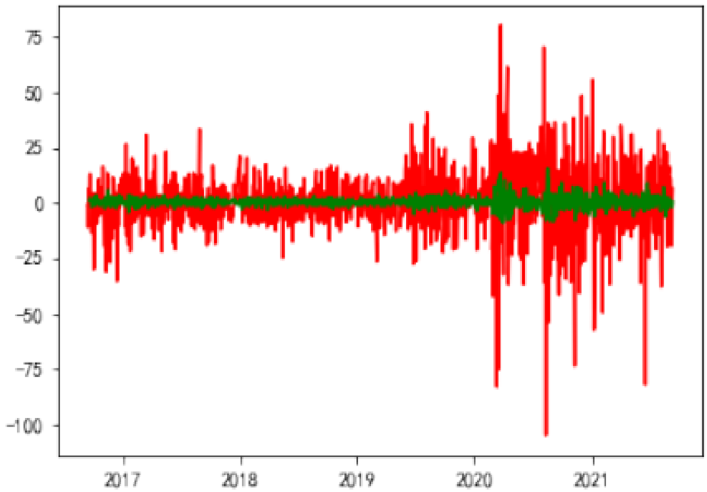
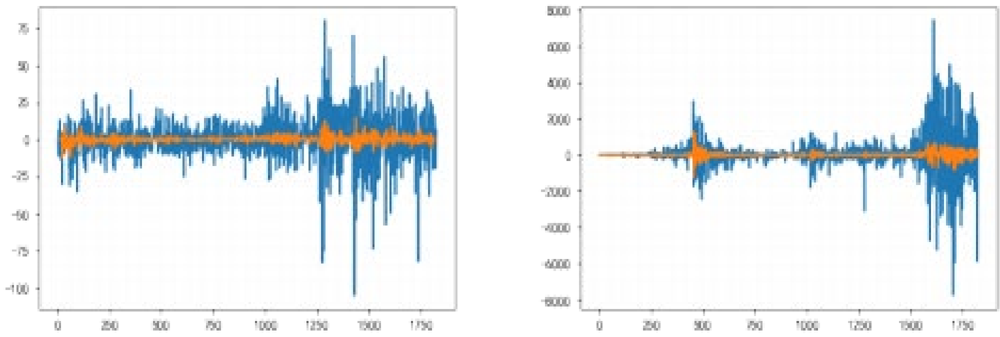


Fig. 12 Comparison map of the gold increase

While the big rally was not predicted, the overall peak was largely predicted, as is Bit coin.Next is expected to rise, get gold (a) and Bit coin (b) is expected to rise:

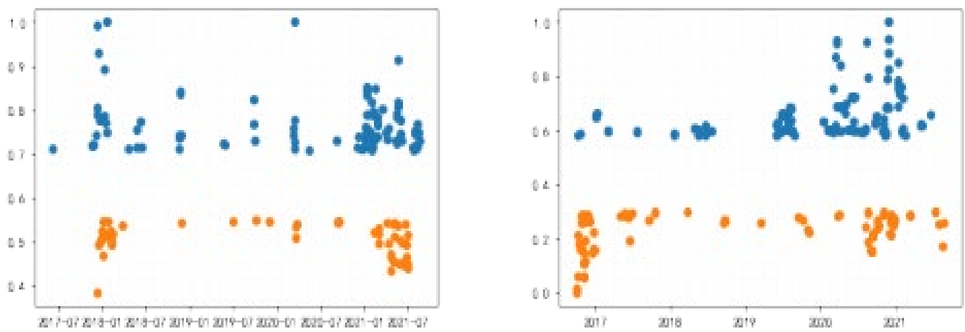


(a) Gold (b) Bit coin

Fig. 13 Overall peak forecast

Calculate the buy score formula: buy score = increase*10 + bull-bear market index * 5-residual + 1 / investment risk index——

Then compare the rationality of the score: through the comparison of gold buying score chart and gold trend chart, you can see that you buy at trough and sell at peak. Bit coin is the same, so the comparison score is reasonable.We then normalized the buy score, and we set a threshold to determine whether to buy or sell (blue needs to buy, orange needs to sell), and compared the Bit coin (a) and gold (b) score:



(a) Bit coin (b) Gold

Fig. 14 Comparing Bitcoin (a) and Gold (b) scores

2.2 Solution

First question trading model core:

Gold score more than 0.58 to buy, sell when less than 0.3, Bit coin score more than 0.71 to buy, sell when less than 0.56, that is, gold buy standard 0.58, sell standard 0.3.

Trading Rules:

• To judge whether it is a golden trading day, consider gold, not not gold

• When Bit coin and gold can be bought simultaneously, if the gold buy score is- 0.58> (Bit coin buy score-0.71) X2 (gold score is lower than Bit coin, more room for rise)

• Purchase Amount = Current Cash Line X Purchase Score X (1-Fee) / Current price

• Sales Line = Holding Share * (1-score + sales standard) We conducted a simulated transaction, and finally got the total asset map, and found that the final results were around 25w, ranging from around 1000 to 25w.

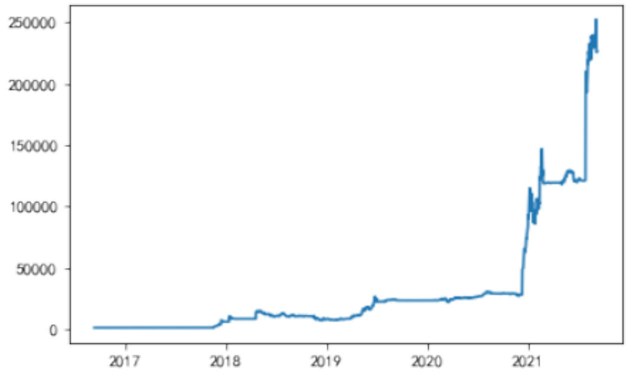


Fig. 15 Bit coin can be traded daily, but gold is only on the opening day (considering the form).

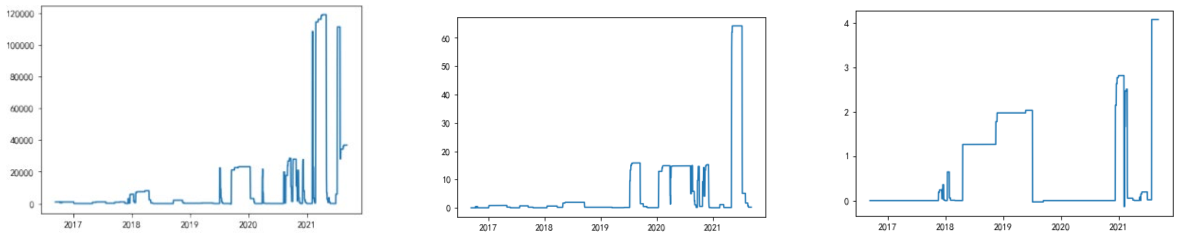


Fig. 16 Comparison of cash, Bit coin and gold holdings

Total assets = Gold held share * USD (PM) + Bitcoin held share * Value + cash held
Total assets = Gold held share * USD (PM) + Bit coin held share * Value + cash held share

3. Problem 2

Return is the ultimate goal for buyers, so we determine the influencing factors most closely related to price fluctuation, and further analyze the key influencing factors.We take gold as an example, and we can argue that the model in the first question provides the best strategy for traders through the correlation analysis, the stepwise regression-wavelet neural network model, and the SRA-WNN model.

3.1 Step 1: The correlation analysis

Correlation analysis refers to the analysis of two or more variables with correlation to measure the close degree of correlation of the two variable factors.We used CY, SR, WJ, HJ, and SS for the number, months, time, geographical location, and type of gold value, respectively.The degree of correlation between the relationship and the measured variables quantified by the study was calculated using the correlation coefficient.Random allocation data set method: the data set is randomly divided into the training set and the test set, usually randomly allocated in proportion, where 3 / 4 of the data set is the training set used to simulate the model, and 1 / 4 of the data set is the test set used to verify the obtained model. To visually represent the correlation between the metrics in different years, in 2020, we represent it using Pearson’s correlation coefficient:

$$R=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})\sum_{i=1}^{n}(y_i-\bar{y})}$$

x and y are different study variables, x and y with i are values of two variables and x and y with "-" are the mean of the study variables.There is no relationship with the degree of change, location and size of the two study variables, so it can be used to measure the nature of the correlation between the two study variables.

When R=0, unrelated,

When 0<R<1, positive correlation
When -1<R<0, negative correlation
When absolute value of R>0.8, high correlation
When 0.5<absolute value of R<0.8, moderate correlation
When 0.3<absolute value of R<0.5, low correlation
When absolute value of R<0.3, very low correlation

Table 2. The correlation coefficient of the five indicators is as follows:

	SR	CY	HJ	WJ	SS
SR	0.534	0.478	0.567	0.439	0.445
CY	0.475	0.640	0.555	0.561	0.535
HJ	0.530	0.546	0.737	0.477	0.481
WJ	0.502	0.638	0.578	0.736	0.603
SS	0.499	0.633	0.579	0.589	0.594

According to the calculation of Pearson’s correlation coefficient, for the three indicators of month, number and position, the monthly data and the number of gold value data are significantly consistent, which can be considered to be consistent with the law mentioned in the question, while the geographical location strategic value is not significant. We also analyzed the growth rate of each index, and showed the results from five types of gold value.

Table 3. The calculated value of each index is as follows:

Number (0-159)	Type (0-5)	SR	CY	HJ	WJ	SS	key
157	3	3677.459	142.9103	171.0768	70.99567	105.9173	141.7785
138	1	1686.243	24.30642	114.9405	78.01256	47.87772	81.37279
129	1	-16.7051	203.0133	129.2076	1083.489	97.51777	159.925
102	3	471.7582	160.1092	174.3204	185.6692	245.5107	198.9721
89	1	125.4604	37.22315	65.64036	24.63325	957.1373	88.14718
89	1	491.7303	52.29399	102.7903	50.77946	47.87772	76.36485
117	3	-5.31837	137.0271	436.5097	63.97179	65.01954	79.31279
96	1	505.1196	27.47039	40.54416	33.59149	79.98074	56.65529
129	1	164.4077	64.21041	177.1202	210.3	63.92522	111.9048

3.2 Step 2: Step regression-wavelet neural network model

Stepwise regression analysis (SRA) is a method to establish the optimal regression equation that contains only significant impact factors, while the remaining factors with no significant effects are removed together. The basic principle of SRA: based on the initial regression equation, statistical model, and based on the influence factors obtained from modeling the market valuation level, the market valuation level gives one yuan regression to each influence factor one by one, and then the partial regression square sum introduces the regression factor from large to small. The index is then F-significance test, introduced if significant, and in the currently introduced factor, find the smallest partial regression square factor, perform the F-significant test, and if not significant, the index is removed from the regression equation. The introduction and elimination cycles were repeated until no significant factor was introduced and no non-significant factor elimination, and the cycle was stopped. The optimal regression equation is then obtained. The factors included in the optimal regression equation, which is the one with a significant influence on the market valuation level, can serve as input factors for the wavelet neural networks.

Wavelet transform refers to selecting a basic wavelet function $\varphi(t)$ (also known as the mother wavelet function) to translate τ and make an inner product with the signal $x(t)$ to be analyzed at different a scales. The equivalent time-domain expression is :

$$f_x = (a\tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(\omega)\varphi(a\omega) \exp^{j\omega} dt$$

Combined with the previous index weight analysis, we establish the Fisher discriminant function, process the given index data, comprehensively analyze the liquidity index data, and bring it into the discriminant function, so as to determine the value effect of the patent and quantify the processing.

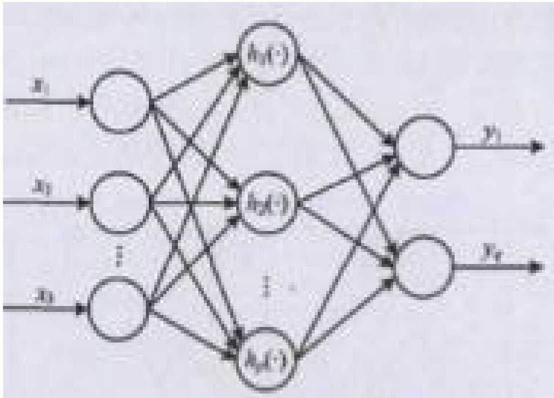


Fig. 17 Topology of Wavelet Neural Network

3.3 Step 3: Solution of the $SRA-WNN$ model

According to the analysis of the model grey association method, and the determination of the correlation significant factors in this model.Taking the fundamental indicators and liquidity indicators of predicting China's patent market as an example, we take the market value, technical value, legal value, strategic value and economic value as the influencing factors, which are labeled... and, respectively $x_1, x_2, x_3, x_4, x_5, x_6$.

Based on the number of gold value, month, time, geographical location, type and actual data, the optimal regression equation of value level was obtained.

The initial factor corresponding data were fed into the WNN model and the correlation significant factor corresponding data into the $SRA-WNN$ model for training, and then the predicted sample for prediction.Prediction analysis was also performed using a stepwise regression.After multiple data training and prediction, better prediction results were obtained for each model.

$$w_1 = -0.95X_1 - 0.42X_2 + 0.03X_3 - 0.77X_4 + 0.56X_5 - 0.57X_6$$

$$w_2 = 0.08X_1 + 0.14X_2 + 0.09X_3 + 0.25X_4 + 0.06X_5 + 0.59X_6$$

4. Problem 3

We examine the impact of commission changes on the results, take the scope of the two, the middle interval of 0.01:

Gold commission: 0.01-0.11

Bit coin commission: 0.01-0.21

After a simulated transaction, we have a commission-total assets data excel.

And drew the maximum total assets trend chart of different fees:

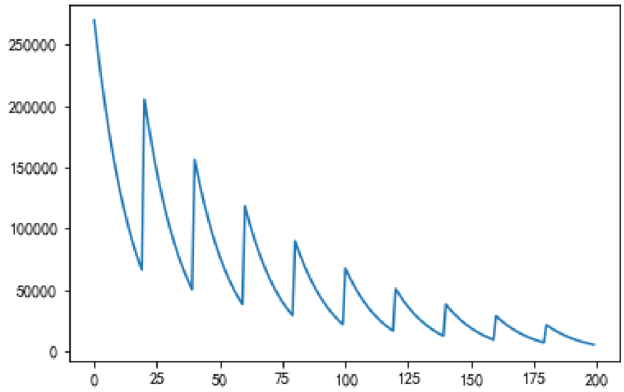


Fig. 18 The maximum total assets trend chart of different fees

5. Conclusion

In order to maximize the investment utility of traders and minimize the value at risk, this paper establishes a bull bear market judgment model, transaction risk analysis and time series prediction model to combine the three investments of gold, cash and Bit coin in the United States and evaluate their future value. It is obtained that the total assets on September 9, 2021 are 25W. Then we use correlation analysis Stepwise regression wavelet neural network model and $SRA-WNN$ model are solved to demonstrate that the first question is the best strategy. Finally, in order to carry out sensitivity analysis, the Commission is set as a variable to investigate the impact of Commission on transaction results.

References

[1] Yan Wang, Yuankai Guo, [J]. Forecasting Method of Stock Market Volatility inTim Series Data Based on Mixed Model of ARIMA and XGBoost. China Communications, 2020, 17(03): 205-221.

[2] Guo Sihan, [D]. Research on Bit coin price prediction and trading strategy based on improved recurrent neural network. Shanghai Normal University, 2021 DOI:10.27312/d.cnki. gshsu. 2021.002181.

[3] Zhang Wenhua, [J]. Algorithm improvement and Simulation of ARIMA model for time series prediction. Information and computer (theoretical Edition), 2021, 33 (05): 53-56.

[4] Luo hongben, [J]. Research on intelligent hybrid prediction of financial time series based on Grey Arima. Financial theory and practice, 2014, 35 (02): 27-34.

[5] Qiusen, [D]. Time series prediction of urban rail transit passenger flow based on wavelet analysis and neural network. Beijing Jiaojong University, 2017.