

# Quantitative trading strategy and portfolio optimization analysis based on LSTM model

Yueyao Li<sup>1</sup>, Ziyao Zhu<sup>1,\*</sup>, Ziyu Liu<sup>2</sup>, Zixuan Zhou<sup>3</sup>, Yan Guo<sup>4</sup>

<sup>1</sup> School of Insurance and Economics, University of International Business and Economics, Beijing, China, 100029

<sup>2</sup> Business School, University of International Business and Economics, Beijing, China, 100029

<sup>3</sup> School of Cultural Industries Management, Communication University of China, Beijing China, 100024

<sup>4</sup> School of Banking and Finance, University of International Business and Economics, Beijing, China, 100029

\* Corresponding Author Email: zzyafr@163.com

**Abstract.** This paper focuses on the optimal risk portfolio problem. Based on the principle of return maximization, an LSTM neural network model is used to predict the price movements of gold and bitcoin; then an objective planning method is used to find the optimal trading strategy for each day. Finally, the final amount of the initial asset is calculated for the end date. To accommodate the huge price changes of Bitcoin, an LSTM neural network model was built for prediction in this paper, which produced well-fitted Bitcoin price data. After obtaining the predicted prices of gold and bitcoin, an objective programming model is built with daily returns as the objective function, constraints are set, and market days (gold can be traded) and non-market days (gold cannot be traded) are discussed to derive the optimal investment strategy; a linear programming solution method is used to derive the optimal solution under the constraints to ensure the maximum amount of daily returns. The LSTM neural network forecasting model used in this paper has a long memory function and good predictability. Meanwhile, the heuristic algorithm greatly accelerates the convergence of the model for the optimal solution.

**Keywords:** QTDM, LSTM model, Goal programming.

## 1. Introduction

To help investors maximize returns and properly mitigate risk in risky investments, appropriate mathematical models are required to advise investors on the best daily portfolio based on the previous day's gold or bitcoin price, so that they know whether they should buy, hold or sell gold or bitcoin[1]. Given two financial products, bitcoin and gold, market traders can choose the appropriate investment strategy to invest in. In order to achieve the goal of maximizing total return, we first need to be clear that an investment strategy is essentially a forecast of an uncertain future market[2]. Therefore, in order to build the quantitative trading decision model required by the market trader, the first step requires building a predictive model of the future market, and the second step can be based on the results derived from the predictive model to determine the best investment option[3].

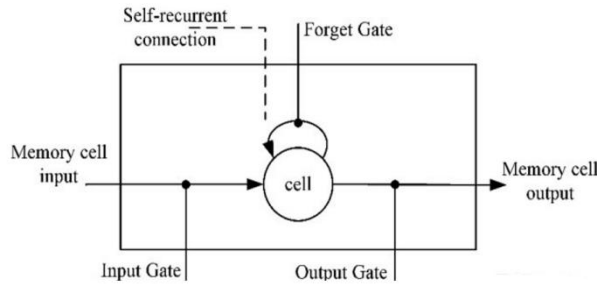
## 2. Model Construction

### 2.1. LSTM Model

The Long Short Term Memory (LSTM) model is an improvement of the Recurrent Neural Network (RNN) model, which is based on the ordinary multilayer BP neural network and adds a horizontal connection between the units in the hidden layer to transfer the value of the previous time series neuron to the current neuron, so that the neural network has a memory function[4]. The LSTM model adds memory units to each neuron in the hidden layer of the RNN model, and controls the

degree of memory and forgetting of previous and current information through controllable gates each time it is passed between units in the hidden layer, thus having a long-term memory function[5].

The LSTM model has obvious advantages over the ARIMA model for bitcoin investment products, where the price data changes dramatically and the investment returns are highly variable: on the one hand, the LSTM model can fully exploit the long-term dependency relationship between variables for bitcoin's time-series structure, and thus apply the time-series relationship to the forecasting process as much as possible[6]. LSTM model also solves the problems of gradient disappearance and gradient explosion of RNN model, so as to make more accurate prediction.



**Figure 1.** LSTM neural network concept diagram

Mathematically, the LSTM is a highly composite nonlinear parametric function that maps one column of vectors to another set of vectors through an implicit layer[8]. Figure 1 depicts the computational process of the judgment of the LSTM model, and at moment, the LSTM neural network is defined by the following equation.

$$f_t = \sigma_g(W_f x_j + U_f h_{t-1} + b_f) \quad (1)$$

$$i_t = \sigma_g(W_i x_j + U_i h_{t-1} + b_i) \quad (2)$$

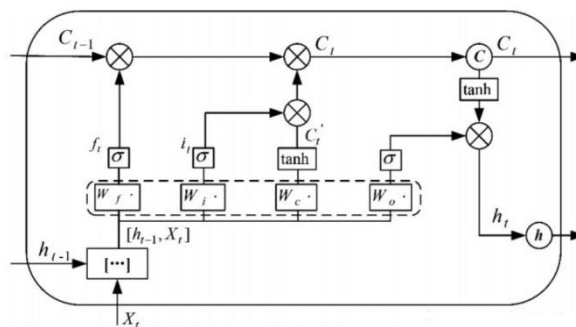
$$o_t = \sigma_g(W_o x_j + U_o h_{t-1} + b_o) \quad (3)$$

$$g_t = \sigma_h(W_g x_j + U_g h_{t-1} + b_g) \quad (4)$$

$$c_t = f_t * c_{t-1} + i_t * g_t \quad (5)$$

$$h_t = o_t * \sigma_t(c_t) \quad (6)$$

The internal structure of the LSTM neural network is shown in the Figure 2. In the training process of LSTM neural network, firstly, the data features at moment t are input to the input layer and the result is output through the excitation function. The output result, the output of the hidden layer at moment t-1 and the information stored in the cell unit at moment t-1 are input into the nodes of the LSTM structure, and the data are output to the next hidden layer or output layer through the processing of Input Gate, Output Gate, Forget Gate and cell unit, and the results of the nodes of the LSTM structure are output to the neurons of the output layer, and the calculation of back propagation error, and update each weight value.



**Figure 2.** Internal structure of LSTM neural network

The Final output value  $y_t$  is a nonlinear function on  $h_t$   $y_t = \sigma_t(W_t h_t + b_t)$ .

In addition to the above mentioned variables  $\sigma_g, \sigma_h$  represents the nonlinear activation function. sigmoid function and tanh function are the two most commonly used activation functions in neural networks, and multiple combinations of these activation functions can approximate complex nonlinear relationships between the input and output vectors.

**2.2. Constraint model**

Based on the above prediction model we can derive the price of a future trading day. However, the price can only determine the direction of a future trading day, but cannot reflect the specific trading strategy. So we need to establish a constraint model and conduct a goal planning. Objective programming is a special application of linear programming, which can be solved by general linear programming, or by the alternative solution method[9].

Assume that the initial portfolio is  $[C, G, B] = [0,500,500]$ . Because Bitcoin is a virtual currency and can be traded daily, while gold is an international currency and is only traded when the market is open. So we discuss it in two cases as follows.

- (1) When both bitcoin and gold can be traded
  - a. Determine the objective function

$$maxZ = \frac{n_1}{m_1}(z_1 + x_1) + \frac{n_2}{m_2}(z_2 + x_2) \tag{7}$$

- b. Determination of constraints

Constraint 1: The total amount of gold and bitcoin transactions cannot exceed the total number of holdings.

$$-z_1 - z_2 < x_1 + x_2 < z_1 + z_2 \tag{8}$$

Constraint 2: Neither gold nor bitcoin can be bought or sold in excess of their respective total values.

$$-z_1 < x_1 < z_1 \tag{9}$$

$$-z_2 < x_2 < z_2 \tag{10}$$

Constraint 3: The commission cost of each transaction is known to be  $\alpha\%$  of the transaction amount. When  $\alpha_{gold}=1\%, \alpha_{bitcoin} = 2\%$  the tomorrow's return of gold and bitcoin must be greater than the respective commission cost.

$$0.01x_1 \leq \frac{n_1}{m_1} - 1 \tag{11}$$

$$0.02x_2 \leq \frac{n_2}{m_2} - 1 \tag{12}$$

In summary, the following target planning model was developed.

$$maxZ = \frac{n_1}{m_1}(z_1 + x_1) + \frac{n_2}{m_2}(z_2 + x_2) \tag{13}$$

$$s. t. \left\{ \begin{array}{l} -z_1 - z_2 < x_1 + x_2 < z_1 + z_2 \\ -z_1 < x_1 < z_1 \\ -z_2 < x_2 < z_2 \\ 0.01x_1 \leq \frac{n_1}{m_1} - 1 \\ 0.02x_2 \leq \frac{n_2}{m_2} - 1 \end{array} \right. \tag{14}$$

- (2) When bitcoin is tradable and gold is not tradable  
a. Determination of the objective function

$$\max Z = z_1 + \frac{n_2}{m_2} (z_2 + x_2) \tag{15}$$

- b. Determination of constraints

Constraint 1: Bitcoin cannot be traded for more than its holdings.

$$-z_2 < x_2 < z_2 \tag{16}$$

Constraint 2: Bitcoin's tomorrow's yield must be greater than its commission cost.

$$0.02x_2 \leq \frac{n_2}{m_2} - 1 \tag{17}$$

In summary, the following target planning model was developed.

$$\max Z = z_1 + \frac{n_2}{m_2} (z_2 + x_2) \tag{18}$$

$$s. t. \begin{cases} -z_2 < x_2 < z_2 \\ 0.02x_2 \leq \frac{n_2}{m_2} - 1 \end{cases} \tag{19}$$

### 3. Model solving and analysis

#### 3.1. Solution and Result

We use Python and Keras to implement the LSTM model prediction for time series. The code structure for solving the LSTM model is shown in the following Figure 3.

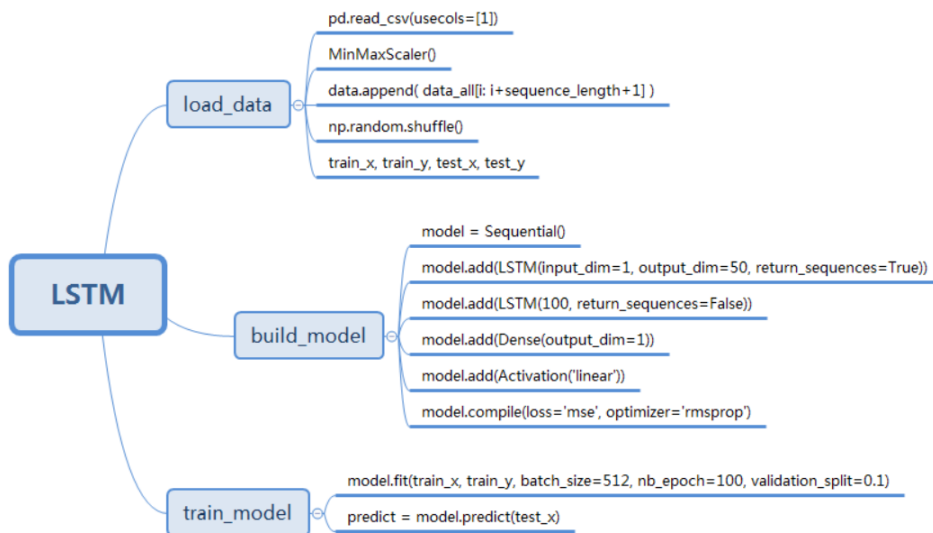


Figure 3. LSTM code structure

(1) Data preparation: introduce sklearn, keras, pylb and other libraries, import the original bitcoin univariate time series data into Python, merge the given data, complete the missing values, and sort them in chronological order.

- (2) Setting up the LSTM model

a. We use the LSTM model mainly to predict the future value based on the existing historical value, the input layer and output layer both have only one neuron, so the Sequential model is used to

construct the LSTM model neural network. The number of neurons in each layer is 50; the number of iterations is taken as 400; the Adam algorithm is used to optimize and improve the learning rate.

b. Set the loss function as the mean square error function RMSE to ensure the consistency of the magnitude, and set ReLU as the activation function to avoid the existence of Vanishing Gradient Problem, so that the convergence rate of the model is maintained in a stable state and the model is activated.

$$ReLU : f(x) = \max(0, x) \tag{20}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_{obs,i} - x_{model,i})^2}{n}} \tag{21}$$

(3) Run LSTM model

a. Train the model, convert the data to 32-bit floating point numbers, apply the MinMaxScaler function for normalization, and input the model. All data are iterated 400 times, the weights are updated, and an optimization fit is performed.

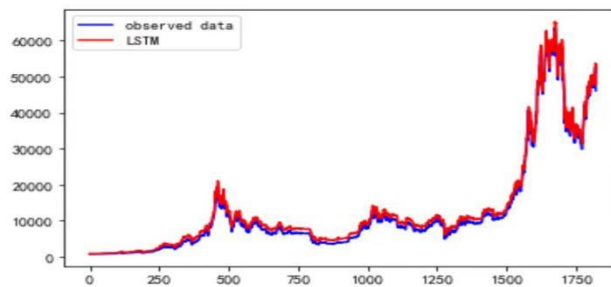
$$X^* = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{22}$$

b. Set the test model parameters: MSE, RMSE, MAE, R2, which are used to test the accuracy of the model prediction fit.

(4) Result output: Output the predicted data and test parameters, and plot the predicted results against the real data, with the horizontal axis being the adjusted time series and the vertical axis being the price.

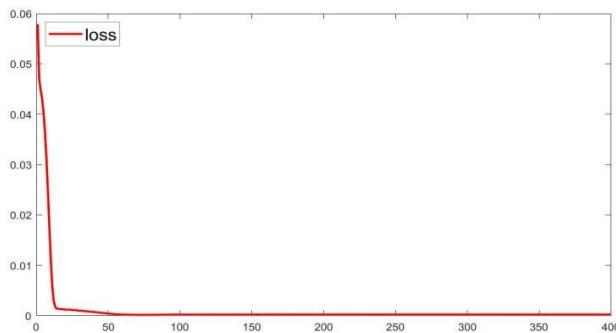
**3.2. Analysis of the Result**

Combined with the above model narrative for solving, we can get the prediction about the price level of Bitcoin, as shown in the Figure 4, the output prediction results and the actual price comparison image are as follows. By analyzing the image, we can see that the LSTM predicted values are basically in line with the actual values, and the fitting and prediction results are excellent.



**Figure 4.** Comparison between LSTM and observed data

The following Figure 5 shows the change of the loss function value in 400 iterations. During the iterative optimization fit to adjust the weights, the model can be fitted quickly, and the loss function value basically converges to 0 after the 50th iteration, and the fit optimization effect is good.



**Figure 5.** loss function value in 400 iterations

Measured by the MSE, RMSE, MAE, and R2 values output by the program of Table 1,  $R2 \approx 99.29\%$  indicates an excellent fit, but from the MSE, RMSE and MAE, there is a large deviation in the LSTM prediction effect. The problem of large error values in three can be explained in two ways: first, the modeling selected a data period covering the period from September 2020 to September 2021, when the value of bitcoin changed dramatically, with a maximum monthly growth rate of about 95%; second, in building the prediction model, the investment risk was not considered, and the model was built only for the price variable.

**Table 1.** Parameters

MSE	RMSE	MAE	R2
1195295.78	1093.2501	976.3825	0.9929

#### 4. Analysis of the Result and Sensitivity

Combining the construction of the target planning model we can combine the forecast data with the historical data to get the daily optimal solution, and run the model code through Python to get the first day's trading strategy. After executing the first day's strategy, the second day's strategy is obtained by combining the second day's holdings with the forecast results and running the program again, and so on. Here we add a round-robin decision model code to do the calculation to get each day's trading strategy[10].

Where, the current day's gold (bitcoin) trading amount is the day's trading strategy, combined with the size of the value has been adjusted to show the number of digits, where a positive number represents a buy, a negative value represents a sell, and a 0 value represents no operation. The current day's gold (bitcoin) holding amount is the value after trading, the price change of the day, and adjusting for commissions.

**Table 2.** Gold, Bitcoin Trading Strategies

Transaction date	Today's number of Gold holdings	Today's number of Bitcoin holdings	Gold trading volume today	Bitcoin trading volume today
2016-9-11	500.0000	502.3623	0.0000	12.2337
2016-9-12	499.5212	516.5271	-0.1125	13.2179
2016-9-13	497.9483	527.8481	-0.5636	12.9426
...	...	...	...	...
2021-9-8	0.0000648890	298872.6157	0.0000614223	0.0000649812
2021-9-9	0.0001194964	302112.2539	0.0000601393	1.4634
2021-9-10	0.0000008050	271132.9826	-0.0001184141	-13.9217

As shown in Table 2, assume that in the last trading day of the forecast period bitcoin, gold value no longer occur any fluctuations. Simultaneously at the end of the day we do not do any operations, After calculation \$ 1000 initial funds, after the above series of transactions, on September 10, 2021 significantly rise to the final value of \$ 271118.9216.

We also test the sensitivity of the previous model by varying the value of commission cost of Bitcoin or Gold, to simulate investors in five-year deal could meet the changes in the commission. Bring the case for  $\alpha_{\text{gold}}=0.5\%$  1% 1.5% 2% 2.5%,  $\alpha_{\text{bitcoin}}=2\%$  to models and bring the case for  $\alpha_{\text{gold}}=1\%$ ,  $\alpha_{\text{bitcoin}}=1.5\%$  2% 2.5% 3% 3.5% to models as well, models can calculate exactly.

We find that the change of commission has a great influence on the predicted maximum return of the corresponding financial product, but compared with the change of gold commission, the bitcoin commission will significantly change the final investment return, as Figure 6 shows.

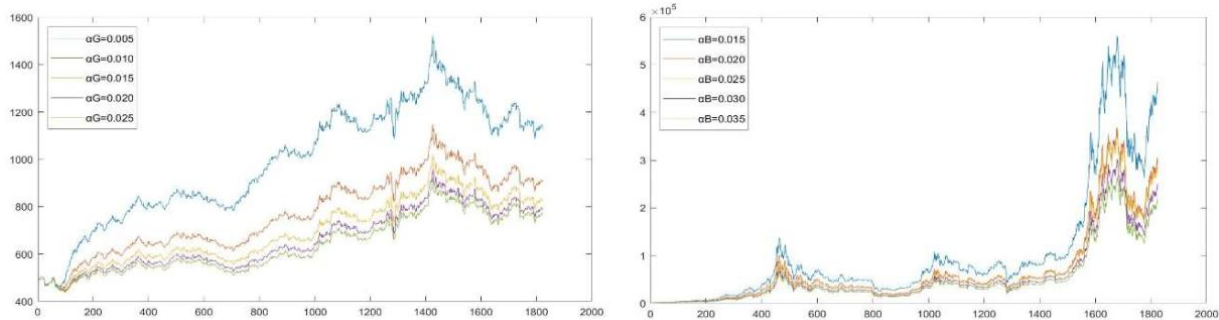


Figure 6. Impact of varying  $\alpha$ -value on the final investment return

## 5. Conclusions

Integrating the above three models from pushing to building to solving the computing process, we purposefully selected the best prediction model that best fits the target financial product, using all known information - the five-year gold (bitcoin) price from 2016.9-2021.9 for construction, and inputting the results into the target planning model to obtain the investment plan that maximizes on the value in the feasible interval.

However, such an investment scenario is not entirely realistic and ignores the investor's consideration of risk, assuming that the investor is "100% risk-averse". The model can be profitable due to the continued increase in the value of both investment products over the forecast horizon, but the consideration of risk is particularly important when considering a more general investment scenario.

## References

- [1] Kim, S., & Kang, M. (2019). Financial series prediction using Attention LSTM. arXiv preprint arXiv:1902.10877.
- [2] Giokas, D., & Vassiloglou, M. (1991). A goal programming model for bank assets and liabilities management. *European Journal of Operational Research*, 50(1), 48-60.
- [3] Aouni, B., Colapinto, C., & La Torre, D. (2014). Financial portfolio management through the goal programming model: Current state-of-the-art. *European Journal of Operational Research*, 234(2), 536-545.
- [4] Kim, M. K., & Zumwalt, J. K. (1979). An analysis of risk in bull and bear markets. *Journal of Financial and Quantitative analysis*, 14(5), 1015-1025.
- [5] Zhang N. Research on the application of stock price prediction based on LSTM neural network[J]. *Modern Business*,2021(16):116-118.
- [6] Meng Y,Xu QJ. Stock price prediction based on LSTM neural network-Marshall chain[J]. *Time Finance*,2021(11):3-6.
- [7] Jiao, Fengxia. Research on stock price prediction based on LSTM neural network[J]. *Digital Technology and Applications*, 2021, 39(03): 220-222.
- [8] Wang Na,He Yiyue,Zhang Shan. Modeling intra-day stock trading distribution prediction based on LSTM neural network[J]. *Wealth Life*,2020(18):9-11.

- [9] Wu Dashuo,Zhang Chuanlei,Chen Jia,Xiang Qihuai. Improved LSTM neural network stock index prediction analysis based on genetic algorithm[J]. Computer Application Research,2020,37(S1):86-87+107.
- [10]Chen Teng-Jin. Time series security price trend forecasting based on LSTM neural network-sampling with exponential moving average mean data[J]. China Business Journal,2021(20):92-94.