Problem Chosen	2022 MCM/ICM Summary Sheet	Team Control Number
C		2200024

Think like a Turtle: A Low-Risk Strategy based on the Past

Summary

Nowadays, bitcoin has huge potential in the financial sector due to its high yield, unregulated and tax-free nature. Thus, bitcoin is sometimes called digital gold, which can replace gold to hedge inflation and become a new hedge asset. It can complement gold for hedging, which is of great interest to all kinds of investors in the financial field.

Our trading strategy for Bitcoin and Gold investments incorporates the ideas of the **Turtle Trading Strategy**, with **model simplification, risk control and trend mastery** as the main goals of our trading strategy. Instead of using prediction about future market prices to decide when to trade, we use **only past data** for training in backtesting to avoid errors similar to recency bias, which allows traders to **minimize the risk** of losing money in the trading process. The trade model will be based on three aspects: **Tendency, Risk and Circuit breaker price**.

As for Tendency, we set seven indicators to measure the trend, which is **Moving Average Convergence and Divergence (MACD), Market Price of the day** and **Exponential Moving Average (EMA)** for a week (7 days), a month (30 days) and a season (90 days) respectively. Then, we used the **Multivariate Linear Models** to train out the weight of each indicator for measuring the trend. However, this model suffers from the problem of overfitting the trend and is not effective in the performance process. Thus, we apply the **Long Short Term Memory (LSTM)** model to improve our training method and get more reasonable indicator weights by minimizing loss value.

To measure the Risk, we use **Variance** and the **Technique for Order Preference by Similarity to Ideal Solution (TOPSIS)** to indicate the risk of asset loss and determine the goods we use to preserve our assets. We use the **K-Means Method** to obtain the Circuit breaker price from the daily market price change.

Then, we build our trading model based on the above information. The essence of the model is to determine the following two aspects: the **Timing of trade** and the **Number of prices traded**.

For the determination of the appropriate trade time, this is determined by the position of the trend line relative to the true value, as analyzed in the trend section. And the number of prices traded is determined by the relationship between the time interval from the last exchange and the Circuit breaker price.

At the end of the article, we perform **a sensitivity analysis of the model relative to commissions.** In general, a reduction in commissions would give us a higher rate of return. The impact of transaction fees on returns is higher for gold compared to bitcoin.

Keywords: Turtle trading strategy, LSTM, Multilinear model, EMA, MACD, K-Means, TOPSIS

Content

1. Introduction

1.1 Background

Because of its unique form and rare characteristics, gold has been regarded as a symbol of wealth since the early days. After the collapse of the "Bretton Woods system", the price of gold is no longer determined by the government, which means that gold really entered the market. And the global economic integration process accelerated, making the gold investment widely sought after. This has made gold a very important investment channel for a long time.

Bitcoin, as a new type of digital virtual currency, which relies on blockchain technology and decentralization, its own anonymity and other characteristics, has also attracted the attention of many investors in recent times. Today, its emergence as a new generation of investment tools has had a big impact on traditional investment tools.

Nowadays, with the development of the social economy, the investment market has developed significantly, which makes more people invest their assets in the form of assets to achieve their own property value preservation and appreciation. However, since investment is a risky behavior, some inexperienced investors often use their past investment experience and investment intuition to make investments, which may cause serious losses to a large extent. Therefore, it is necessary to designate a plan to guide investors to achieve reasonable and correct investment through professional analysis.

Gold and Bitcoin represent traditional and new investment tools respectively, and they have different advantages. In comparison, bitcoin has a high yield but high volatility, while gold is less volatile but has a relatively low yield. Thus, the strategy of choosing between these two assets based on the information known at the moment is of great practical significance.

1.2 Restatement of the Tasks

Our task is to develop the model to determine the trader's investment behavior each day for the assets in their portfolio using only the past daily bitcoin and gold price flows to date.

In order to be able to specify the best trade strategy based on known information, we will address the following aspects.

1) Provided **best daily trading strategy** according to the model-based only on the previous price data.
2) Evaluation of the model to demonstrate the **feasibility** of the provided strategy.
3) Determine the **relationship** between the **strategy and transaction costs**, and articulate the impact of transaction costs on the strategy and its outcomes through sensitivity.

2. Assumptions

1. **There is no cost for a trader to hold an asset.** For each trader, the overhead for holding the property is varied under different conditions. If we take this instability into account, it may result in a model that is not very applicable and thus difficult to apply to most traders.
2. **Ignore inflation.** The value of cash in the hands of traders will not be affected by inflation. Excluding the effect of inflation on the value of cash reduces the complexity of the model and makes it easier for traders to understand. At the same time, cash inflation has had relatively little impact on the price of cash in recent years, so this also has less impact on the accuracy of the model.
3. **The market price of bitcoin and gold will fluctuate between its true value.** Although the price of goods can change from moment to moment, there is often a value that the goods themselves should have due to their unique properties. Therefore, the price of goods fluctuates up and down around their value. In the process of analysis, we also base our model on such assumptions.
4. **The total value of the trader's asset holdings is considered directly as the trader's ultimate total return.** In the final calculation of the return to the trader, since both gold and bitcoin have their own value at that point in time, which means that both instruments have not only their market value but also the opportunity cost of gaining future benefits. Therefore, we do not convert these non-cash assets to cash and then compare them.
5. **A commission will be charged for each sale and purchase, and the percentage of the commission is fixed - 2% for bitcoin, 1% for gold.** If the rate of costs changes too frequently in various cases, it is difficult to build a stable and valid model. Moreover, since changes in the rate of costs are not very common, the model does not take into account the possibility of changes in the rate of costs within this 5-years time frame.
6. **Assume that the maximum acceptable loss to the trader is capped at $1000 and don't use financial leverage.** Since the return of investment is proportional to its risk, an excessive pursuit of return without consideration of risk may lead to a significant reduction in the viability of the trading strategy. Therefore, setting an upper limit on the maximum acceptable amount of loss helps us to ensure that the risk is acceptable to a certain extent. In this case, we use a much safer method to control the maximum loss will not over $500.
7. **Future is hard to predict.** In addition to the value of the goods themselves, there are many factors that may have an impact on the market price of the goods, including the international situation and national policies.

3. Analysis of the Issue Paper

Through the topic analysis we can get that if we need to develop an effective trade policy, we need to build models to identify the right time to trade and the right number of prices to trade. And to determine the model, we will analyze three aspects, the market price, EMA, MACD to estimate the trend based on historical data, the variance and TOPSIS to analyze the risk, and K-means and interval length to set the Circuit breaker price.

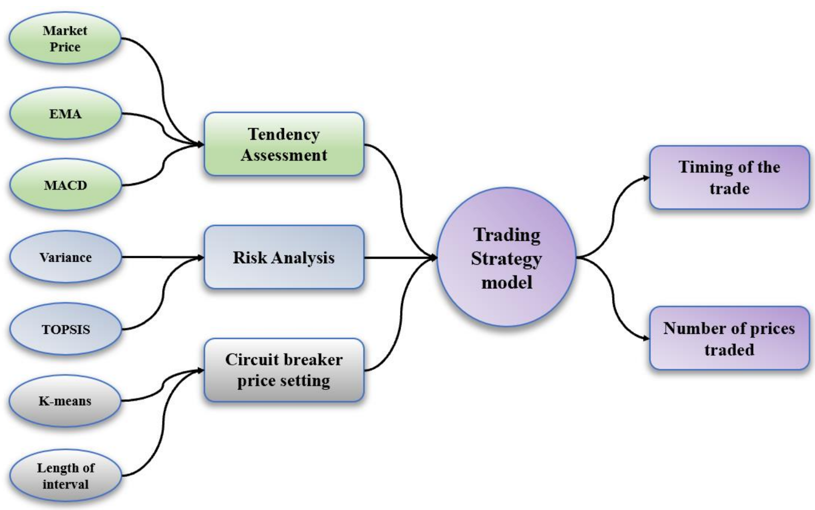


Figure 1 Architecture of the whole article

3.1 Preliminary analysis

The provided bitcoin and gold historical price information is shown as follow:

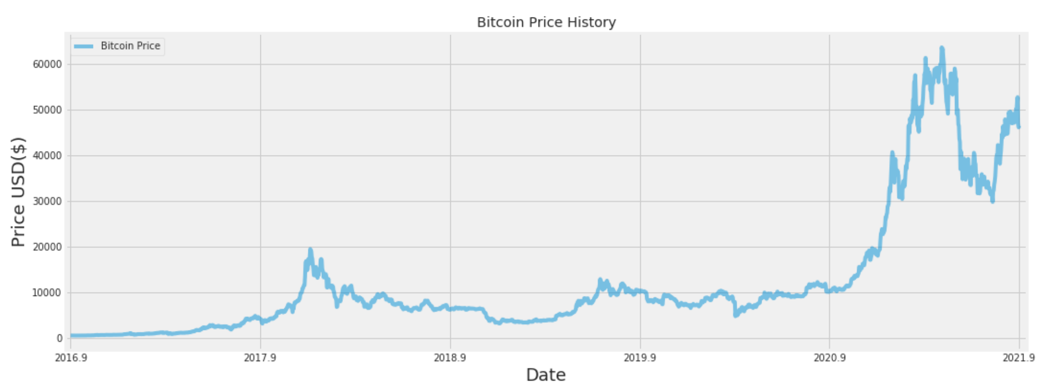


Figure 2 Bitcoin historical information

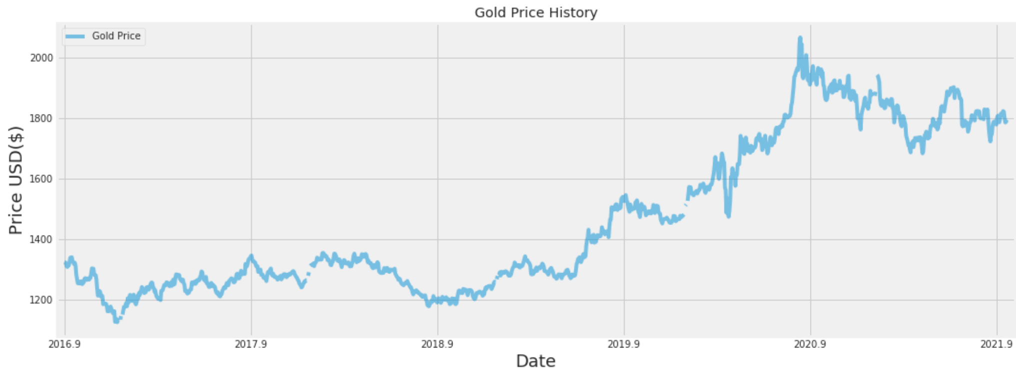


Figure 3 Gold historical information

It can be easily found that at the beginning, bitcoin almost cost no money while gold's price was pretty high relatively, but no soon, the price of bitcoin rose dramatically. So, intuitively, investing more money in bitcoin at the beginning brings more profit. And we should build a model continuously investing a large amount of money in the bitcoin market. However, we came to this conclusion from the perspective of gold, while in a real-world situation we cannot do that, so we need to find some metrics telling our model - "It's time to invest!"

3.2 Correlation analysis

Since we want to invest in a combination of gold and bitcoin, for these two goods we can first perform correlation analysis to find out if there is a linear relationship between them. The correlation coefficient is a quantity that studies the degree of linear correlation between variables and can indicate exactly the **degree of correlation** between two variables. It can be calculated by the following formula:

$$\boldsymbol{r}(\boldsymbol{X},\boldsymbol{Y})=\frac{\boldsymbol{Cov}(\boldsymbol{X},\boldsymbol{Y})}{\sqrt{\boldsymbol{Var}[\boldsymbol{X}]\boldsymbol{Var}[\boldsymbol{Y}]}} \tag{1}$$

After the calculation, we get the correlation coefficient between gold and bitcoin price from 2016 to 2021 is 0.59396007, which is quite small and it means that gold and bitcoin are weakly correlated in a linear space. Therefore, we can consider investing in gold and bitcoin **separately** when we develop our trading strategy. The relationship between the two is represented by the following figure:

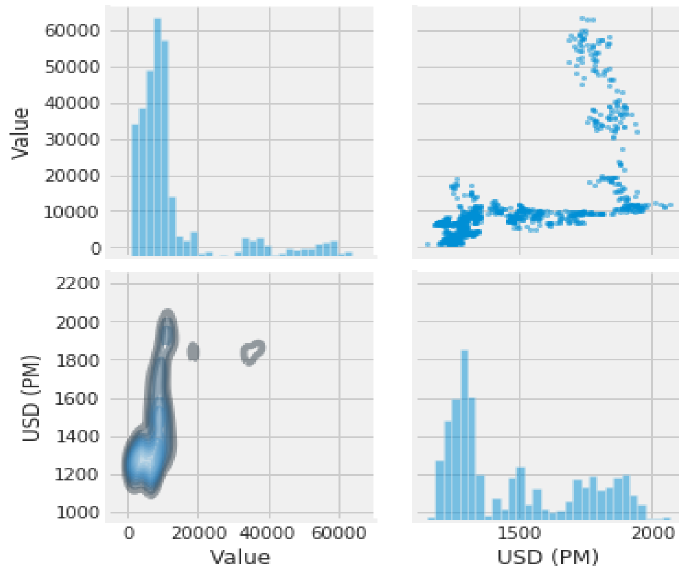


Figure 4 Correlation Analysis

The canvas in the upper left corner and the canvas in the lower right corner represent the **histograms** of bitcoin and gold values, respectively, from which we can know the value and general changes of their prices. The canvas in the upper right corner is a **scatterplot** of the correlation between the gold price and bitcoin price, with the gold price on the x-axis and bitcoin value on the y-axis. The lower-left corner is the **kernel density estimation graph** which is used to calculate the density probability between the two.

4. Notation

Table 1 Notation

Symbol	Abbreviation
$\boldsymbol{T}$	The length of the time window
$\boldsymbol{r}$	Correlation coefficient
$\boldsymbol{D}$	Second normal form distance
$\boldsymbol{\vartheta}$	The average rate of ROI change under different transaction costs
$\boldsymbol{\kappa(A)}$	The condition number for A
$\boldsymbol{\mathcal{R}}$	The rate of change at least
$\boldsymbol{\Omega}$	Lower-Bound
$\boldsymbol{n}$	Data Length
$\boldsymbol{Cash_{init}}$	The initial cash we have (\$1000)
$\boldsymbol{currentPrice}$	The current price of the bitcoin or gold
$\boldsymbol{numHold}$	The unit of goods (bitcoin/gold) we hold
$\boldsymbol{Wealth}$	Our all wealth
$\boldsymbol{ExpectRatio}$	The proportion of the total specific goods' value we hold over our all wealth

5. Data Preprocessing

5.1 Data Integrity Detection

Through mapping observation, it was found that there were 10 days of missing data on gold prices between September 2016 and September 2019, which need to be dealt with.

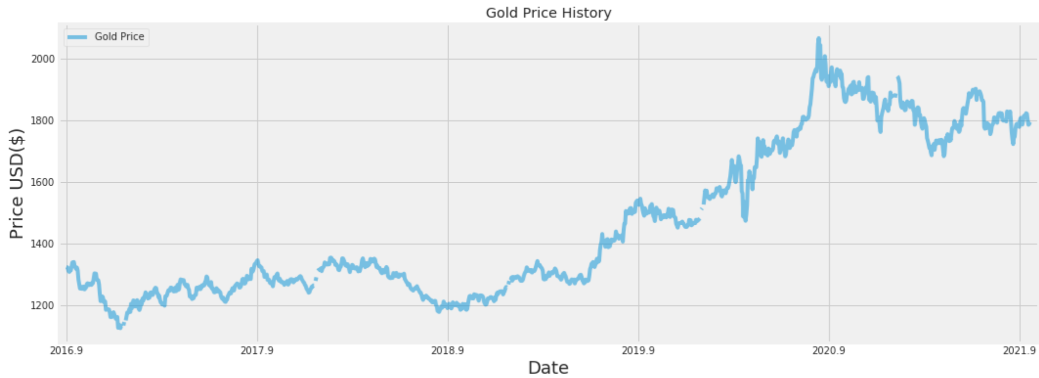


Figure 5 Gold historical information

Because the price data is very random, which is vary according to market changes, filling data can destroy the accuracy of the data either through interpolation or edge padding. Therefore, we believe that **removing null data** is the best choice.

5.2 Data Noise Reduction

5.2.1 Noise Reduction Requirements

As the representative of financial time series, transaction price cannot avoid the noise problem, so the research on denoising problem is very important. In another hand, the **cleanliness** of the data can have a significant impact on the results, both for later data metrics and for the input to the neural network.

5.2.2 Wavelet Denoising

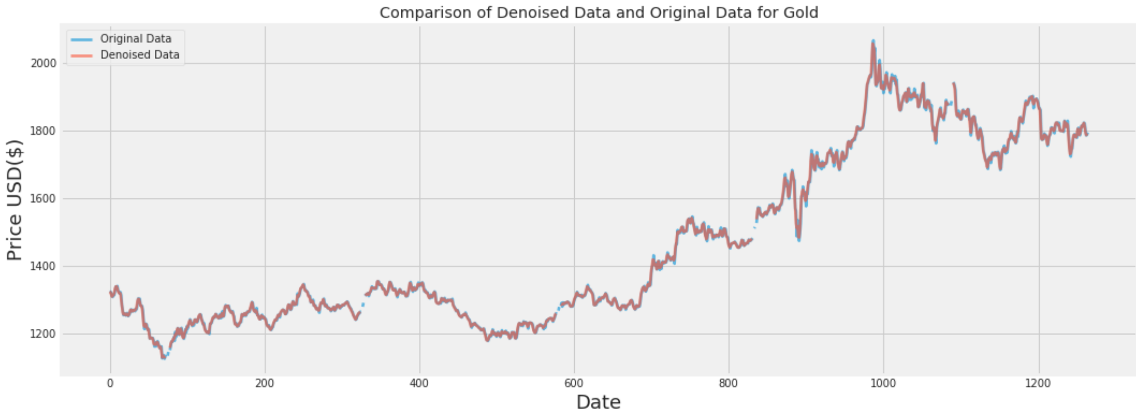
The price information is regarded as a **signal**, and the wavelet coefficient generated by the signal contains the important information of the signal. After the signal is decomposed by the **wavelet**, the wavelet coefficient is larger, and the wavelet coefficient of the noise is smaller, and the wavelet coefficient of the noise is smaller than that of the signal.

1. Wavelet decomposition of the signal

We select a wavelet and determine the level of a wavelet decomposition as 5, and then perform a **5-level wavelet** decomposition calculation on the signal.

2. Threshold selection

A suitable threshold, the wavelet coefficients larger than the threshold are considered to be generated by signals and should be retained, and those smaller than the threshold are considered to be generated by noise and are set to zero to achieve the purpose of denoising. According to the tuning result, we choose **0.5 as the parameter**. Similarly, the discriminant function of the threshold is **the Garrote function**. The noise reduction effect on gold and bitcoin is as follows:



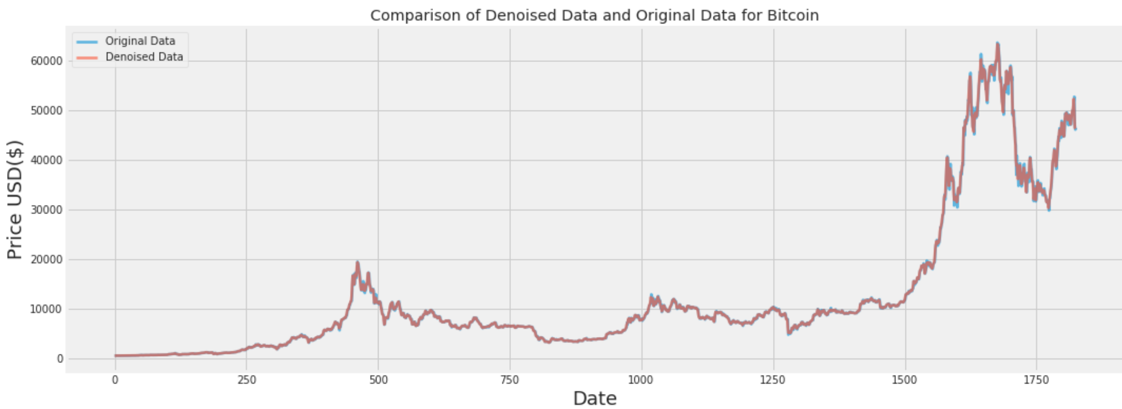


Figure 6 Comparison of Denoised Data and Original Data

6. Building Investment Models

For how to develop the optimal strategy for trading to maximize returns. The model essentially needs to solve the problem of deciding when to buy and sell, and how much each transaction volume is. And therefore, we start analyzing with three aspects – **Trend**, **Risk** and **Trading curb**, to create the metrics in order to build up our final trading model.

6.1 Trend of trading prices

6.1.1 Rationality of indicators

Value: The basic information given accurately reflects the specific information of the data.

Exponential Moving Average (EMA): A trend-type indicator that performs a weighted arithmetic average of the value to determine the changing trend of the price in the future and focuses on the weight of the current price (current) market. The formula is as follows:

$$\boldsymbol{EMA_N(x_n) = \frac{2}{N+2}\sum_{k=0}^{\infty}\left(\frac{N-1}{N+1}\right)^k x_{n-k}} \quad (2)$$

which N is the (window) period and n^{th} data. The window size we set a week, a month a quarter and represent by $EMA_7, EMA_{30}, EMA_{90}$ to represent respectively.

Moving Average Convergence and Divergence (**MACD**): A comprehensive class indicator, on the basis of moving average, a class penalty term is added to offset the influence of some false information, which can also reflect the speed of price changes to a certain extent. T MACD turns from negative to positive, which is a buy signal. When the MACD turns from positive to negative, it is a sell signal. When the MACD changes at a large angle, it means that the gap between the fast-moving average and the slow-moving average is widening very rapidly, representing a change in the general trend of the market. The formula is as follows:

$$\begin{aligned} \boldsymbol{DIF(x_n)} &= \boldsymbol{EMA_F(x_n) - EMA_S(x_n)} \\ \boldsymbol{MACD(x_n)} &= \boldsymbol{2[DIF - (EMA_M(DIF(x_n))]} \end{aligned} \quad (3)$$

S, L, M are set of arguments to this function, respectively represent fast movement number, slow movement number and difference value movement number, the default parameter Settings are 12, 26 and 9.

Trend Mean Line: It is a custom indicator used to reflect the price trend to a certain extent, expressed by the average value of a period before and after the period. The calculation formula is:

$$\mathbf{P_T = \frac{1}{2t} N(P_{T-t} + P_{T+t})} \quad (4)$$

6.1.2 Model establishment

6.1.2.1 Multivariate linear model

A multiple linear regression model is a **linear regression model** that contains multiple explanatory variables, and it is used to explain the linear relationship between the explained variable and multiple other variables.

We use a multivariate linear model to fit the trend mean to calculate the optimal weight for each parameter. First, we need to divide the test set and the training set. Take 80% of the data in the previous period as the training set, and the remaining 20% of the future (later) data as the test set, the independent variables are Value, EMA_7, EMA_{30}, EMA_{90} , and MACD.

The dependent variable is the trend mean line. The advantage of this score is that the model has a certain predictive ability because the data we can use in each transaction is only the past so that it can be closer to the future. The form of writing the target as a matrix is as follows, where the matrix X is our existing data, and the five columns represent the five indicators. Our goal is to train W which can fit y as much as possible so that the global EMSE (loss) is minimized.

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{14} & x_{15} \\ x_{21} & x_{22} & x_{23} & x_{24} & x_{25} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{n-1,1} & x_{n-1} & x_{n-1,3} & x_{n-1,4} & x_{n-1,5} \\ x_{n1} & x_{n2} & x_{n3} & x_{n4} & x_{n5} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_{n-1} \\ w_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \\ y_n \end{bmatrix} \tag{5}$$

With W, the proportion of each parameter, we can more sensitively represent the trend of the data, rather than a fixed mean algorithm. The following data is the result of training:

Figure 7 Gold linear model training result

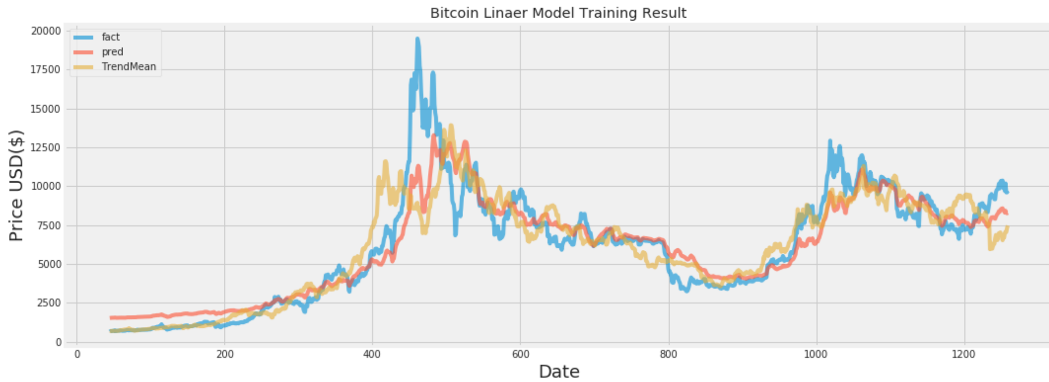


Figure 8 Bitcoin linear model training result

Bitcoin and gold test results as follows：

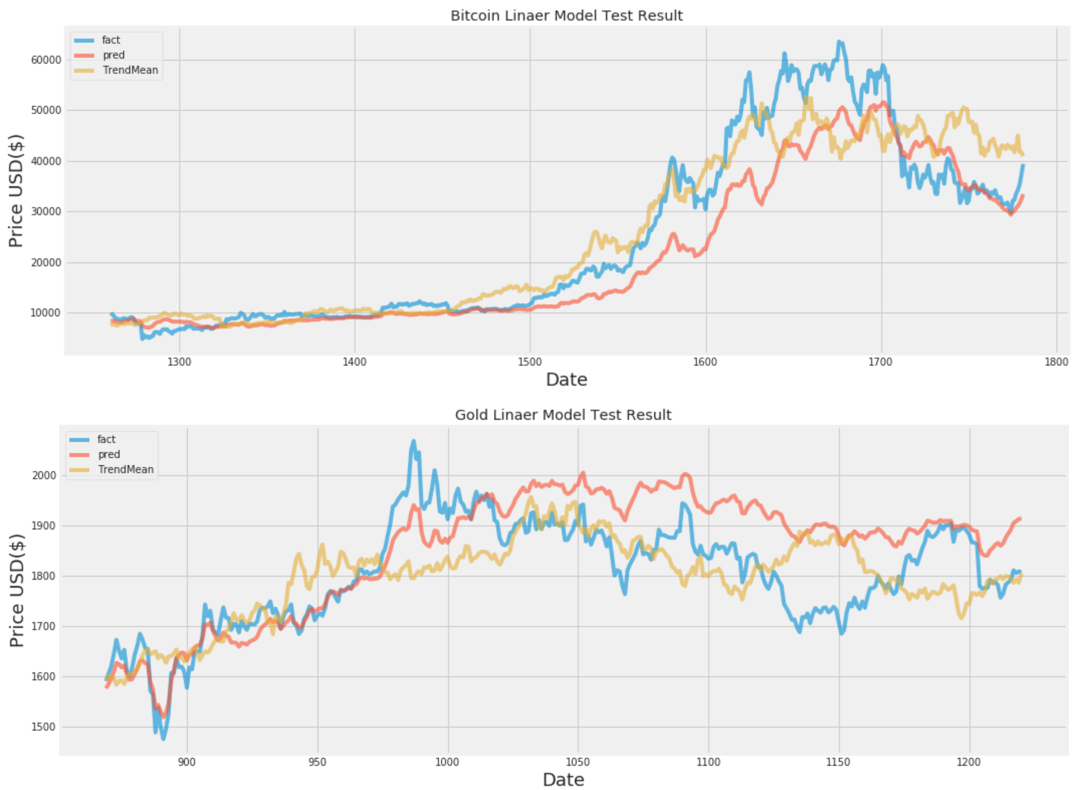


Figure 9 Linear model Test result

As a result, the parameters are adjusted, and the final accuracy rate is above 93%, so there is enough evidence to show that this result can represent the model trend.

6.1.2.2 LSTM model

Since the Multilinear model is kind of overfitting and seems not perform well for the trend indicating (fluctuates too much), we try to construct an LSTM model to give a better performance.

The **Long Short Term Memory** (LSTM) model is a specific form of **Recurrent Neural Network (RNN)**. The LSTM model is based on the RNN model, which makes the RNN really effective to use the long-range temporal information with 3 types of gates, which also solves the gradient dispersion problem in RNN.

LSTM model uses memory cells, like long-term and short-term memory modules, instead of using traditional neurons. Each memory cells have input gates, forgetting gates, memory units and output gates.

- Input gate: It's used to memorize some information of the current input information with layers whose activating function are sigmoid and tanh. The sigmoid layer determines which value we will update, and the tanh layer is used to create new candidate values vector $\widetilde{C_t}$

$$\begin{gathered}\boldsymbol{i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)} \\ \boldsymbol{\widetilde{C_t} = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)}\end{gathered} \tag{6}$$

- Forgotten gate: It’s utilized to selectively forget some information of the past. The gate will output a number between 0 to 1 to discard the degree of information where 1 means “completely reserved” and 0 means “completely discarded”

$$\boldsymbol{f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right)} \tag{7}$$

- Memory unit: It’s like a bridge connecting the past and present information and combines these memories.

$$\boldsymbol{C_t = f_t * C_{t-1} + i_t * \widetilde{C_t}} \tag{8}$$

- Output gate: It’s used for selectively outputting some information with sigmoid layer and tanh layer where sigmoid layer determines which values can be output, and the tanh layer shrinks the value to be ranging of -1 and 1.

$$\begin{gathered}\boldsymbol{o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o),} \\ \boldsymbol{h_t = o_t \cdot \tanh(C_t)}\end{gathered} \tag{9}$$

Among them, h_t is t-time hidden state, W_i, W_C, W_f, W_o are the weight metrics, and b_i, b_C, b_f, b_o are bias for the corresponding units. These are all model training parameters. The LSTM model is shown as follow (Notice that here $z = W \cdot [h_{t-1}, x_t]$):

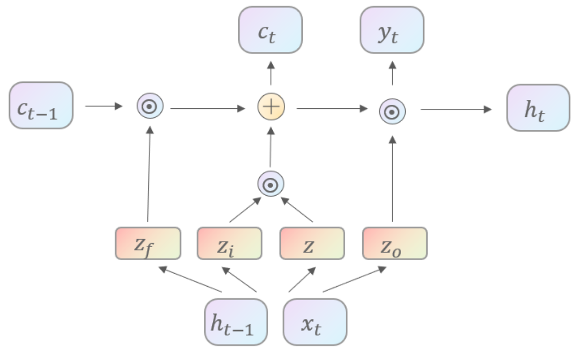


Figure 10 LSTM Architecture

To train the LSTM model, we take a sliding window over the datasets to create time-step frame data. In the following picture, the rectangles are our processed data, and the raw date is on the top. We take 60 days as the width of the window, using it to catch a data frame for one time-step, and slide the window each step for 1 day. Finally, we got a 3-d matrix as the input, and each dimension means "time-step", "batch" and "features" correspondingly.

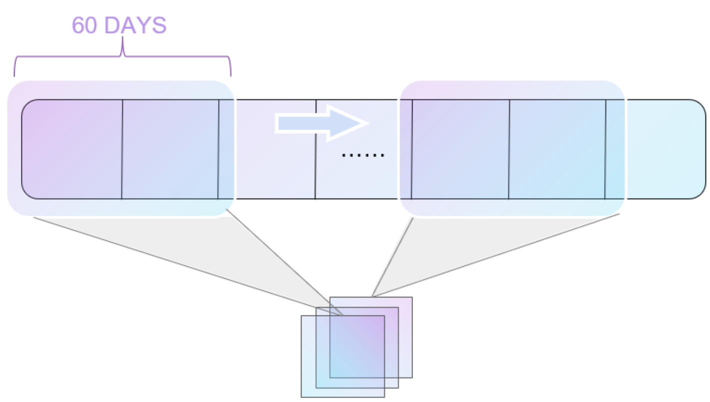


Figure 11 Data preprocessing for LSTM training

Then we construct our neural network in the following architecture:

Figure 12 LSTM model architecture

We construct a 3 layers neural network to detect the trend feature, and use a 7 cells LSTM layer in our first layer as the features detector fetching 7 features in our training data. Then we use the mean value of a period as the central point's true value (y-label). After tuning the hyperparameters, the model running result is shown as below.

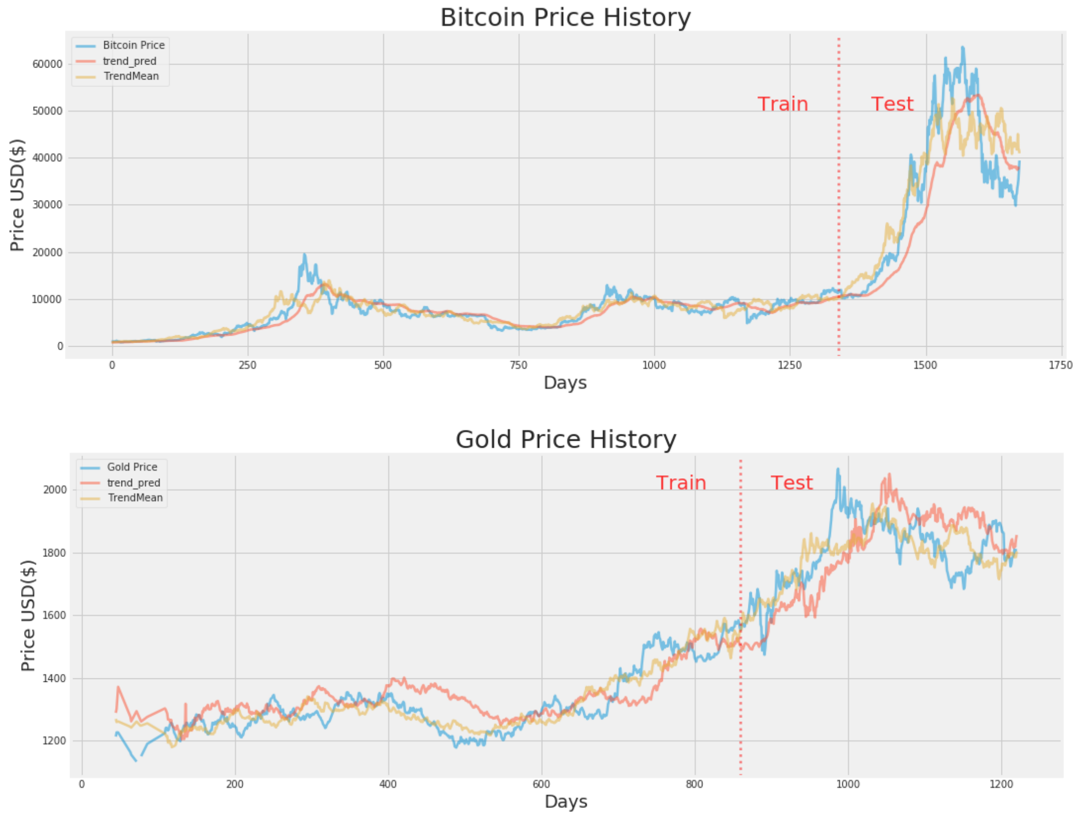


Figure 13 LSTM model fitting result

As we can see, the predicted trend is generally smoother than the original data, and period of time that the blue line over the red one is always at a pretty high price – so we can buy gold or bitcoin when the current price is below our predicted trend line; otherwise, we can consider selling some of them.

6.2 Risks in the investment process

6.2.1 Variance

Variance is a measure of the degree of dispersion when measuring a random variable or a set of data, and is meant to be the average of the squared values of the difference between each sample value and the average of all sample values. The formula is shown below.

$$\mathbf{s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - x)^2} \tag{10}$$

Since variance measures the magnitude of fluctuations in the data, we used variance for the stability of the market price of bitcoin and gold in our risk measurement for them. The following graph shows the logarithmic comparison of the variance between the two goods:

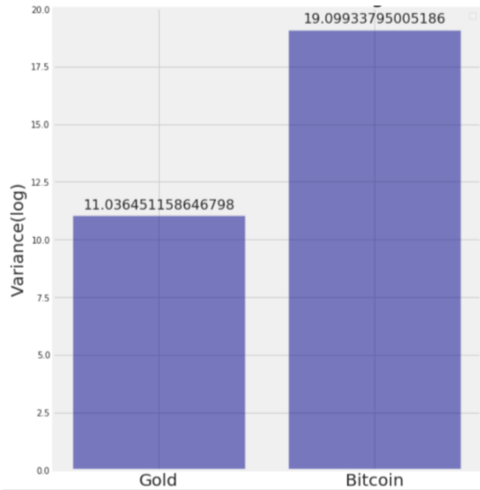


Figure 14 Variance of bitcoin and gold

As can be seen from the chart above, the variance of gold is very small compared with bitcoin. This also means that the risk of **gold in the market is relatively low**. Therefore, in our investment strategy, we **use gold as a hedging asset**.

6.2.2 TOPSIS Risk Analysis

Technique for Order Preference by Similarity to an Ideal Solution(Topsis) is a commonly used comprehensive evaluation method, which makes full use of the information from the original data, and its results can accurately reflect the gaps between the evaluation programs. Thus, it is also suitable for analyzing the risk score of this data.

6.2.2.1 Build Feature

In the establishment of the trend model, we have obtained a trend line that can represent the model, and subtract the trend line from the true value to obtain a data near 0. When the data is very small (it can be negative), it means that there is a future A larger probability value will increase, which means that the risk of loss will be smaller; on the contrary, if the data is larger, the risk of loss will be greater. Therefore, this is a very small metric (smaller is better) and then use it as a feature as an input to the model.

6.2.2.2 TOPSIS Score Calculation

1. Positive indicator: Because the smaller the relative distance, the smaller the risk. Therefore, it is necessary to convert the minimum value index into the maximum value index.

This method uses the distance scale to measure the sample gap, and the use of the distance scale requires the same direction processing of the index attributes. The algorithm is:

$$X_{i_{new}} = MAX(X) - X_{x_{old}} \tag{11}$$

2. Data Standardization: Make feature variables have the same scale, control the value within a certain range. The algorithm is:

$$\boldsymbol{Z_i = \frac{X_i}{\sqrt{\sum_{i=1}^{n} x_i^2}}} \tag{12}$$

3. Determine the best and worst options: The optimal solution is a parameter calculated by taking the ideal value (maximum value) of each indicator:

$$\boldsymbol{Z^+ = \{max(Z_i)\}} \tag{13}$$

The worst solution is a parameter calculated by taking the most unsatisfactory state (minimum value) for each indicator:

$$\boldsymbol{Z^- = \{min(Z_i)\}} \tag{14}$$

4. Proximity to the optimal solution and the worst solution (second normal form distance): Among them, w is the attribute weight (importance). According to the actual situation, only the distance value and the trend line distance need to be considered here, so the weight is set to 1.

$$\boldsymbol{\begin{aligned} D_i^+ &= \sqrt{\textstyle\sum_1^n w(Z^+ - Z_i)^2} \\ D_i^- &= \sqrt{\textstyle\sum_1^n w(Z^- - Z_i)^2} \end{aligned}} \tag{15}$$

5. Calculate the **closeness Score** of each evaluation object to the optimal solution. The algorithm is:

$$\boldsymbol{Score = \frac{D^-}{D^+ + D^-}} \tag{16}$$

The range of Score is $0 \leq \boldsymbol{Score} \leq 1$. The closer the result is to 1, the closer the distance to the optimal solution and the farther the distance to the worst solution, the better the evaluation object is. Finally, we use the final score as an indicator of risk.

6.3 Trading Curb

As the saying goes, “Don't put all eggs in one basket”, investing most of your money in one goods at a time is also not a good idea, although sometimes the price is very low, it can be even much lower, as we never know when the price hits rock bottom. Therefore, we need trading curb to provide a guarantee that we won’t spend too much in a goods at a time so that we have cash to buy more if the price comes even lower.

We make a tiny program to solve out the value of trading curb for bitcoin and gold's buying and selling based on the following equation. The number of transactions (T) required under the assumption that there is no loss or profit, and only transaction costs ($\varphi * H^r$) should be calculated until the left cash down to some certain percentage of the original (δ). Each time the cash will update by the function:

$$\boldsymbol{H^{r+1} = H^r \left(1 - (\varphi * 2)\right)} \tag{17}$$

When $H^r < \delta H$, then stop the iteration and get the number of transactions. In this case, the commission is 1% for gold transactions and 2% for bitcoin. Taking the higher commission for our computation, let's say 2% for φ and 50% for δ, and solve the problem, we got the result – 32 times, which means if the price keeps fixed, after 32 trades, our wealth will fall to a half of the initial situation, $500.

Then the interval (I) between two transactions can be calculate as follows:

$$\boldsymbol{I = \frac{5 \times 365}{2 \times T}} \tag{18}$$

After getting the interval (I), the value of trading curb (L) can be calculated by:

$$\boldsymbol{L_{ij} = E_{ij} \times I_{ij} \times 1.25} \tag{19}$$

where $i = bitcoin\ or\ gold,\ \ j = up\ or\ dow$, and E in the formula means the expectation of the corresponding types' value, for example, the expectation of the situation that the second day's bitcoin price is higher than the first day.

6.3.1 K-means

To find the expectation of the corresponding types' value, we use K-means, and considering its clusters' kernels as the expectations. We got the result shown in figure 14 and figure 15:

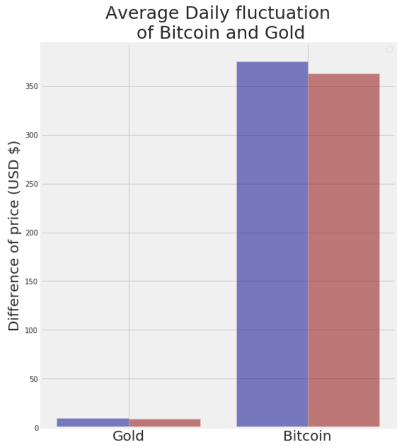


Figure 15 Average Daily fluctuation

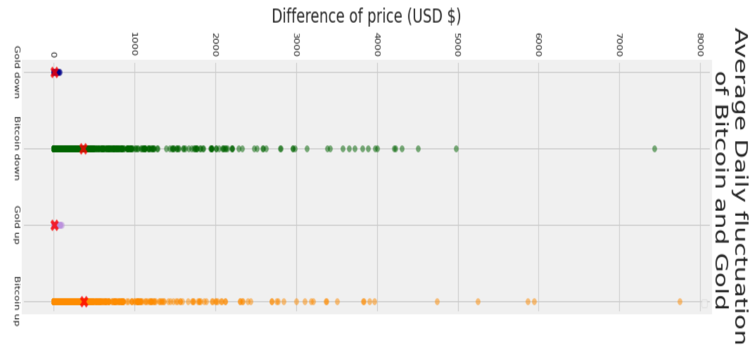


Figure 16 Average Daily fluctuation (K-Means)

After getting the expectation, we can calculate the trading curbs by formula (19) which are 9.301254180602015 for $gold\_up$, 9.154984423676016 for $gold\_down$, 375.45288125270105 for $bitcoin\_up$, 362.4866431582406 for $bitcoin\_down$.

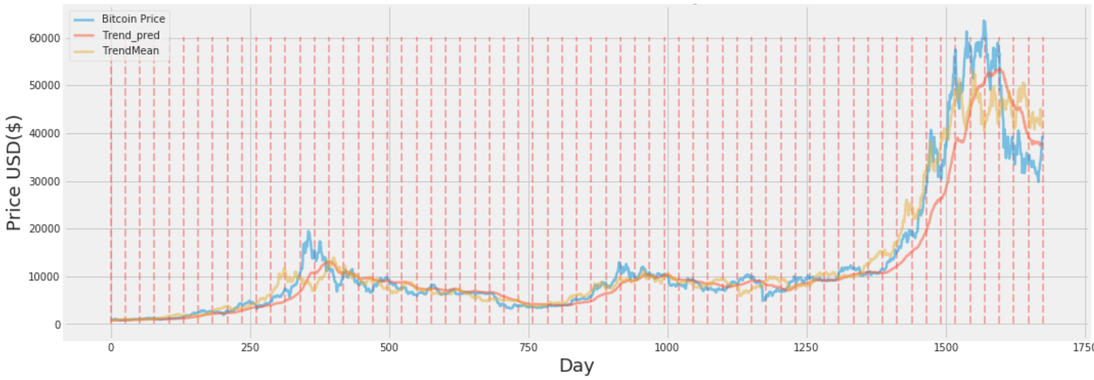
6.4 Trading model

Till now, we constructed all the metrics we need, so it is time to build up our final trading model. Referring to the idea of **The Turtle Trading Strategy**, we take the trend and risk into account, deciding when to buy and how much to invest. Also, for convenience, we build a **Simulating Trading System** to help tune the hyperparameters.

Taking the bitcoin price history as an example. The blue line is the actual price, and the red line is our trend line. We first divide the whole period into I intervals. Each interval allows trading only once. When the price is below the red line, it's time to invest; otherwise, we can consider if it can be sold. Since bitcoin price fluctuates a lot and its commission is pretty high, we don't want to trade frequently. So when the bitcoin's price is over the tendency line, we keep observing, if it keep over the tendency line TS days (TS is a hyperparameter), which means it is seriously overpriced and it time to sell it – after our tuning, here is the result:

$$\begin{cases} \boldsymbol{TS_{bitcoin,sell} = 120} \\ \boldsymbol{TS_{bitcoin,buy} = 1} \\ \boldsymbol{TS_{gold,sell} = 7} \\ \boldsymbol{TS_{gold,buy} = 90} \end{cases} \tag{20}$$

Figure 17 Gold historical information



After knowing when to trade, we use **TOPSIS Risk Analysis** to estimate how much we should trade.

$$\begin{cases} Trade_{buy} = lowest_{amount} \times topsis \times \omega + \dfrac{lowest_{amount} \times remainPeriod}{futurePeriod} \\ lowest_{amount} = \dfrac{Cash_{init}}{I} \end{cases} \quad (21)$$

Among these notations, topsis is the TOPSIS risk analysis result score, $Cash_{init}$ is the initial cash we have (in this case, it's \$1000), ω is a hyperparameter (after tuning, we got $\omega_{bitcoin} = 337$, $\omega_{gold} = 90$).

$$\begin{cases} Trade_{sell} = \begin{cases} currentValue - ExpectValue, & currentValue - ExpectValue > 0 \\ 0, & currentValue - ExpectValue \le 0 \end{cases} \\ currentValue = currentPrice \times numHold \\ ExpectValue = ExpectRatio \times Wealth \end{cases} \quad (22)$$

In the formula above, currentPrice is the current price of the bitcoin or gold, numHold is the unit of goods (bitcoin/gold) we hold, Wealth is our all wealth, ExpectRatio is the proportion of the total specific goods' value we hold over our all wealth.

Follow the rules mentioned above, we finally got **\$115802.155** in **5 years**, and got a **ROI** (Return On Investment) of around 83.231%. Here we can easily get the **annually average ROI** is 20.81%. (The ROI tracking is shown as follows)

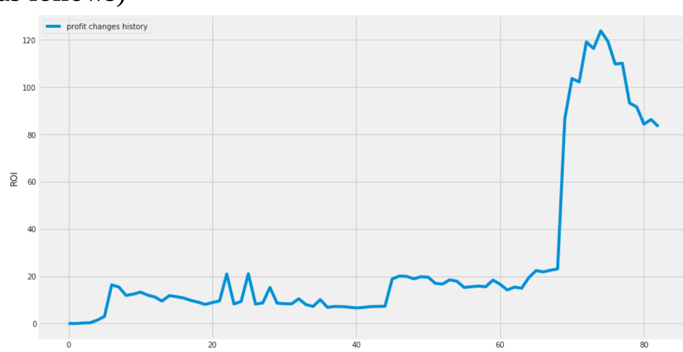


Figure 18 ROI tracking

7. Model Evaluation

7.1 Model sensitivity analysis

7.1.1 Problem analysis and solution

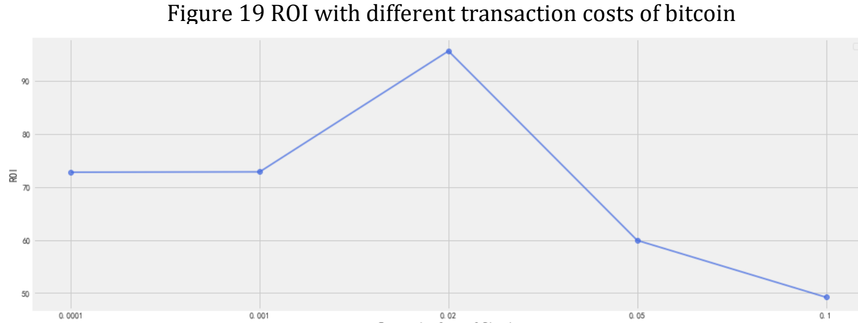
The question we face is how sensitive the strategy is to transaction costs. So we need an indicator to measure the sensitivity. In the field of numerical analysis, sensitivity analysis is an important step to measure the stability and sensitivity of the system. Sensitivity analysis can be used to determine which parameters have a greater impact on the system or model. Thus, the question can be transformed into how to measure the sensitivity of the transaction model to transaction costs. In other words, how much of an impact do transaction costs have on the model. ROI is the factor we need to maximize, and that is the ultimate goal and has great authority as a judge of whether the trading model is good or bad. The transaction cost affects various parameters and hyperparameters of the transaction model and ultimately affects the return on investment. Therefore, the problem can be transformed into the sensitivity of ROI to transaction costs.

The condition number is the norm of the matrix times the norm of the inverse of the matrix. From linear algebra analysis, we know that the condition number of a matrix is always greater than 1. The condition number of an orthogonal matrix is equal to 1, and the condition number of a singular matrix is infinite. According to the above definition, the condition number is suitable to measure the sensitivity of the matrix output to the input error. The larger the condition number is, the worse the sensitivity is.

7.1.2 Model establishment

First, we choose some typical transaction costs as the relatively small changes from the subjects of 0.02 (bitcoin transaction costs) and 0.01 (gold transaction costs). The following figure shows the change in return for the different costs of gold when the bitcoin costs are fixed and the change of investment with different Bitcoin costs when the gold costs is fixed:

Figure 19 ROI with different transaction costs of bitcoin



Through observation, it can be clearly confirmed that there is a non-linear relationship between the level of transaction costs and the final return on investment, so that the conditional number can represent the calculation of the non-linear mapping taking into account the sensitivity of the output (return) to the input (transaction costs).

As for the sensitivity, according to the definition of the condition number, A is understood as our entire trading system. The formula for calculating the condition number is as follows. We can take the first normal form, the second normal form or the infinite normal form.

$$\boldsymbol{\kappa}(\boldsymbol{A}) = \|\boldsymbol{A}\| \cdot \left\|\boldsymbol{A}^{-1}\right\| \tag{23}$$

The entire model system is A, but we cannot accurately capture the specific parameter values, so we do not use the definition method to solve it. Set x as ROI， b as transaction costs, then we can have $Ax = b$

Numerically, $\frac{\|\Delta b\|}{\|b\|}$ is as the relative backward error, $\frac{\|\Delta x\|}{\|x\|}$ is as the relative forward error. Δx indicates the difference between different costs， Δb indicates the gap in results (ROI). By the way, the ratio of the relative backward error to the relative forward error must be less than or equal to $\kappa(A)$, which means that we can use the ratio of the relative gap of x (transaction costs) to the relative gap of b (return rate) to measure the sensitivity of the linear system.

$$\frac{\|\Delta \boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leq \boldsymbol{\kappa}(\boldsymbol{A}) \frac{\|\Delta \boldsymbol{b}\|}{\|\boldsymbol{b}\|} \Rightarrow \boldsymbol{\kappa}(\boldsymbol{A}) \geq \frac{\|\Delta \boldsymbol{x}\|}{\|\boldsymbol{x}\|} : \frac{\|\Delta \boldsymbol{b}\|}{\|\mathbf{b}\|} \tag{24}$$

Using the inequality above, we can obtain a lower bound on the condition number to assess the sensitivity of the model. The closer the condition number is to 1, the less singular the matrix is, and the larger the condition number is, the more singular it is. That is, if the condition number of A is large, a small change in b can cause a large change in the solution x and numerical stability is poor. When the condition number of A is small, there is a small change of b, and the change of x is also small, and the numerical stability is good. It can also represent the change of x when b is unchanged and A has a small change. The final lower bound of κ(A) for gold transaction cost is: 3.196068217475405 and the final lower bound of κ(A) for bitcoin transaction cost is: 5.102926631314048.

Looking at this inequality again, we can draw some intuitive conclusions. First of all, the condition number of the matrix A uniquely determines how much the solution x of the linear equation is affected by the noise of the observed value b. The larger the condition number, the more strongly x is affected by the noise, which means that the rate of change of x differs from the rate of change of b. For gold transaction costs, the condition number is 3.196068217475405 and the average rate of change of observations is 0.8599290976655869, which means that the solution b(ROI) will change by at least 26.905843028120%. For bitcoin transaction costs, it causes the solution b(ROI) to change by at least 6.544377429801619%. The formula for the minimum rate of change is as follows: $\mathcal{R} = \frac{\vartheta}{\Omega \cdot \kappa(\mathrm{A})} \times 100\%$

7.1.3 Reasonable results

From the results, it appears that the entire transaction system is more sensitive to the change in the transaction cost of gold than to the transaction cost of Bitcoin. One reason is that gold is mainly used as a store of value in the trading system algorithm, which requires more frequent transactions to fund potentially more profitable Bitcoin investments. The second reason is that Bitcoin does not require as many buys and sells, as simply buying and selling unstable data causes more losses.

8. Conclusions

8.1 Strengths

(1) Safer. We make all decisions based only on the data from history, rather than predicting the future, because the future is unknown and history is pretty safe for us. Making decisions based on the known information can guarantee the trader to make money.

(2) Convincing. Compared to the model based on predictions, the simulation result of our model is much more convincing. This is because models based on predictions use 5-year data to train and the same data to simulate, which certainly never occurs in the real world. However, our model is based on the historical data - which means our simulation result is convincing.

(3) Efficient. In the model, we use simple and efficient methods such as the similarity matrix, principal component analysis, K-Means clustering algorithm, and so on. These methods are not only effective, but also avoid the model being redundant.

(4) High Robustness. This model predicts trends based only on historical data and does not require a large amount of data for training, so the model has high adaptability to data.

8.2 Weaknesses and Model Extension

(1) Cash inflation was not considered, while the real world has inflation.

(2) Before using data to train the model, we can do pre-training using the data before the specified time period to effectively improve the generalization ability and accuracy of the model parameters.

(3) When more relevant data can be used, more relevant metrics can be used to achieve better results in measuring risk and return, such as Sharpe ratio, Jensen, Traynor ratio.

(4) If more characteristics can be provided, there is a possibility to make predictions about the future, which can further optimize the investment strategy.

9. Reference

[1] Baur, Dirk G., and Lai Hoang. "The Bitcoin Gold Correlation Puzzle." Journal of Behavioral and Experimental Finance, vol. 32, 2021, p. 100561., https://doi.org/10.1016/j.jbef.2021.100561.

[2] Conrad, Jennifer, and Gautam Kaul. "An Anatomy of Trading Strategies." Review of Financial Studies, vol. 11, no. 3, 1998, pp. 489–519., https://doi.org/10.1093/rfs/11.3.489.

[3] Farmer, J.Doyne, and Shareen Joshi. "The Price Dynamics of Common Trading Strategies." Journal of Economic Behavior & Organization, vol. 49, no. 2, 2002, pp. 149–171., https://doi.org/10.1016/s0167-2681(02)00065-3.

[4] Hochreiter, Sepp, and Jürgen Schmidhuber. "Long Short-Term Memory." Neural Computation, vol. 9, no. 8, 1997, pp. 1735–1780., https://doi.org/10.1162/neco.1997.9.8.1735.

[5] Liu, Fang, et al. "Exploring Cluster Stocks Based on Deep Learning for Stock Prediction." 2019 12th International Symposium on Computational Intelligence and Design (ISCID), 2019, https://doi.org/10.1109/iscid.2019.10107.

[6] Vezeris, Dimitrios, et al. "Take Profit and Stop Loss Trading Strategies Comparison in Combination with an MACD Trading System." Journal of Risk and Financial Management, vol. 11, no. 3, 2018, p. 56., https://doi.org/10.3390/jrfm11030056.

[7] Xu, Jingxian, and Junhao Zheng. "Improved Turtle Trading Strategy and Empirical Analysis." Proceedings of the 2018 2nd International Conference on Education Science and Economic Management (ICESEM 2018), 2018, https://doi.org/10.2991/icesem-18.2018.279.

[8] Zaremba, Wojciech et al. "Recurrent Neural Network Regularization." ArXiv abs/1409.2329 (2014): n. pag.

10. Memorandum to traders

TO：Trader
FROM：MCM Team
DATE：February 22, 2022
SUBJECT：A Low-Risk Strategy based on the Past

Model - Referring to the idea of **The Turtle Trading Strategy**, we take the trend and risk into account, deciding when to buy and how much to invest. We built Multilinear model, LSTM model and TOPSIS risk analysis model as the sub-models providing the key metrics. Also, for convenience, we build a **Simulating Trading System** helping us tune the hyperparameters and find the best strategy.

Strategy - First, divide the 5 years into a few intervals. Each interval allows only one trade. Then buy goods if their price is below the expected true value for several days; otherwise sell goods. The trade amount is calculated as follows ($Trade_{buy}$ and $Trade_{sell}$ is the amount we're going to trade):

$$\begin{cases} \mathrm{Trade_{buy}} = \mathrm{lowest_{amount}} \times \mathrm{topsis} \times \omega + \dfrac{\mathrm{lowest_{amount}} \times \mathrm{remainPeriod}}{\mathrm{futurePeriod}} \\ \mathrm{lowest_{amount}} = \dfrac{\mathrm{Cash_{init}}}{\mathrm{I}} \\ \omega_{\mathrm{bitcoin}} = 337, \qquad \omega_{\mathrm{gold}} = 90 \end{cases}$$

$$\begin{cases} \mathrm{Trade_{sell}} = \begin{cases} \mathrm{currentValue} - \mathrm{ExpectValue}, & \mathrm{currentValue} - \mathrm{ExpectValue} > 0 \\ 0, & \mathrm{currentValue} - \mathrm{ExpectValue} \le 0 \end{cases} \\ \mathrm{currentValue} = \mathrm{currentPrice} \times \mathrm{numHold} \\ \mathrm{ExpectValue} = \mathrm{ExpectRatio} \times \mathrm{Wealth} \end{cases}$$

But for each time, the trading amount can't over the value of trading curb:

$$\begin{cases} \mathrm{Gold_{buy}} = 194.5968 \\ \mathrm{Gold_{sell}} = 191.5366 \\ \mathrm{Bitcoin_{buy}} = 5776.7291 \\ \mathrm{Bitcoin_{sell}} = 5577.2302 \end{cases}$$

Result - Above rules, in 5 years we have earned about $115802,155 and a ROI (Return On Investment) of about 83,231%. Here we can easily find that the annual average ROI is above 20.81%. (The ROI tracking is shown as follows).