Market trading: LSTM-based forecasting and decision making

Ziqi Dai[1, †], Tianle Shen[1, †], Yu Hou[2, †], Xin Luo[1], Suping Zhang[3, *]

[1]School of Electronic and Information Engineering, Anhui Jianzhu University, Hefei, Anhui, 230601, China

[2]Faculty of Mechanical and Electrical Engineering, Anhui Jianzhu University, Hefei, Anhui, 230601, China

[3]School of Mathematics and Science, Anhui Jianzhu University, Hefei, Anhui, 230601, China

* Corresponding Author Email: 592745039@qq.com

[†]These authors contributed equally.

Abstract. Power In this paper, we address the decision problem of whether a trader should trade two assets, gold and bitcoin, daily under different circumstances, and use LSTM, neural network models, to obtain daily predicted price data for gold and bitcoin based on the knowledge of past market data. Accordingly, different planning strategies are applied to gold and bitcoin to obtain the total assets we can hold after five years. We provide evidence of the optimality of the strategy in several ways and specifically analyze the sensitivity of the strategy to transaction costs and the impact of transaction costs on the strategy and the results.

Keywords: LSTM neural network forecasting, Matlab software, sensitivity analysis, risk factor, integrated evaluation.

1. Introduction

Based on the rapid growth of the global financial market in the past 20 years and the intensified competition behavior in the market (programmed trading, algorithmic trading, and high-frequency trading brought about by the rapid development of computers), more institutional investors have joined this market, which has had a huge impact on the trading market in some US states. In the asset trading process, each transaction requires the payment of a corresponding commission, and the total return on investment is affected by various factors and is risky. If there is a large deviation in the capital trading forecast, it may cause a loss of the principal based on a negative return; there is a need to establish appropriate models and strategies to forecast the future asset situation at short and medium-term levels based on the past performance of assets, and accordingly, to decide the trading flow of assets.

In the early economic theory studies, most of the studies gathered on the questions of analyzing whether macro factors have an impact on the asset trading market and whether the asset trading market produces an effective response to changes in macro factors. During this period, asset trading markets were viewed as an important component of macroeconomics and included in the overall study of macroeconomics.

Engle (1982) used the ARCH model [1] and its extensions in his study to describe the intrinsic transmission and distribution lags of price fluctuations in asset trading markets; Frech, Schwert, and Stambaugh (1987) used GARCH [2] to analyze the relationship between expected returns and volatility in the stock market based on U.S. asset trading market data and showed that expected returns are positively correlated with the predictable volatility of stock returns. scholars such as Kuttner (2001), Bernanke (2005), and Leiter (2009) analyzed the impact of interest rate changes on asset trading market volatility based on data from studies in the United States, Japan, and European countries, and found that interest rates can significantly affect the stock and bond markets [3]. Bohl, Mayes, and Siklos (2011) explore the impact of interest rate changes on stock market returns based on EU and UK data, and their findings show that expected interest rate shocks hurt asset trading market returns. Hyeuk and Sang (2016) used the random forest method for the analysis and prediction of stock market movements, and the results of the study showed that the method is more effective compared to existing methods [4]. Machine language shows better effectiveness in stock market simulation and forecasting processes [5-8], which inspires us to use machine language for model design, data analysis, and backtesting [9-10].

To meet the profitability of asset trading, we need to develop a model that predicts asset prices at a later time from known market transactions at a certain time. Considering that most market traders use Gold and Bitcoin as liquid assets, we pay particular attention to how these two trade over long periods.

What we need to do is as follows:

(1) Develop a model that gives the best daily trading strategy based only on price data up to that day. And use your model and strategy to derive the value of the investment on September 9, 2021, at an initial cost of $1,000.

(2) Provide appropriate data to prove that the model provides the best strategy for market traders.

(3) Determine the sensitivity of the strategy to transaction costs, and discuss the impact of transaction costs on strategy and outcomes.

(4) How these findings and models can be generalized to other tasks will be discussed.

We use figure 1 to show our work:

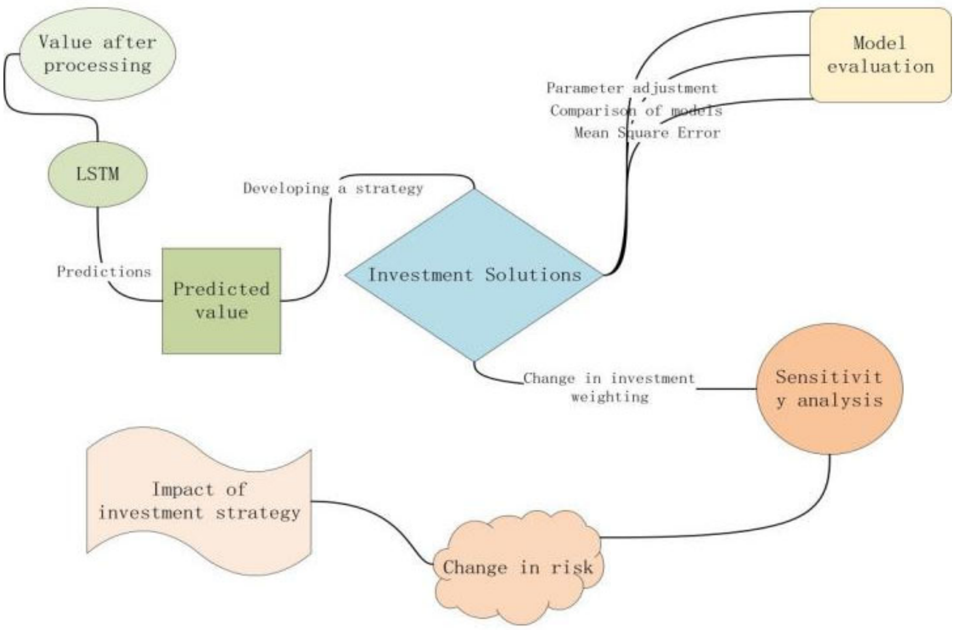

Figure 1. Workflow diagram

2. Model building and solving

To simplify our model and eliminate complexity, we have made the following main assumptions in this paper. All assumptions will be re-emphasized once they are used in our model construction.

(1) Before each trade, the commission payable is deducted from the principal.

(2) During the gold trading process, the daily unit price that is known to be missing from the data set directly defaults to the previous day's price.

(3) Market traders pay more attention to the rate of return when buying, holding, or selling assets in their portfolios than to the risk of investment.

(4) The impact of natural disasters and national policies on the market is not considered during the five years of capital flows.

We use figure 2 to show the idea of this question.

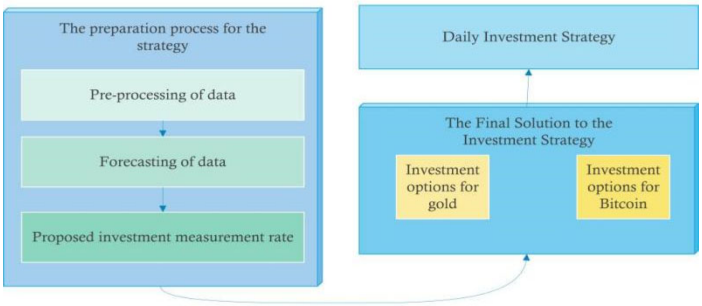

Figure 2. Ideas for Question 1

2.1 Introduction to the model

2.1.1 LSTM neural network

The given data shows the daily prices of both gold and bitcoin over five years, but we can only know the prices of both assets at the beginning of a given day, i.e. there is this limitation of only being able to use the past daily prices to date to determine whether we should buy, hold or sell the assets in our portfolio each day. To ensure a return on our investments, we need to use known data to make predictions of the daily prices of gold and bitcoin.

We will build a "sequence-to-sequence" regression LSTM neural network, in which the LSTM network will learn to predict the cell state value at the previous moment of the sequence after inputting the cell state value at that moment, as shown in Figure 3; to predict the cell state value at multiple moments in the future, we use the predict And Update State function to make predictions and update the network state at each prediction. And the daily prices of gold and bitcoin in the past so far are used as the dataset, aiming to predict the future prices of both assets based on the known prices of both assets.

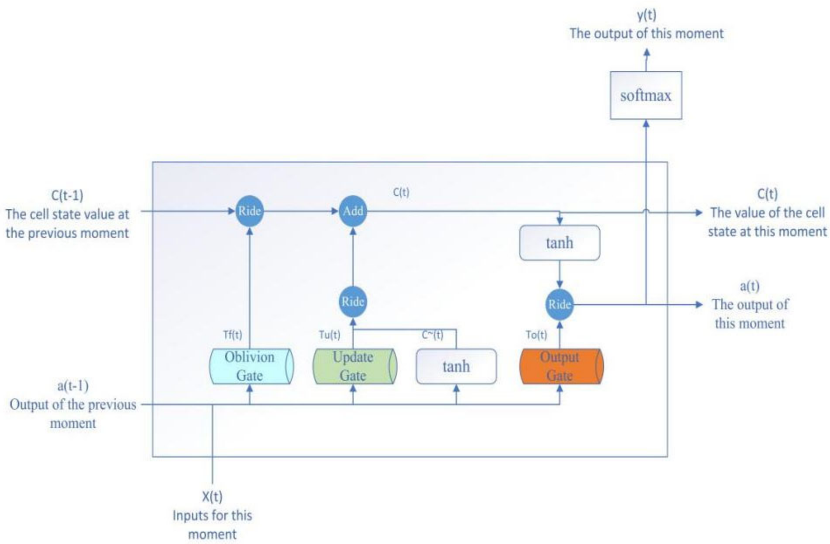

Figure 3. Principle of LSTM neural network

The 1 neuron of the LSTM neural network contains 1 cell state and 3 gate mechanisms. The cell state is the memory space of the LSTM model. The cell state changes over time and the recorded information is determined and updated by the gate mechanism. the LSTM has three gates to protect and control the cell state: forget gate, update gate, and output gate. Some of the functions and principles of each gate are as follows.

Forgetting gate: the internal sigmoid function determines what information is discarded from the cell state and is calculated as follows:

$$T_f = \sigma(W_f[a^{<t-1>}, x^{<t>}] + b_f) \tag{1}$$

Among them, $a^{<t-1>}$ is the output at the moment of $(t-1)$, $x^{<t>}$ is the input to this layer at the moment of t, W_f is the weight of each variable, b_f is a bias term, σ is a function of $sigmoid$, The form is: T_f between 0~1, indicates the value output to each in cell state $c^{<t-1>}$,1 Indicates "fully reserved",0Indicates "complete abandonment".

Update gate: the internal $sigmoid$ function that determines which values we will update, and the tanh function creates a new $\tilde{c}^{<t>}$ vector of candidate values, added to the cell state, again by multiplying the old cell state by the oblivion gate (T_f), forgetting part of the old information; then add $T_u * \tilde{c}^{<t>}$, updating the current cell status, calculated as follows:

$$T_u = \sigma(W_u[a^{<t-1>}, x^{<t>}] + b_u) \tag{2}$$

$$\tilde{c}^{<t>} = \tanh(W_c[a^{<t-1>}, x^{<t>}] + b_c) \tag{3}$$

$$c^{<t>} = T_u * \tilde{c}^{<t>} + T_f * c^{<t-1>} \tag{4}$$

Among them, T_u between0～1, tanh which is the hyperbolic tangent excitation function, Outputs a value between -1 and 1, $c^{<t-1>}$ denotes the cell state value $(t-1)$, $\tilde{c}^{<t>}$ indicates that the information to be recorded is extracted from the input information at the moment t, $c^{<t>}$ indicates the updated cell state value.

Output gate: the internal $sigmoid$ function will determine how much information to output, Use function tanh to process $c^{<t>}$, Get a value between -1 and 1, $\mathrm{T_o}$ Multiplying $\mathrm{T_o}$ and $c^{<t>}$ Obtain the output value at the time t,The calculation formula is as follows:

$$\mathrm{T_o} = \sigma(W_o[a^{<t-1>}, x^{<t>}] + b_o) \tag{5}$$

$$a^{<t>} = T_o * c^{<t>} \tag{6}$$

So far, the internal processing of 1 neuron is completed by 3 control gate mechanisms, which allows the LSTM model to effectively use the input data to form a memory of long past data.

2.1.2 Data processing

Standardized value, is to speak of a single number in the set and the set of the mean value of the result of subtraction divided by the standard deviation of the set to obtain the standardized results, the method is similar to the normal distribution standardized conversion, the conversion function formula is.

$$z = \frac{x - \mu}{\sigma} \tag{7}$$

$$z = \frac{x - \mu}{\sigma} \tag{8}$$

The formula x is the original value to be standardized, μ is the mean, σ is the standard deviation, and is not equal to 0. A fraction of the standardized value represents the example between the original value and the mean of the set, calculated as the standard deviation. The value exists between positive and negative values, below the mean are negative values, and vice versa are positive values, the range is $[-\infty,+\infty]$, the data mean is 0 and the variance is 1.

The loss function (Loss function) is used to estimate the degree of inconsistency between the predicted value X and the true value Y of the network model, which is a non-negative real-valued function, usually denoted by L（Y，f（x））. The smaller the loss function is, the better the robustness of the model is. The loss function is a core part of the empirical risk function and an important component of the structural risk function.

The model takes the very commonly used mean squared error loss, squared loss can also be understood as the least squares method, which is generally more common in regression problems and is calculated as follows.

$$L=\sum_{i=1}^{N}(Y-f(x)) \tag{9}$$

$$MSE=\frac{1}{N}\sum_{i=1}^{N}(Y-f(x)) \tag{10}$$

The relative deviation is the absolute deviation of a measurement as a percentage of the mean value and is expressed as follows.

$$\text{Relative Deviation} = [(\text{Label value}-\text{Predicted value}) / \text{Label value}]*100\% \tag{11}$$

Pick the right parameters.10 neurons for 250 training rounds. To prevent gradient explosion, set the gradient threshold to 1. Specify an initial learning rate of 0.005 and reduce the learning rate after 125 training rounds by multiplying by a factor of 0.2.

3. Results and Discussion

3.1 The establishment of the simulation model

Gold, Bitcoin daily price forecast results: A comparison of the forecast price of gold with the true price is shown in figure 4.

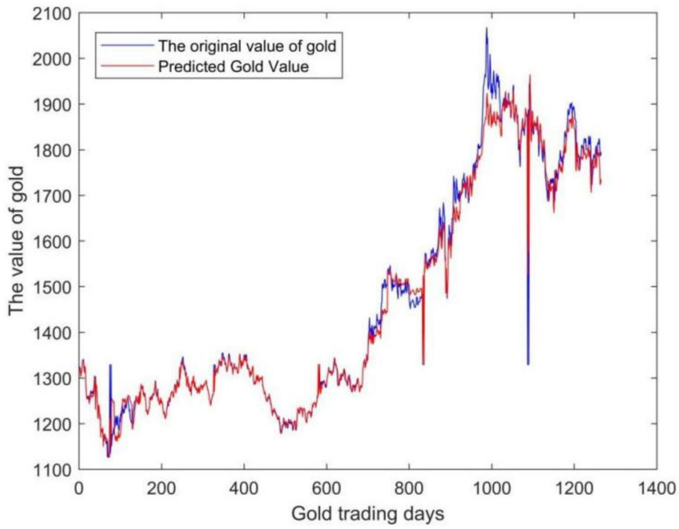

Figure 4. Gold price comparison chart

The resulting graph shows that the trend of the forecast results is approximately the same as the trend of the original data, and the forecasted data is used to determine the return on gold. A comparison of the predicted price of Bitcoin with the true price is shown in figure 5.

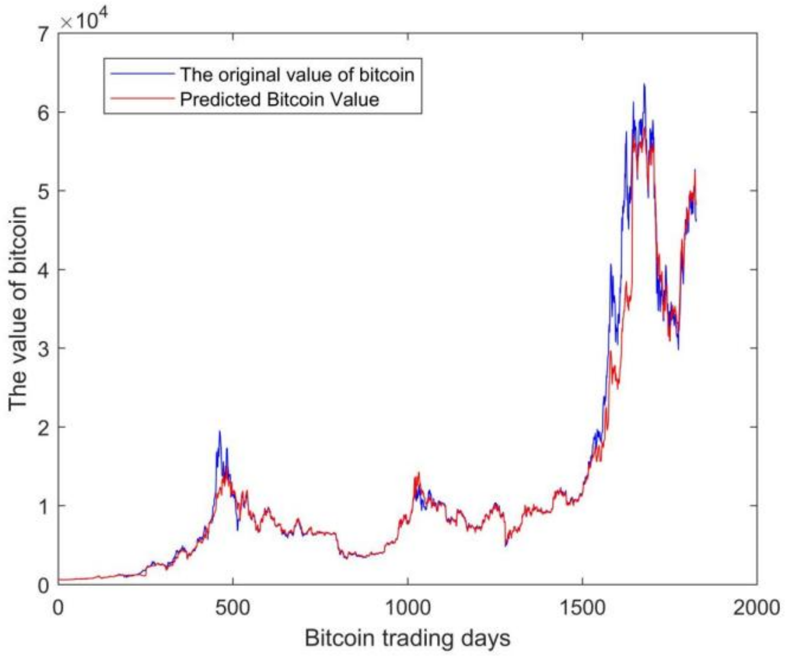

Figure 5. Bitcoin Price Comparison Chart

The graph shows that the trend of the predicted result is approximately the same as the trend of the original data, and the predicted data is used to determine the bitcoin return.

Gold, bitcoin yields to solve: The formula for predicting the return on the two assets is as follows.

$$K = \frac{M_{I+1} - M_{\mathrm{i}}}{M_{\mathrm{i}}} \tag{12}$$

K denotes the rate of return, M_{I+1} denotes the asset price on the day after, and denotes the asset price on the day of.

The forecast for gold yields is shown in figure 6.

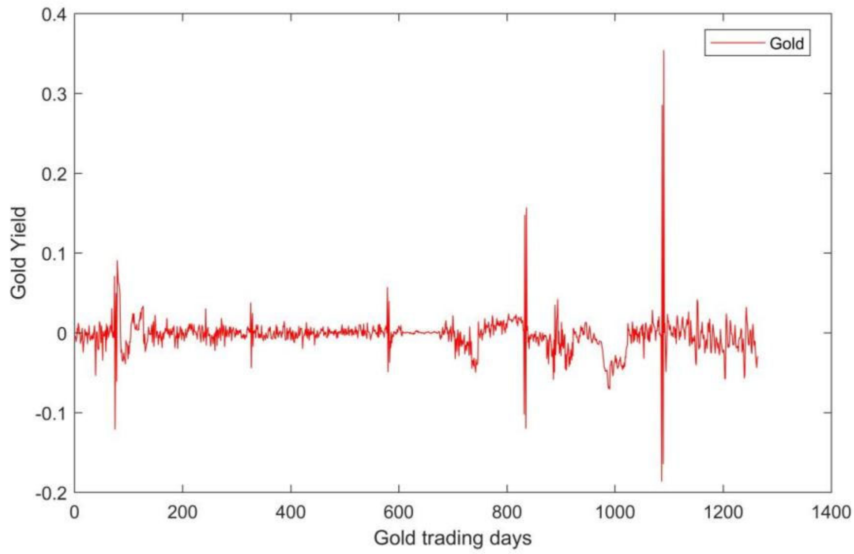

Figure 6. Gold Yield Forecast Chart

The graph shows that the return on gold fluctuates more smoothly. The projected bitcoin returns are shown in figure 7. The resulting graph shows that bitcoin's return is more volatile.

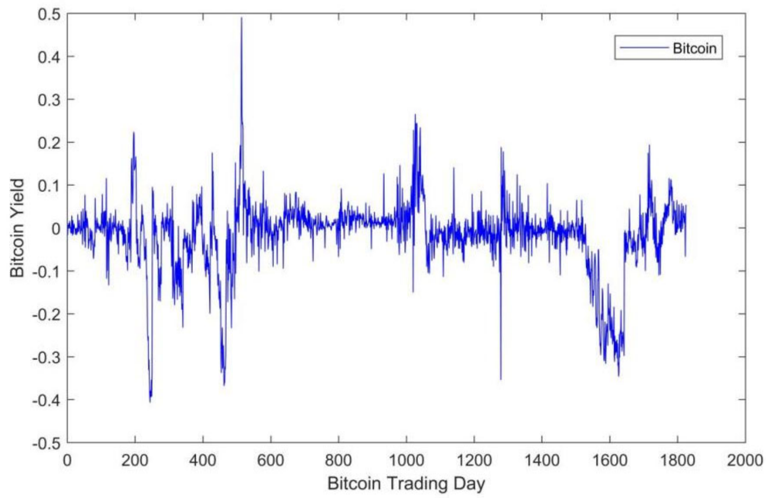

Figure 7. Bitcoin Yield Forecast Chart

Fluctuation of total assets: As the corresponding principal moves through the two trading markets for gold and bitcoin, the total assets owned by the trader change as each day the trader buys, holds, or sells assets in his portfolio, as shown in figure 8.

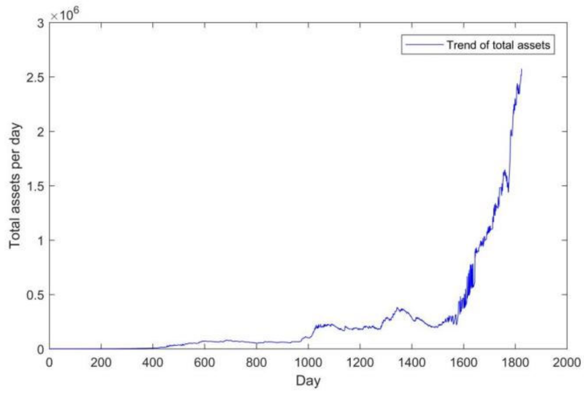

Figure 8. Trend of total assets

The graph shows that as the trading strategy is implemented, the total assets owned by the trader gradually increase at an increasing rate.l

The final assets can be calculated based on the corresponding total program table, as shown in Table 1: $257,4973.177.The large data prediction model for the user's electricity consumption is implemented in the Clementine software.

Table 1. the corresponding total program table

Date	Bitcoin Shares	Bitcoin unit price	Number of gold shares	Gold unit price	Total Funds
2016-12-19	0	780.3468123	0	1136.25	1000
2016-12-20	0.869818809	790.324157	0	1125.7	987.4388169
……	……	……	……	……	……
2018-4-6	7.710953801	6685.130426	0.232140066	1331.2	51857.75672
2018-4-7	0	7014.455514	0.232140066	1331.2	53339.59468
2018-4-8	7.757551579	6856.429435	0.232140066	1331.95	53498.30396
……	……	……	……	……	……
2018-7-30	0	8069.433271	0.232140066	1220.95	78573.41491
2018-7-31	10.13413152	7812.5556	0.232140066	1219	79456.44467
2018-8-1	0	7946.072303	0.232140066	1215.45	75837.63833
2018-8-2	0	7798.317774	0.232140066	1216.3	75837.83565
……	……	……	……	……	……
2021-9-8	0	52354.5123	0.357924086	1802.15	2514934.105
2021-9-9	0	49181.25634	0.357924086	1786	2514928.325
2021-9-10	53.47417359	48141.61578	0.357924086	1788.25	2574973.177

3.2 Discussion

There are two ways to prove that a model can provide an optimal strategy, namely by rationalizing the parameters of the model itself and by optimizing it compared to other models. As the daily prices of gold and bitcoin in the dataset are provided over five years under real conditions, we use the error of the dataset predicted by the model compared to the original dataset as part of the judging evidence.

we selected three aspects as the judging evidence: rationalization of the model parameter setting, optimization of the model compared to the original data set, and error of the predicted data set - original data set. In the optimization proof of the prediction model, we elaborate on the three parts of evidence on the advantages that the tuning of the parameters of the neural network (the number of neurons, the number of performance iterations), the errors (the value of gold, bitcoin, the average relative error between the real value of the return and the predicted value, and the root mean square error) have compared to other models; in the optimization proof of the strategy model, we demonstrate the advantages of the model by the proportional placement of funds and the In the proof of optimality of the strategy model, we demonstrate the superiority of the model by the proportional placement of funds and the financial efficiency.

First, we build a sensitivity analysis model without considering risk, and get the total asset change under different strategies by changing the initial proportion of funds invested in gold and bitcoin; then we introduce risk, and we build a risk factor evaluation model again with the help of the sensitivity analysis model and judge the sensitivity of the strategy to transaction costs based on the two models together; finally, we conclude that the more bitcoin we buy, the higher the risk, and therefore we need to adjust the fund allocation.

4. Conclusions

To address the problem of investing with a certain initial principal and maximizing returns over five years, we propose a new series of models. From predicting daily returns on gold and bitcoin over five years to planning a daily money flow strategy. These models have high accuracy and robustness.

The LSTM neural network-based forecasting models can combine known transaction data of assets with further processing to obtain asset returns and investment risk rates, which is a good solution to the problem that asset forecasting is highly nonlinear and suffers from gradient dissipation and gradient explosion phenomena. We can use the proposed model to forecast the daily price of gold bitcoin over five years and get more specific about the daily returns of the two assets. And based on the predicted dataset the decision of the capital flow is made through the planning model with the goal of optimal return, giving the daily trading strategy for five years and calculating the final value of the principal held by the trader.

We optimize the model in terms of both the reasonableness of the model's parameter settings and optimality compared to other models, and the error of the model's predicted dataset compared to the original dataset is also used as one of the judging evidence. the LSTM neural network can better control how it handles the current input and previous cell states when calculating the final hidden state, and this mechanism makes the final hidden state even in This mechanism allows the final hidden state to fluctuate relatively smoothly even when the state of the cell is changing rapidly. The subsequent sensitivity analysis not only effectively justifies the model, but also illustrates the extent and process of the impact of transaction cost on the strategy and outcome.

References

[1] BudihartoI Widodo. Data science approach to stock prices forecasting in Indonesia during Covid-19 using Long Short-Term Memory (LSTM)[J]. Journal of Big Data,2021(1).

[2] Vidal, Andres, Kristjanpoller, Werner. Gold volatility prediction using a CNN-LSTM approach[J]. Expert Systems with Application,2020(Nov.): 113481.1-113481.9, 2020.

[3] Khalil, Kasem, Eleash, Omar, Kumar, Ashok. Economic LSTM Approach for Recurrent Neural Networks[J]. IEEE Transactions on Circuits and Systems, II. Express briefs,2019(11):1885-1889,2019.

[4] Kang, Sang Hoon, Mciver, Ronp, Hernandez, Jose Arreola. Co-movements between Bitcoin and Gold: A wavelet coherence analysis[J]. Physica, A. Statistical mechanics, and its applications,2019.

[5] Musshoff Oliver, Hirschauer Norbert. Investment planning under uncertainty and flexibility: the case of a purchasable sales contract[J]. The Australian Journal of Agricultural and Resource Economics,2008(1):17-36,2008.

[6] Martin Hall. Risk-based investment planning[J]. Water & Waste Treatment, 2005 (10).

[7] Ge Zhang, Xin Zhang, Hao Guo. The Relationship Between Investor Sentiment and Stock Market Volatility: Based on the VAR Model[C]. 2018:173-180.

[8] Jean-Marie Dufour, Tarek Jouini. Finite-sample simulation-based inference in VAR models with application to Granger causality testing [J]. Journal of Econometrics, 2006 (1/2) :229-254,2006.

[9] Alexander Mironychev. International Federation of Automatic Control, Stability of ecological economic linear models (sensitivity analysis) [C]. 1998:29-33.

[10] Benedetto Manganelli, Pierluigi MORAorano, Francesco Tajani. Risk assessment in estimating the capitalization rate[J]. WSEAS Transactions on Business and Economics,2014(Pt.1):199-208.