A quantitative venture capital strategy based on LASTM-DP

Hao Wang [a], Xi He [b], Peiru Li [c, *]

School of economics and management, BeiHang University, Beijing, China

[a] wanghao_sem@buaa.edu.cn, [b] hexi13954@gmail.com, [c] 1374535568@qq.com

* Corresponding Author

Abstract. With the development of quantitative trading technology such as deep learning, more and more investors participate in quantitative trading. LSTM (long short term memory network) belongs to one of the deep learning models, which has better performance in predicting time series data. Therefore, this paper takes the mixed trading strategy of gold and bitcoin from 2018 to 2021 as the research object, forecasts the price through LSTM-EMA algorithm, uses dynamic programming to solve the optimal investment strategy, finally constructs the factors to measure the investment risk by using data mining method, calculates the daily investment risk by using analytic hierarchy process, and combines it with the obtained strategy to effectively reduce the volatility of daily income, At the same time, the annualized yield of the final strategy was increased to 39.9%, leading the vast majority of fund returns in the same period.

Keywords: Quantitative Trading LSTM-EMA; Dynamic Programming; Analytic Hierarchy Process.

1. Introduction

Price prediction has always been a research hotspot in the financial field, and its change trend is even compared with the "weather table" in the financial field. At present, the existing stock forecasting methods mainly include basic analysis method, technical analysis method, combination analysis method, time series method, neural network method and so on [1]. In the research of price data with highly nonlinear characteristics, the deep learning method suitable for extracting nonlinear characteristics gradually occupies the mainstream [2]. The deep neural network has more layers and more complex structure, which can transform the shallow information in the data into more abstract high feature information [3] LSTM is a cyclic neural network with control threshold structure for units. Hochreiter and Schmidhuber proposed LSTM neural network model in 1997 [4]. It can process long-time time sequential data and forget those unimportant feature information [5], which is just in line with the characteristics of price data prediction. Because LSTM has great advantages in temporal feature extraction, many studies regard it as a benchmark model [6], Through the improvement of LSTM algorithm, we can get better prediction accuracy. However, most studies end only after price prediction, and price prediction is a part of the specified trading strategy. For different financial assets, the fundamental factors affecting the price trend are also different. For example, the fundamental factors affecting stock index futures include macro-economy, interest rate level, money supply, etc [7]. Pierdzioch et al can predict the excess return of gold futures by using fundamental information such as inflation rate and term spread. [8] In real trading scenarios, more than one period of prediction is often required, and the more periods of prediction, the greater the error. Therefore, when formulating strategies, we should also consider the number of transactions (too many transactions mean more handling fees), investment proportion, investment risk, etc. Based on the above considerations, this paper realizes a feasible whole process transaction strategy model, selects bitcoin and gold as the representatives of high-frequency and medium and low-frequency transactions, uses LSTM-EMA to predict the price, and obtains the optimal strategy during the prediction period through dynamic programming [9]. Finally, six risk related indicators [10] are selected from the literature [11], the AHP algorithm is used to measure the daily transaction risk [12], and the investment proportion is set in combination with the required strategy, Based on this, we can get the specific trading strategy of each period. The annualized yield of this strategy in 2018-2021 is as high as 39.9%.

2. Generate policy model

Three models are established to complete the generation of the whole transaction strategy. Model 1: Price Prediction Model, which uses LSTM algorithm for time series prediction, and continuously improves the prediction accuracy, providing a premise for subsequent strategy formulation; Model 2: Trading Strategy Model, Based on the results of the prediction model, the dynamic programming algorithm is used to obtain the optimal portfolio on the prediction results, so as to generate the preliminary trading strategy. Model 3, risk assessment model. According to the price data, the risk characteristics are extracted, the comprehensive evaluation model is constructed by using AHP method, the weight of each characteristic is obtained and the risk score is calculated. Finally, the risk score is combined with the preliminary trading strategy to obtain the final investment strategy.

The flow chart of the whole strategy generation is as follows:

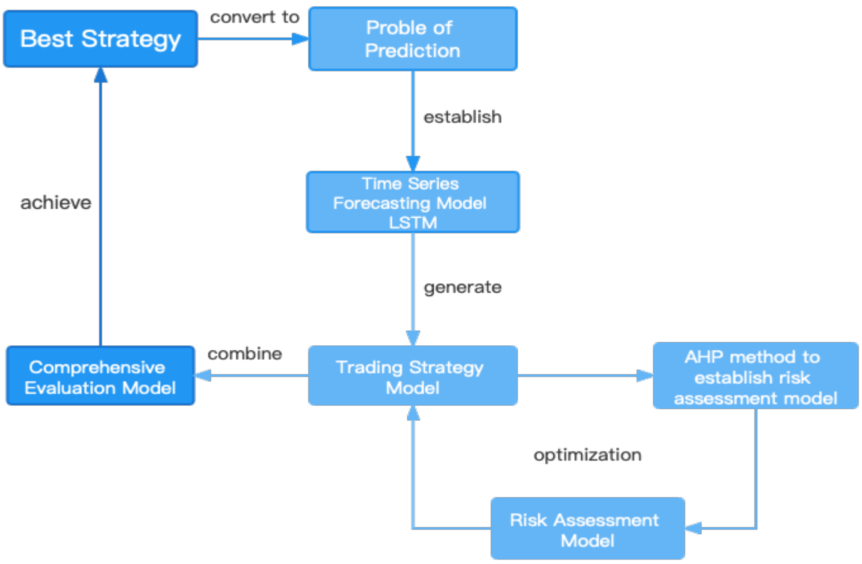

Figure 1. Strategic model flow chart

2.1 Price Prediction Model

In order to calculate the investment strategy, we obviously need to predict the future price of BTC and Gold based on the historical data, and then specify the future investment strategy based on our prediction results. The first thing we need to do is to predict backwards for a few days at a time. Obviously, the smaller the number of days of backward prediction, the more accurate it is. However, knowing the price of the next day alone can easily make investors miss the opportunity of long-term benefits due to uncertain prices, and a longer prediction window means inaccurate. After reading relevant materials, we finally set the prediction window to 5 days. The next step is how to predict. Time series analysis is a means to analyze and judge the future development trend through mathematical modeling based on the existing historical data, which is widely used in the prediction of price trends. A highly accurate time series analysis model for price analysis and prediction can bring great profits for investors. By referring to relevant literature, our prediction model adopts LSTM-EMA method, which is mainly divided into two steps. First, historical data is processed by EMA method, and then the processed data is trained by LSTM.

2.1.1 EMA principle

EMA is a trend index, which is constructed on the principle of arithmetic average of closing prices and connecting the average values at different times into a smooth curve to judge the future trend of security prices. The calculation formula for the exponential moving average $EMA_N(x_n)$ of the sequence $\{x_n\}$ whose period is N is defined as follows (up to item n):

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

From the definition, we can see the characteristic of weighted average EMA. In EMA indicators, the closer the time is to the time of the day, the more weight is assigned to the price of a security. This shows that EMA function gradually increases the weight ratio of recent securities prices, which can better reflect the recent fluctuations of securities prices in a timely manner. As for the choice of period N in the calculation formula, we can see from the broken line chart of the historical price data of BTC and Gold that the price fluctuation of BTC is much greater than that of Gold, so BTC needs a smaller N. Based on the comprehensive data, we set N_{btc} as 5 and N_{gold} as 15. The final data are obtained according to the above formula. It is important to note that the model assumes that the price of Gold on a non-trading day is the same as on the previous trading day.

2.1.2 LSTM model principle

The next step is to train the processed data using Long Short Term Memory (LSTM). LSTM is a kind of artificial neural network, derived from recurrent artificial neural network (RNN), which is usually used for prediction purposes because of its ability to calculate the dependence between various observation values in time series. And because of its inherent ability to quickly adapt to sharp changes in trends, LSTM can work well in fluctuating time series. Traditional time series analysis models, such as differential integrated moving average autoregression model (ARIMA), cannot adapt to sudden and drastic changes of security prices as soon as possible. In addition to the RNN structure, the LSTM also has a valve node layer. There are three types of valves: forget gate, input gate and output gate. These valves can be opened or closed to determine whether the output of the memory state of the model network (the state of the previous network) has reached a threshold to be added to the current calculation of the layer, as shown in the figure below for the individual cell structure of the LSTM:

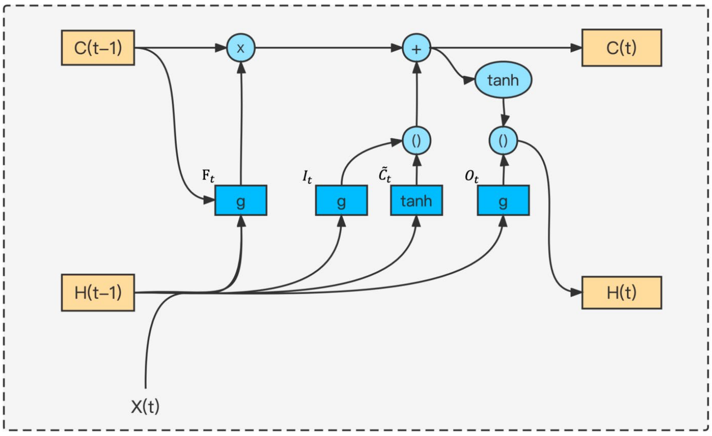

Figure 2. LSTM algorithm structure diagram

f_t, i_t, o_tt and C_t represent the corresponding three gate structures and cell states at time T respectively. W_f,W_i,W_o and W_c represent the weight matrix of the cyclic layer.σ represents the activation function, usually the Sigmoid or TANh function. Bf, BI, BO and BC represent bias vectors, and the calculation formula of LSTM model is as follows:

$$\begin{cases} f_t = g(W_f \cdot [h_{t-1}, x_t] + b_f) \\ i_t = g(W_i \cdot [h_{t-1}, x_t] + b_i) \\ \tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \\ C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \\ o_t = g(W_o \cdot [h_{t-1}, x_t] + b_o) \\ h_t = o_t \cdot \tanh(C_t) \end{cases} \tag{2}$$

2.1.3 LSTM model training

We choose Mean Squared Error (MSE) as the loss function to measure the training Error. Data from EMA formula from 2016/9/12 to 2018/9/11 are used as the training set to predict the price of BTC and Gold respectively, This also means that our investment strategy will start from September 12, 2018.

First, the price of BTC is predicted. In order to better adapt the model to the data, we continuously reduce validation loss by adjusting the number of hidden layer neurons, hidden layer number, Epochs, batch size and other parameters, and obtain the final model as follows:

In a 4-layer network, the first layer is Lstm layer (320 neurons), the second layer is Lstm layer (128 neurons), the third layer is Lstm layer (128 neurons), and the fourth layer is dropout layer (Dropout =0.1). Finally, a full connection layer with 5 neurons is used to predict the BTC price of the next 5 phases. The Epochs of the model was 140 times, and the Batch Size of the input data was 3. The network structure and obtained Loss curve is as follows:

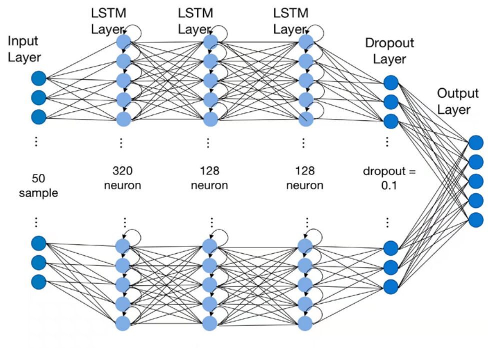

Figure 3. Network structure

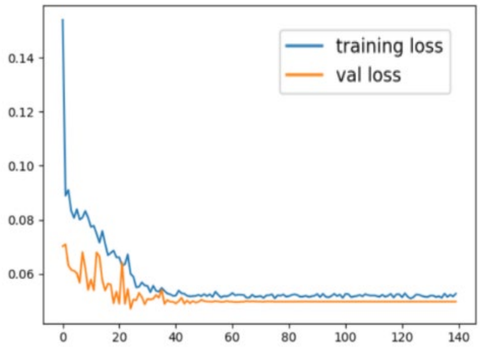

Figure 4. Loss of BTC

It can be seen that the value of Validation Loss finally stabilizes at about 0.03, which means that the prediction effect of the model is about 97%. The prediction curve is:

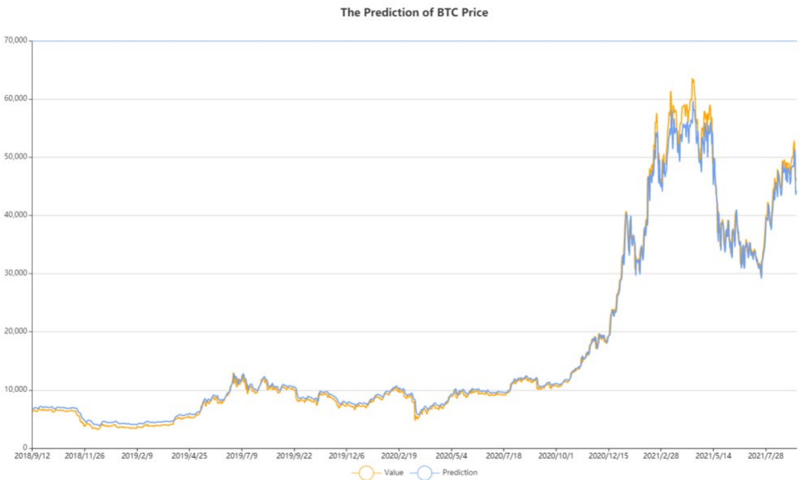

Figure 5. The prediction of BTC price

The same method as above is applied to predict Gold data, and the final model structure is as follows:

In a 4-layer network, the first layer is Lstm layer (128 neurons), the second layer is Lstm layer (64 neurons), the third layer is Lstm layer (64 neurons), and the fourth layer is dropout layer (Dropout =0.15).Finally, a full connection layer with 5 neurons is used to predict the Gold price for the next 5 periods. The number of Epochs used in the model is 100, and the Batch Size of input data is 3.

Figure 6. The prediction of glod price

From the complexity of the two LSTM models,We can clearly see that fitting the BTC price data requires more neurons and more training times(epoch).This shows that the price of BTC is more difficult to fit, which confirms our previous idea that the price of BTC is more volatile than that of Gold. The selection of period N in our EMA model is reasonable.

2.2 Strategy calculation model

We will optimization problem into the prediction of decision problem, and to establish forecasting model, using the dynamic programming, the maximum benefit of decision, as a result of the prediction accuracy will reduce gradually, as the prediction time so we are considering cycle under the premise of accuracy and the dynamic programming algorithm, We choose the price of the five periods predicted as a computing unit, and use the dynamic programming algorithm to calculate the optimal solution of the five periods and iterate continuously to obtain the optimal strategy in the prediction model, which will be used as the strategy basis in the trading strategy model and wait for further algorithm optimization.

In the dynamic programming algorithm, we record the initial status of the five days, including three states (0: cash, 1: COINS, 2: gold) and the conversion of cash assets, composed of an array (state, total assets), the equation of state for the new status updates a day, according to the trading strategy of assets, and update the total assets, total assets changes, If the total asset becomes smaller, it will be returned. After traversing all cases, it will return the decision strategy of the optimal decision within five days, and save the optimal decision within five days. Through continuous iteration, all the strategies will be finally obtained.

Let the state variable s_k be the state of the assets in the previous period, the decision variable u_k be the post-decision state of the assets in the current period, the indicator function is the cash value of the total assets $v_k(s_k, u_k)$, and the basic equation is:

$$f_k(s_k) = \max[v_k(s_k, u_k) + f_{k+1}(s_{k+1})],\ k = 5,4,3,2,1 \quad (3)$$

The state transition equation is:

$$s_k = u_k \quad (4)$$

Take the 20 days from October 1, 2018 as an example to explain the implementation of the strategy. When the strategy stays at 1, it means to hold bitcoin all the time; when the status changes from 1 to 2, it means to sell bitcoin and buy gold; when the status changes from 2 to 0, it means to sell gold.

Table1. Statement of strategic income

Date	Strategy	Total return
2018/10/6	1	1028.170821
2018/10/7	1	1026.941877
2018/10/8	1	1033.963681
2018/10/9	1	1034.454838
2018/10/10	1	1035.257557
2018/10/11	1	976.1723087
2018/10/12	2	948.8852546
2018/10/13	2	948.8852546
2018/10/14	2	948.8852546
2018/10/15	2	956.8201836
2018/10/16	2	957.4036342
2018/10/17	0	946.5588424

The trend chart of overall asset change results is as follows:

Figure 7. Asset change chart

2.3 Model of risk assessment

In addition to the time prediction model, we also establish a risk assessment model to evaluate the credibility of the time prediction and influence the choice of strategy.According to the data given by the topic, we extracted features to score, and adopted strategies according to the score.

2.3.1 Selection of influencing factors

The risk degree of whether the time prediction is correct or not is comprehensively influenced by the previous predictive risk and systemic risk, and each factor contains numerous determinants, which together constitute the characteristic factors of the risk assessment model. Through consulting materials and combining the given data, we selected the following 6 factors as the main influencing factors of the risk assessment model.They are Prediction fitting effect , Trend prediction accuracy , Blas , Variance , Expected risk rate and Absolute volatility.The following are their calculation formulas:

$$PF = \frac{\sum_{i=1}^{5}(X_{ip}-X_{ir})^2}{5} \tag{5}$$

X_{ip} represents the forecast value of phase i,X_{ir} represents the true value of phase i,PF response to the prediction and fitting effect of the previous stage.

$$TP = \frac{CP}{5} \tag{6}$$

CP indicates the correct times of trend prediction in the previous stage,TP correct rate of trend prediction in the previous stage.

$$MA = \sum_{i=1}^{5} X_{ir} \tag{7}$$

$$Blas = \frac{X_r - MA}{MA} \times 100\% \tag{8}$$

X_r represents the average value of the real price in the previous stage.Blas reflect the real price volatility of the previous stage.

$$VA = \frac{\sum_{i=1}^{5}(X_{ir} - \bar{X})^2}{5} \tag{9}$$

VA measures the price fluctuation of the previous stage from another angle.

$$ERR = \frac{2 \times (max((X_{ir}) - \min((X_{ir})}{max((X_{ir}) + \min((X_{ir})} \times 100\% \tag{10}$$

$Max((X_{ir})$ represents the maximum real price of the previous stage,$\min((X_{ir})$represents the minimum real price of the previous stage.ERR measured the maximum volatility.

$$AV = \frac{\sum_{i=1}^{5}|X_{ir} - X_{(i-1)r}|}{5} \tag{11}$$

AV represents the average value of absolute error in the previous stage.

On the basis of the above index system, the risk degree of the predicted value of each day (or trading day) is calculated according to the index data and the following construction model, as follows:

$$V_k = \sum_{i=1}^{5} X_i^* \times W_i \tag{12}$$

Where, V_k is the risk degree of the predicted value on day K (trading day); X_i^* is the initialization value of the i_{th} factor; W_i is the weight of the i_{th} factor. The determination of the weight of each factor can be determined by a variety of analysis methods, such as Delphi method and analytic hierarchy process.

2.3.2 Weight calculation

Our model uses analytic hierarchy process to determine the weight of each index. AHP is used to determine the weight of each factor in the Risk assessment model. Firstly, Risk indicators are taken as the overall target, namely layer A, and Predictive Risk and Systemic Risk are taken as layer B. Prediction fitting Effect, Trend Prediction accuracy, Blas and Market risk are used as layer C. Variance, Expected Risk rate, and Absolute Volatility are regarded as Layer D.

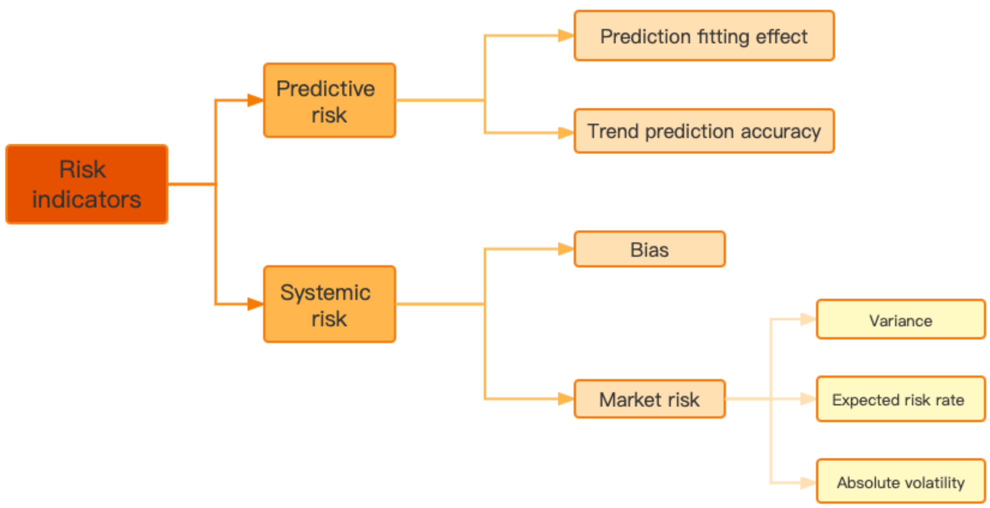

Figure 8. AHP structure diagram of the risk assessment model

(1) For target layer A, the accuracy of historical predicted value is more convincing, so the importance of historical predictive risk is greater than the systemic risk of the system itself, so the construction matrix is:

$$\mathbf{A} = \begin{bmatrix} 1 & 3 \\ 1/3 & 1 \end{bmatrix}$$

(2) For Predictive risk, fitting effect of Prediction depends on the difference between the predicted value and the real value, and Trend Prediction accuracy depends on the Trend accuracy rate of Prediction. It is concluded that the size of prediction gap is equally important to the direction of prediction trend, so the construction matrix is:

$$\mathbf{B1} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

(3) For a Systemic Risk, Blas represents the deviation between the price of the current day and the average price of the system, and Market Risk represents the volatility of the price before the current day. As Blas is linked to the price of the current day, it is more important. Market Risk only represents the volatility of the previous price without any direct impact, so the construction matrix is:

$$\mathbf{B2} = \begin{bmatrix} 1 & 3 \\ 1/3 & 1 \end{bmatrix}$$

(4) For Market risk, the matrix is constructed by reading literature and in-depth analysis and study of data:

$$\mathbf{C} = \begin{bmatrix} 1 & 3 & 2 \\ 1/3 & 1 & 1/2 \\ 1/2 & 2 & 1 \end{bmatrix}$$

After calculation, all discriminant matrices passed the consistency test.According to the weight of the influence factors of each layer, the total weight of each factor can be obtained, as shown in the following table.

Table 2. Total weight of index

The primary indicators	Weight	The secondary indicators	Weight	The tertiary indicators	Weight	Total weight
Predictive risk	0.75	Prediction fitting effect	0.5			0.375
		Trend prediction accuracy	0.5			0.375
Systemic risk	0.25	Blas	0.75			0.1875
		Market risk	0.25	Variance	0.54	0.03375
				Expected risk rate	0.16	0.01
				Absolute volatility	0.3	0.01875

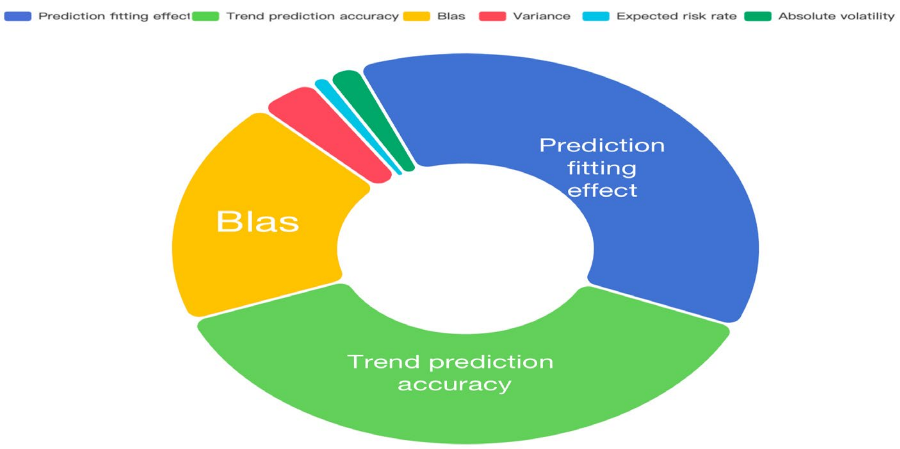

Figure 9. Weight fan chart

2.3.3 Venture capital strategy

Due to the deviation of the prediction results of the previously established prediction model, there are limitations in the results of strategy formulation only through the prediction model. Therefore, we establish a risk assessment model, carry out risk assessment, combine the two models, optimize the strategy obtained from the prediction model, and get the final strategy.

The figure below shows the impact of risks on returns. The part of the return curve in May 2021 is selected. The red area in the figure has a higher risk assessment, indicating that investment based on the predicted results is more risky. This diagram visualizes the impact of risk and therefore demonstrates the need to include risk assessment in the prediction model.

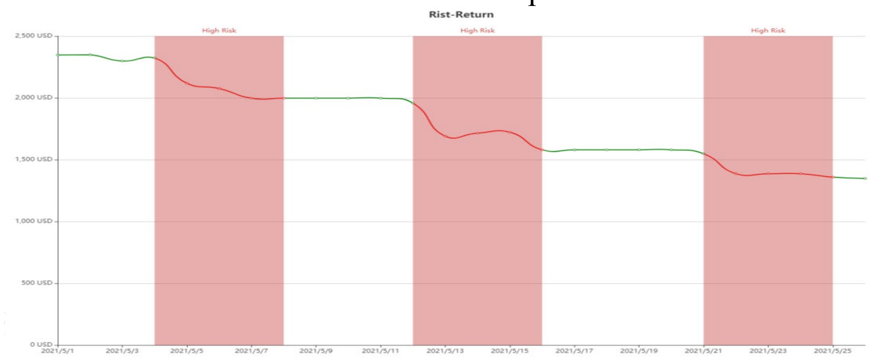

Figure 10. Risk warning chart

By reading literature and combining the scoring results of risk assessment model, we divide risk score into five levels: [0, 0.1), [0.1, 0.3), [0.3, 0.5), [0.5, 0.7), and [0.7, 1]. These five levels correspond to different transaction ratios: 100%, 70%, 50%, 30%, 0%. In the trading strategy combined with the scoring results of the risk assessment model, the specific number of transactions at each time is the product of the number of transactions in the original strategy and the proportion of transactions; when the risk score is less than 0.1, the risk is considered extremely low and the original strategy is traded; when the risk score is greater than 0.7, the risk is considered extremely high and the transaction is not carried out.

The figure below clearly and clearly shows the total return curve with and without risk (yellow with risk, blue without risk).

Figure 11. Total revenue curve

In the figure, we can clearly see that after considering the risk, the fluctuation of the model is much smaller than before, and better returns are obtained most of the time.

3. Summary

Through the establishment of the above prediction model, strategy model and risk evaluation model, we have completed the formulation of the whole venture capital strategy only by using the past price data of bitcoin and gold, and maintained a higher than average rate of return in most of the back test. Our research is mainly based on the prediction model and developed to the formulation of the whole investment strategy, which makes the research not only stay in the single task of predicting the price, but use the predicted price to further determine the investment portfolio. Comprehensive price fluctuations and prediction accuracy measure investment risks, dynamically generate investment strategies, and reasonably allocate assets, so that the final annualized rate of return has reached an amazing 39.9%. The main contribution of this study is to propose a feasible investment strategy generation method using only historical price data, which has a very good performance in the portfolio of high-frequency and low-frequency.

References

[1] Hu Juzhou, LAN Qiujun. Comparison between data mining technology of financial time series and classical statistical model [J]. Systems engineering, 2005, 23 (6): 95-98.

[2] Fang Yiqiu, Lu Zhuang, Ge Junwei Stock price prediction of joint RMSE loss lstm-cnn model [J] Computer engineering and application: 1-10.

[3] Wang Ting, Xia yangyuxin, Chen Tieming, stock short-term trend prediction based on multi category feature system [J] Computer science, 2020,47 (S2): 491-495.

[4] Hochreiter S and Schmidhuber J. Long short-term memory. [J]. Neural computation, 1997, 9(8): 1735-80.

[5] Wu Dashuo, Zhang chuanlei, Chen Jia, Xiang Qihuai Prediction and analysis of stock index based on improved LSTM neural network based on genetic algorithm [J] Computer application research, 2020,37 (S1): 86-87 + 107.

[6] Kumar K, Haider M T U. Enhanced Prediction of Intra-day Stock Market Using Metaheuristic Optimization on RNN-LSTM Network [J]. New Generation Computing, 2021, 39(1): 231-272.

[7] Murphy J. Technical analysis of the futures market [M]. London: Prentice Hall Press,1986.

[8] Christian Pierdzioch and Marian Risse and Sebastian Rohloff. On the efficiency of the gold market: Results of a real-time forecasting approach [J]. International Review of Financial Analysis, 2014, 32: 95-108.

[9] Zhao Xiaoli Research on optimal pension investment strategy based on dynamic method [D] China University of metrology, 2018.

[10] Apoorv Gupta and Prabina Rajib. Do VaR exceptions have seasonality? An empirical study on Indian commodity spot prices [J]. IIMB Management Review, 2018, 30(4): 369-384.

[11] Madhusudan Karmakar and Samit Paul. Intraday portfolio risk management using VaR and CVaR: A CGARCH-EVT-Copula approach [J]. International Journal of Forecasting, 2019, 35(2): 699-709.

[12] Shen Liangfeng, fan Xiangru Evaluation and decision-making of venture capital projects based on analytic hierarchy process [J] Infrastructure optimization, 2002, (04): 20-22.