Portfolio Decision Model Based on NIWPSO-LSTM

Xinyi Feng[1](✉), Mingshang Chen[2], and Zhengrong Hou[2]

[1] School of Economics and Management, North China Electric Power University, Beijing, China
ncepuFXY@ncepu.edu.cn

[2] School of Energy, Power and Mechanical Engineering, North China Electric Power University, Beijing, China

Abstract. As a cross product between computer science and financial science, the primary purpose of quantitative investment is to explain the formation principle of financial asset prices and to predict the future price of financial assets. With gold and bitcoin daily price data from London Bullion Market Association and NASDAQ, we develop a model that uses only the past stream of daily prices to date to determine each day if the trader should buy, hold, or sell their assets in their portfolio. Firstly, we choose the Long Short-Term Memory neural network based on the improved particle swarm algorithm (NIWPSO-LSTM) to predict the price. We use metabolic grey model (MGM(1,1)) to correct the price in the initial period, to make up for the shortcomings of LSTM which needs a large number of training sets to achieve better prediction results. Secondly, we quantify the return and risk of portfolio investment, establish a nonlinear programming model, and use Monte Carlo simulation method to solve the initial solution. Last, we use indicators to verify the accuracy of our model. All the results show that our model is able to provide extraordinary decision support in a real investment environment.

Keywords: Long Short-Term Memory · Portfolio Decision · Nonlinear Programming · Particle Swarm Optimization

1 Introduction

Financial market plays an important role in a country's economic development and social development. Correspondingly, some financial activities generated through the financial market affect all aspects of national economic and social activities, and affect the economic development of other countries in the world.

Investment strategy research has always been a hot issue in the financial field, which aims to help investors allocate funds rationally, balance income risk, and achieve the purpose of accumulating wealth. With the increasing prosperity and development of the capital market, stock, funds, bonds and other diversified investment products emerge in endlessly, a large number of investors into the market, different investment concepts and strategies came into being. In the complex and diverse real financial markets, the performance of a single investment strategy often fluctuates greatly. Therefore, portfolio investment strategies are proposed and widely used. At the same time, a good data support is indispensable.

Z. Zeng et al. (Eds.): ECIT 2022, AHE 11, pp. 215–227, 2023.
https://doi.org/10.2991/978-94-6463-005-3_22

2 Overview of Research Technique

2.1 Research Content

Market traders buy and sell volatile assets frequently, with a goal to maximize their total return. There is usually a commission for each purchase and sale. Two such assets are gold and bitcoin.

In this paper, we develop a model that uses only the past stream of daily prices to date to determine each day if the trader should buy, hold, or sell their assets in their portfolio.

2.2 Data Description

We've got two documents, historical price data for gold and bitcoin from London Bullion Market Association and NASDAQ.

2.2.1 Data Preview

See Figs. 1 and 2.

2.2.2 Data Pre-processing

Gold trading does not open at the weekend. In addition, there are some missing date and price data in the document. Our work on pre-processing gold price data is shown as Fig. 3.

Fig. 1. Gold daily prices, U.S. dollars per troy ounce. Source: London Bullion Market Association, 9/11/2021.

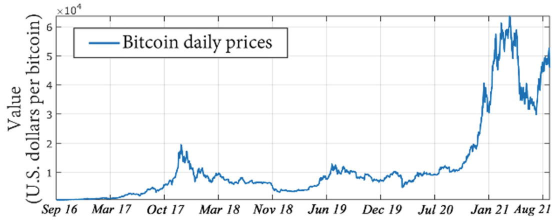


Fig. 2. Bitcoin daily prices, U.S. dollars per bitcoin. Source: NASDAQ, 9/11/2021.

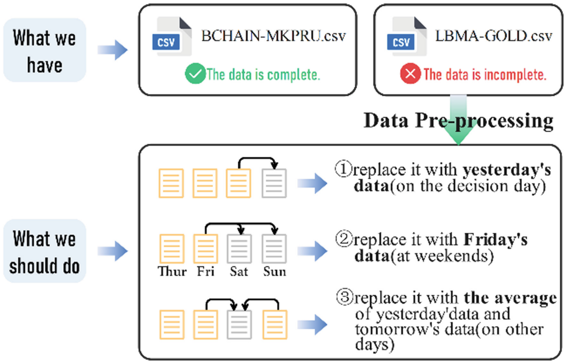


Fig. 3. Data Pre-processing diagram.

Table 1. Notations used in this paper.

Symbol	Description	Unit
N_i	the i_{th} investment scheme	
$S_{i,j}$	the risk of the i_{th} investment scheme in the j_{th} day	
S_j	total risks of the investment scheme in the j_{th} day	
P_i	trade costs to be paid of the i_{th} investment scheme	dollar
$X_{i,j}$	the amount of the i_{th} investment scheme in the j_{th} day	dollar
Q_j	optimization degree of the investment scheme in the j_{th} day	
B_j	trade costs to be paid in the j_{th} day	dollar
R_i	expected rate of return of the i_{th} investment scheme	
$L_{i,j}$	the amount's proportion of total of the i_{th} investment scheme in the j_{th} day	

2.3 Our Work

We use the Long Short-Term Memory neural network (LSTM) suitable for studying time series problems in deep learning to predict the price trend of gold and bitcoin. We added grey prediction and particle swarm optimization to correct LSTM, which greatly improved the prediction accuracy of LSTM.

After forecasting the price, we quantify the return and risk of portfolio investment, establish a nonlinear programming model to give the best trading strategy for every day, and calculate the final return after five years.

We establish *RMSE*, *MAE* and *S* to verify the prediction accuracy. And by constantly changing the risk tolerance, the relationship between risk tolerance and optimal return is obtained.

2.4 Notations

The key mathematical notations we used are listed in Table 1.

3 Stock Price Forecast Model Based on NIWPSO-LSTM

In this section, we simulate the real situation of investment in gold and bitcoin, and predict the future price only based on the existing previous price data every day. Our daily forecast work is shown in Fig. 4.

3.1 Why LSTM?

Before investing, we need to have a full understanding of the characteristics and properties of the product, so as to determine the trend of future price changes and provide good data support for investment decisions.

The stock price curve is highly nonlinear and lacks regularity. Figure 5 shows the historical closing price of NASDAQ. We know that gold and bitcoin, like stocks, have strong volatility and randomness, so they are not suitable for financial time prediction models based on statistical methods, such as ARIMA, GARCH, and so on.

Recurrent Neural Network (RNN) is a kind of neural network used to process sequence data, such as translation problem, prediction of time series data problem, whose performance is better than the traditional neural network. The standard RNN structure is shown on the left side of Fig. 6(a), and the RNN structure based on time expansion is shown on the right side of Fig. 6(a).

However, when trying to deal with the prediction with large amount of data, RNN will miss the important information at the beginning (Huang 2022). Another basic problem

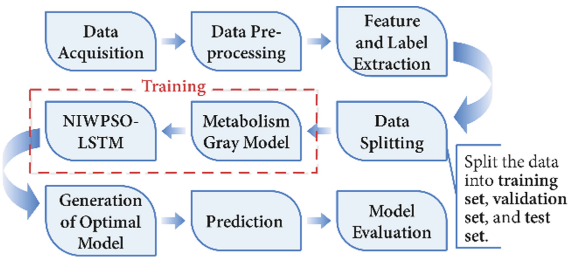
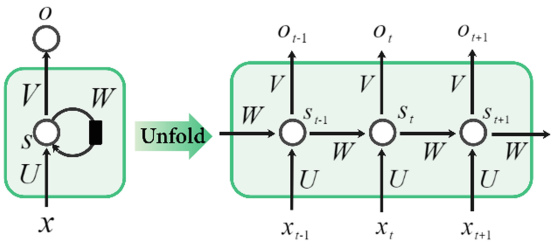
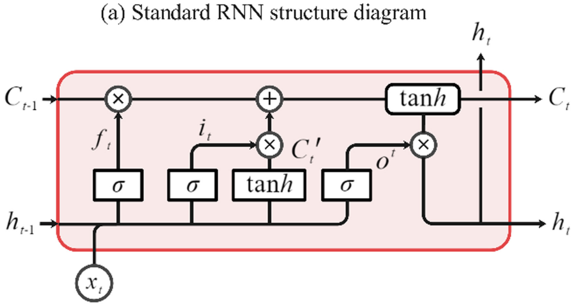


Fig. 4. Daily work of the Forecast Model

Fig. 5. NASDAQ daily prices, U.S. Dollars per share. Source: National Association of Securities Dealers Automated Quotations, 9/11/2021



(a) Standard RNN structure diagram



(b) Standard LSTM structure diagram

Fig. 6. Standard RNN and LSTM structure diagram.

of RNN is gradient vanishing and gradient explosion, which results in low learning rate or inability to learn when the network is deepened.

In order to overcome the problem of RNN's gradient disappearance, Hochreiter and Schmidhuber proposed Long Short-Term Memory (LSTM) neural network units to replace RNN units (Sepp 1997). LSTM network is a special recurrent neural network, which is improved based on RNN. Figure 6(b) shows the expansion structure of LSTM with only one nerve unit in one circulation layer.

3.2 The Improvement of LSTM

3.2.1 Price Correction for the Initial Period—MGM(1,1)

Gray Model is a prediction method that establishes a mathematical model to predict through a small amount of incomplete information. It is based on the past and present development law of objective things, with the help of scientific methods to describe and analyse the future development trend and situation, and form scientific assumptions and judgments.

In the early stage of investment, the number of samples is small, we can get little information. LSTM needs a large number of training samples, so we use Metabolism Gray Model (MGM(1,1)) to modify the model before investment.

3.2.2 Parameter Optimization—NIWPSO

Combining PSO algorithm with LSTM network can effectively realize the optimization of LSTM network parameters. In the PSO algorithm, the only connection between different dimensions of optimization problem space is the introduction of objective function. The position of the i particle in the D-dimensional search space is expressed as $x_i = \left(x_{i,1}, x_{i,2}, \cdots x_{i,D}\right)$, and the best position it has experienced (fitness value) is denoted as $p_i = \left(p_{i,1}, p_{i,2}, \cdots p_{i,D}\right)$. The velocity of the i_{th} particle is expressed as $v_i = \left(v_{i,1}, v_{i,2}, \cdots v_{i,D}\right)$. For each generation, its $(d+1)_{\text{th}}$ dimension $(1 \le d+1 \le D)$ changes according to the following formulas:

$$\begin{aligned} v_{i,d}^{t+1} &= \omega v_{i,d}^{t} + c_1 rand_1 \left(p_{i,d}^{t} - x_{i,d}^{t}\right) \\ &\quad + c_2 rand_2 \left(p_{g,d}^{t} - x_{i,d}^{t}\right) \end{aligned} \tag{1}$$

$$x_{i,d}^{t+1} = x_{i,d}^{t} + v_{i,d}^{t+1} \tag{2}$$

where, $v_{i,d}^{t+1}$ is the velocity of the i_{th} particle at the $(t+1)_{\text{th}}$ iteration; $x_{i,d}^{t+1}$ is the position of the i_{th} particle at the $(t+1)_{\text{th}}$ iteration; ω is the inertia factor. c_1 and c_2 are individual and social learning factors, $rand_1$ and $rand_2$ are random functions in the range [0,1].

In order to match the network structure of the model with the characteristics of stock price data, we introduce the nonlinear dynamic inertia weight improved particle swarm optimization (NIWPSO) (Borowska 2017). According to the adjustment of ω, the algorithm can flexibly adjust the global optimization ability and local optimization ability, and improve the accuracy of LSTM parameter optimization. The updated formula is as follows:

$$\omega = \omega_{\max} - (\omega_{\max} - \omega_{\min}) \ln\left(1 + 1.5\left(\frac{t}{t_{\max}}\right)^{k}\right) \tag{3}$$

The fitness function formula is as follows:

$$f_{it_i} = \frac{1}{K} \sum_{i=1}^{K} \frac{y_i - \hat{y}_i}{y_i} \times 100\% \tag{4}$$

where, K is the number of data sets; y_i is the actual value of the verification sample; $\hat{y}_i$ is the predicted value of the verification sample. In the LSTM algorithm based on particle swarm optimization, the optimal position vector of particles in particle swarm is used as the initial value of each super parameter in LSTM network. Then according to the formula 4, we can calculate the fitness value of each particle with the smallest error.

3.3 The Solution of Forecast Model

Now, we have a complete system that can predict future prices according to the known historical prices. We use Python to code the Forecast Model. After successfully running the program, we get the final prediction results as shown in Fig. 7.

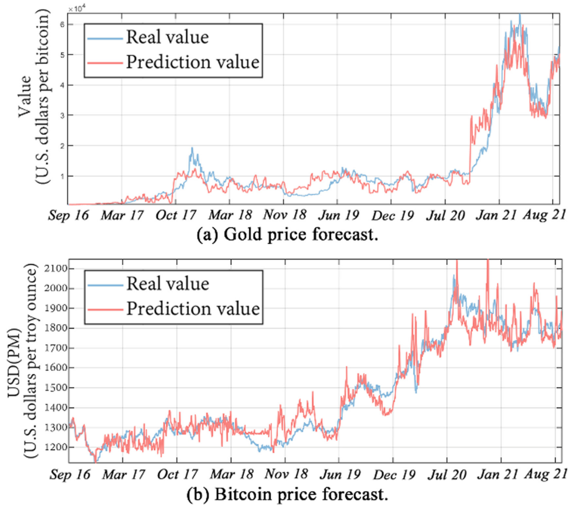


Fig. 7. Gold and bitcoin price forecast.

4 Portfolio Decision Model Based on Markowitz Theory

We use Markowitz's mean-variance model to quantify portfolio investment, and then use linear programming and nonlinear programming to determine the optimal solution of investment. Furthermore, we allocate the funds held by investors to different investment projects according to a certain proportion in order to achieve an optimal equilibrium relationship between overall returns and risks.

4.1 Quantification of Income

(1) The income of the two portfolios.
We assume that the return rates of the two investment schemes (N_1 and N_2) are respectively, and investors invest funds in proportion R_1 and R_2. Then the return rate of the portfolio, R can be expressed as:

$$R = L_1 R_1 + L_2 R_2 \tag{5}$$

(2) Portfolio returns of n investment schemes.

$$R = N_1 L_1 + N_2 L_2 + \cdots + N_n L_n \tag{6}$$

4.2 Quantification of Risk

4.2.1 Risk of Two Portfolios

We use the variance of return to measure the size of the investment risk, and use the σ^2 to identify the variance. The variance of the two investment options is:

$$\sigma^2 = L_1^2 \sigma_1^2 + L_2^2 \sigma_2^2 + 2 L_1 L_2 \mathrm{Cov}(N_1, N_2) \tag{7}$$

where, $\mathrm{Cov}(N_1, N_2)$ is the covariance between the returns of two investment schemes. In order to better measure the correlation between the two investment schemes, we introduce the first relation number to overcome the problem of inconsistent data magnitude, which is as follows:

$$\rho_{12} = \frac{\mathrm{Cov}(R_1 R_2)}{\sigma_1 \sigma_2} \tag{8}$$

Thus portfolio risks can also be recorded as:

$$\sigma^2 = L_1^2 \sigma_1^2 + L_2^2 \sigma_2^2 + 2L_1 L_2 \rho_{12} \sigma_1 \sigma_2 \tag{9}$$

The correlation coefficient can better measure the correlation between the changes in the returns of the two investment schemes. When the correlation coefficient is positive, the return of the two investment schemes (N_1 and N_2) are positively correlated, that is, their return change in the same direction. The correlation coefficient is negative, and the return of N_1 and N_2 change inversely. When the correlation coefficient is 0, the return of N_1 and N_2 is not related.

4.2.2 Risk of n Portfolios

$$\begin{aligned} \sigma^2 = {} & L_1^2 \sigma_1^2 + \cdots + L_n^2 \sigma_n^2 + 2L_1 L_2 \rho_{12} \sigma_1 \sigma_2 \\ & + \cdots + 2L_{n-1} L_n \rho_{n-1,n} \sigma_{n-1} \sigma_n \end{aligned} \tag{10}$$

4.3 Linear Programming and Nonlinear Programming

Based on the analysis of the appeal, we can establish the planning model with the following variables:

- We have n investment strategies. We take X_i as the proportion of the i_{th} investment.
- We take B as the proportion of transaction costs in total investment amount.
- We take p_i as the transaction costs of the i_{th} investment strategy.
- We take R_i as the indicated yield of the i_{th} investment strategy.
- We take S_i as the risk of the i_{th} investment strategy, and take S as the total risk.
- Thus we can obtain the planning model as follows.

$$\max \left(\sum_{i=1}^{n} X_i R_i \right)$$

s.t.

$$\begin{cases} B = \sum_{i=1}^{n} P_i X_i \\ \sum_{i=1}^{n} + B = 1 \\ 0 \le X_i \le 1 \\ L_i = \frac{X_i}{\sum_{i=1}^{n} X_i} \\ \sum_{i=1}^{n} L_i^2 \sigma_i^2 + 2 \sum_{i=1}^{n} \sum_{j=1}^{n} L_i L_j \sigma_{i,j} < 30\% \end{cases} \tag{11}$$

4.4 Investment Programme Decision-Making

4.4.1 Determination of Initial Value—Monte Carlo Simulation

In the process of solving nonlinear programming, the selection of initial value x_0 plays a vital role in the planning results. Therefore, before the nonlinear model, we first use Monte Carlo simulation method to find a Monte Carlo solution, and then use this solution as the initial value of nonlinear programming, in order to improve the speed and accuracy of nonlinear programming.

4.4.2 Change of Investment Plans

Considering the ever-changing investment market, it is extremely unreasonable to simply maintain an investment proportion, which will lead to risk amplification. So we use the known data to plan the new investment ratio after the daily closing price announcement, and then judge whether to adjust the investment proportion by the value of the scheme optimization Q.

$$Q = M_{j+1} - M_j - \sum_{i=1}^{n} \left| X_{i,j+1} - X_{i,j-1} \right| P_i \tag{12}$$

When $Q > 0$, that is, changing investment increases more income than transaction costs, we choose the new investment plan. When $Q \le 0$, we keep the original investment plan.

The daily best trade strategy and the daily real value curve are shown in Figs. 8 and 9.

Red dots represent buying, blue dots and yellow dots represent selling. Only some trading schemes are shown in the Figs. 8 and 9.

In Fig. 8 and Fig. 9, we can see that our total value curve shows an overall upward trend. The final total value is 4578 dollars, the total return rate is 357%, and the average annual return rate is 35.56%. We can also find that Bitcoin is a high-risk and high-return investment product, so our total income curve is greatly affected by the price changes of Bitcoin. Gold has less volatility than Bitcoin in five years. The greater the risk can be borne, the more likely it is to obtain high returns.

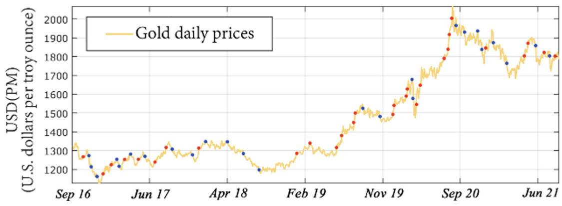


Fig. 8. Daily best trade strategy for gold.

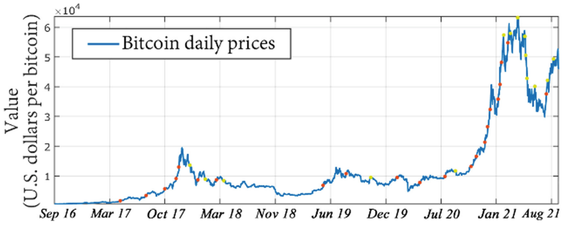


Fig. 9. Daily best trade strategy for bitcoin.

5 Model Evaluation

5.1 Indicator Test

In order to test the effect of prediction results, we use root mean square error (*RMSE*), mean absolute error (*MAE*) and accuracy (*S*) to evaluate, which are as follows.

$$E_{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2} \tag{13}$$

$$E_{MAE} = \frac{1}{N}\sum_{i=1}^{n}|x_i - \hat{x}_i|^2 \tag{14}$$

$$S = \frac{S_{\mathrm{r}}}{S_{\mathrm{all}}} \tag{15}$$

where, x_i indicates the real closing price; $\hat{x}_i$ indicates the forecast closing price; N indicates the total days; S_{r} indicates the correct days, and S_{all} indicates the total forecast days.

RMSE is the square root of the ratio of the square error between the predicted closing price and the real closing price to the total number of days N, which is generally used to explain the dispersion degree in the prediction. *MAE* is the deviation between the predicted value and the real value of the closing price, which is generally used to compare the predicted value and the real value. *S* is the accuracy of prediction. If the error between the predicted value and the real value is less than 5%, the prediction is considered to be correct, otherwise it is considered to be incorrect. The results are shown in Table 2.

Table 2. Error analysis of prediction results.

Product	E_{RMSE}	E_{MAE}	S_{r}	S_{all}	S
Gold	64.13	48.61	1002	1296	77.31%
Bitcoin	3710.58	2479.28	1201	1814	66.21%

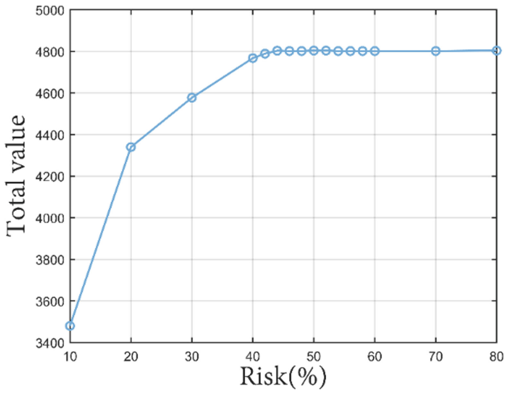


Fig. 10. The Relationship Between Risk and Total Value.

The *RMSE* and *MAE* of bitcoin are much larger than those of gold, which reflects that the price of Bitcoin fluctuates greatly and the prediction is more difficult. We also found that the prediction accuracy is high, so we can think that our prediction model has good prediction effect.

5.2 Risk Assessment

We analyse the income obtained from the same case at different risk levels, and then obtain the most suitable risk level of the model. At this risk level, we can obtain the optimal investment scheme.

From Fig. 10 we can see that when the risk rate changes from 10% to 20%, the total value increases faster. When the risk rate changes from 20% to 41%, the total value has slowed down, but the speed is still not slow. When it comes from 40% to infinity, the total value do not change significantly. Therefore, we suggest that when the investment is conservative, the optimal risk rate is 20%, and when the investors are radical, the optimal risk rate is 41%.

5.3 Strengths and Weakness

5.3.1 Strengths

Strong universality. Our model can not only predict the price trend of gold and bitcoin, but also predict other financial products. As long as we get a certain amount of initial data and indicators, we can get a better model and give a reasonable and robust prediction. Secondly, we establish a portfolio investment model in the decision-making part, which can be used not only in gold and Bitcoin portfolio investment, but also in other financial products portfolio investment.

High accuracy. We establish the LSTM prediction model based on improved particle swarm optimization algorithm, the accuracy of the model is greatly improved compared with the single LSTM prediction.

5.3.2 Weakness

Our current model results are mainly for short-term trading crowd, we only predict the closing price of the day to determine the day's trading strategy. But in fact, long-term holding is more likely to produce high returns than frequent transactions. In addition, since we cannot use the data other than the annex, we do not compare the model horizontally, which cannot prove that our model can improve the best trading strategy.

6 Conclusion

Financial market plays an important role in a country's economic development and social development. With the development of computer networks, deep learning algorithms are more and more frequently used in financial problems. In this paper, based on the existing daily data, we use NIWPSO-LSTM to predict the future data, and then use MGM(1,1) to correct the previous price data.

Then, we quantify the return and risk of portfolio investment, establish a nonlinear programming model, introduce Monte Carlo simulation method to solve the initial solution. The final asset is \$4578, the total yield is 357%, and the average annual yield is 35.56%.

We construct the evaluation indexes of prediction accuracy, including root mean square error ($RMSE$), mean absolute error (MAE) and accuracy (S). Then through analysis and calculation, we obtain $RMSE$, MAE, and S, which verify the accuracy of our prediction model. Secondly, we also explore the relationship between risk tolerance and return rate by constantly changing risk tolerance, and obtain that the maximum return can be obtained when the risk is 41%. Those prove that our mathematical model can provide the best strategy.

References

Borowska B (2017) Nonlinear inertia weight in particle swarm optimization. In: 2017 12th international scientific and technical conference on computer sciences and information technologies (CSIT), vol 1, pp 296–299. IEEE

Dumoulin V, Belghazi I, Poole B, et al (2017) Adversarially Learned Inference

Graves A, Schmidhuber J (2005) Framewise phoneme classification with bidirectional LSTM and other neural network architectures. Neural Netw 18(5–6):602–610

Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780

Huang J, Zhong M, Hu Q (2022) LSTM stock prediction model based on improved particle swarm optimization [J/OL]. J East China Univ Sci Technol (Nat Sci Ed) 1–12. https://doi.org/10.14135/j.cnki.1006-3080.20210616001

Bui V, Nguyen, VH, Pham, TL, Kim J, Jang YM (2020) RNN-based deep learning for one-hour ahead load forecasting. In: 2020 international conference on artificial intelligence in information and communication (ICAIIC), Fukuoka, Japan, pp 587–589. https://doi.org/10.1109/ICAIIC48513.2020.9065071