ISSN 2616-5775 Vol. 5, Issue 10: 60-65, DOI: 10.25236/AJCIS.2022.051010

# Research on Stock Trading Strategy Based on Deep Reinforcement Learning

**Li Tan[1,*], Shuqi Liang[2]**

*[1]School of Information Management, Xinjiang University of Finance and Economics, Urumchi, 830000, China.*
*[2]School of Information Engineering, Zhujiang College, South China Agricultural University, Guangzhou, 510900, China.*
**Corresponding author: tanlifrstentum@qq.com.*

***Abstract:*** *The stock market occupies an important position in the country's overall economic system, and changes in the stock market affect the investment returns of investors and investment institutions. With the advancement of science and technology and the development of the economic market, the price trend of stocks is disturbed by more and more factors, and the traditional forecasting method is becoming more and more difficult to meet people's requirements. Among them, the stock price time series as a dynamic, irregular, non-linear, non-parameterized complex data, brings great challenges to stock price prediction, so The study proposed to use CNN and LSTM models to predict stock prices. First, the data features are extracted using the CNN network structure, and then the stock price is predicted using LSTM. Experiments have proved that the model can be successfully applied to the study of stock forecasting and has a high accuracy in forecasting.*

***Keywords:*** *CNN; LSTM; Stock forecasts*

## 1. Introduction

The financial market occupies an important position in China's modern economic system, with the gradual development and growth of China's financial market, the system is becoming more and more sound, more and more people participate in the field of financial investment, making the demand for financial information services more and more urgent. Analyzing and predicting the behavior of financial markets can effectively help investors reduce investment risks and improve investment returns, so investors will analyze and predict stocks through a lot of data on the stock market to formulate investment plans. In the past, it was generally assumed that financial time series were generated linearly, and models such as ARIMA were often used to predict stock prices. With the development of neural networks, researchers are increasingly applying classic models such as CNN, ICA, and BP to stock price or trend prediction problems. The prediction method also develops more from linear prediction model to nonlinear prediction model that relies on neural networks, such as Wang Dong et al. [1] who combine the principal component method with LSTM, Liu Heng [2] separately studies and compares the effects of LSTM-CNN, RF, CNN, LSTM model prediction. It can be said that stock price prediction based on deep learning is becoming more and more popular.

Most deep learning methods use neural network architectures, so deep learning models are often referred to as deep neural networks. Deep learning models are trained by using large amounts of labeled data, while neural network architectures learn features directly from the data without the need for manual feature extraction [3]. One of the most popular types of deep neural networks, convolutional neural network CNNs, operate by extracting features directly from images without pre-training the relevant features. The network learns relevant features as it trains a set of images. This automated feature extraction enables deep learning models to provide high accuracy for computer vision tasks such as object classification. As a network model, LSTM is designed to solve long-term problems with gradients, so LSTM is good at handling and predicting events with long intervals and large time delays in time series. Combining the above characteristics, the study choose the CNN-LSTM model to complete the prediction of stock prices or trends.

ISSN 2616-5775 Vol. 5, Issue 10: 60-65, DOI: 10.25236/AJCIS.2022.051010

## 2. Data processing

The study selected the more comprehensive CSI 300 index data as the data basis for the research, and selected the stock data of Ping An Bank (000001) from January 20, 2020 to July 19, 2022. 80% of the data is used as the training set, and the remaining 20% is used as the test set to verify the generalization ability of the model [4]. Through Python's tushare financial data interface package, the data obtains five basic data of the selected stock: opening price, closing price, high price, low price and volume. Among them, some of the data of Ping An Bank are shown in Table 1:

*Table 1: Some data of Ping An Bank*

| date | open | high | close | low | volume |
|---|---|---|---|---|---|
| 2020-01-20 | 16.43 | 16.61 | 16.45 | 16.35 | 746074.75 |
| 2020-01-21 | 16.34 | 16.34 | 16.00 | 15.93 | 896603.12 |
| 2020-01-22 | 15.92 | 16.16 | 16.09 | 15.71 | 719464.88 |
| 2020-01-23 | 15.92 | 15.92 | 15.54 | 15.39 | 1100592.12 |
| 2020-02-03 | 13.99 | 14.70 | 13.99 | 13.99 | 2259195.00 |
| ... | ... | ... | ... | ... | ... |
| 2022-07-13 | 14.42 | 14.47 | 13.97 | 13.89 | 2175996.00 |
| 2022-07-14 | 13.70 | 13.71 | 13.37 | 13.29 | 2632120.75 |
| 2022-07-15 | 13.48 | 13.56 | 13.24 | 13.22 | 1598637.50 |
| 2022-07-18 | 13.25 | 13.44 | 13.39 | 13.21 | 1153505.88 |
| 2022-07-19 | 13.34 | 13.43 | 13.42 | 13.28 | 755597.12 |

## 3. Model Introduction

### *3.1 Convolutional neural networks CNN*

As a feed-forward neural network that contains convolutional computation and deep structure, convolutional neural networks are best at processing pictures. Inspired by the human visual nervous system, CNN has two major characteristics, namely, it can effectively reduce the dimension of large data amounts of pictures into small data volumes, and it can effectively retain picture features under the principle of image processing [5]. In simple terms, there are many kinds of network models of convolutional neural networks, but a convolutional neural network model is generally composed of several convolutional layers, pooled layers and fully connected layers, where the convolutional layer is responsible for extracting local features in the image; The pooling layer is used to greatly reduce the order of magnitude of parameters (dimensionality reduction) and effectively avoid overfitting; The fully connected layer resembles the part of a traditional neural network and is used to output the desired result.

### *3.2 Long short-term memory networks LSTM*

LSTM is a special variant of RNN, and RNNs can only have short-term memory due to gradient disappearance. With the further development of research, the LSTM architecture has been continuously improved [6], and the difference between RNN and LSTM today is that there is only one state inside the single cyclic structure of the RNN. The single loop structure of the LSTM has four states inside, as shown in Figure 1 and Figure 2 below.

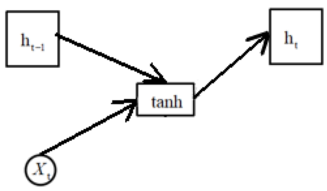


*Figure 1: Structure of RNN unit*

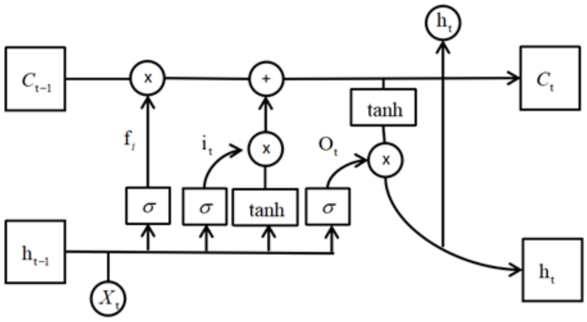


*Figure 2: Structure of LSTM unit*

LSTM cells consist of an input gate, a forgetting gate, an output gate, and a cell state. The most critical of these three gated devices is the forgotten door that can decide to retain or forget the control information [7]. Figure 3 below shows the process of applying the output $h_{t-1}$ and the current data input $x_t$ at the previous moment, and obtaining $f_t$ through the forgetting gate. The evaluation expression is as follows:

$$f_t = \sigma(W_f \bullet [h_{t-1}, x_t] + b_f) \tag{1}$$

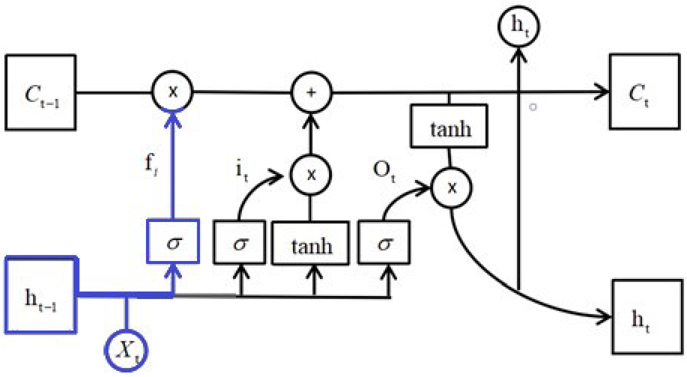


*Figure 3: The forgetting gate in the LSTM unit structure*

Wherein, $h_{t-1}$ and $x_t$ are both used as inputs to multiply with the weight matrix W, and then add up to the deviation matrix b, and finally a vector $f_t$ is obtained under the calculation of the activation function σ, and generally speaking, the sigmoid is used as the activation function [8].

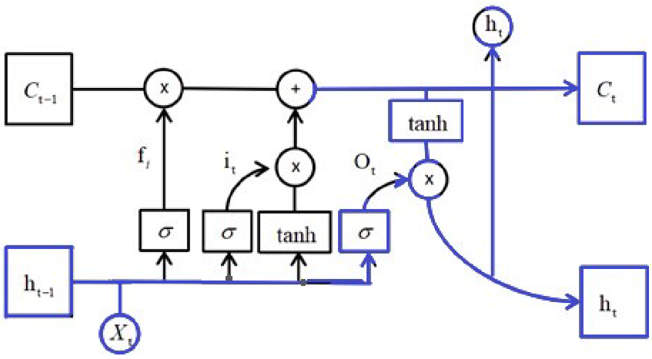


*Figure 4: The input gate in the LSTM unit structure*

After forgetting the information, as shown in Figure 4, it is necessary to determine how much of the input data of the network at the current moment needs to be saved to the LSTM unit through the input gate. That is, the output of the previous moment $h_{t-1}$ and the current data input $x_t$ are applied, $i_t$ is obtained through the input gate, and the process of obtaining the temporary state of the current moment $\tilde{C}_t$ through the unit state. Enter the gate to update the cell state, first determine the value that needs to be updated through the sigmoid layer, and the candidate value vector created by the tanh layer, and multiply the two to get a new candidate value formula as follows:

ISSN 2616-5775 Vol. 5, Issue 10: 60-65, DOI: 10.25236/AJCIS.2022.051010

$$i_t = \sigma(W_i \bullet [h_{t-1}, x_t] + b_i) \tag{2}$$

$$\tilde{C}_t = \tanh(W_C \bullet [h_{t-1}, x_t] + b_C) \tag{3}$$

After the information passes through the input gate and the forgetting gate, the information is integrated, the information that needs to be forgotten is discarded, the useful information is retained, and the cell state Ct at that moment is finally obtained, the formula is as follows:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{4}$$

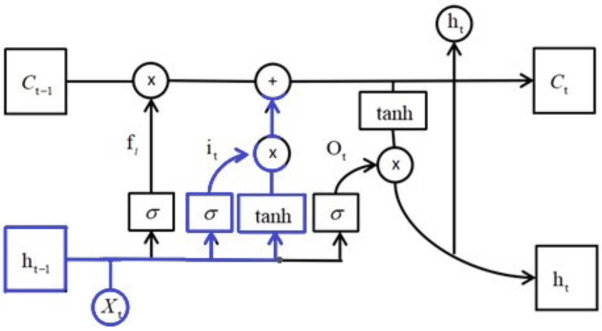


*Figure 5: The output gate in the LSTM unit structure*

Finally, as shown in Figure 5, the output $h_{t-1}$ of the previous moment and the current data input $x_t$ are applied, and the part of the $o_t$ to be output is determined by the sigmoid activation function, and then the resulting $o_t$, with the current cell state $C_t$, is processed by the tanh function to obtain the final output part $h_t$. The formula is as follows:

$$o_t = \sigma(W_o [h_{t-1}, x_t] + b_o) \tag{5}$$

$$h_t = o_t * \tanh(C_t) \tag{6}$$

$h_t$ represents the LSTM unit output information for the current moment.

## 4. Model results

The study obtained data for a total of 591 trading days from January 15, 2020 to July 15, 2022, of which I selected data from the first 493 days as the training set and the remaining 98 trading days as the test set, as shown in the following figure 6.

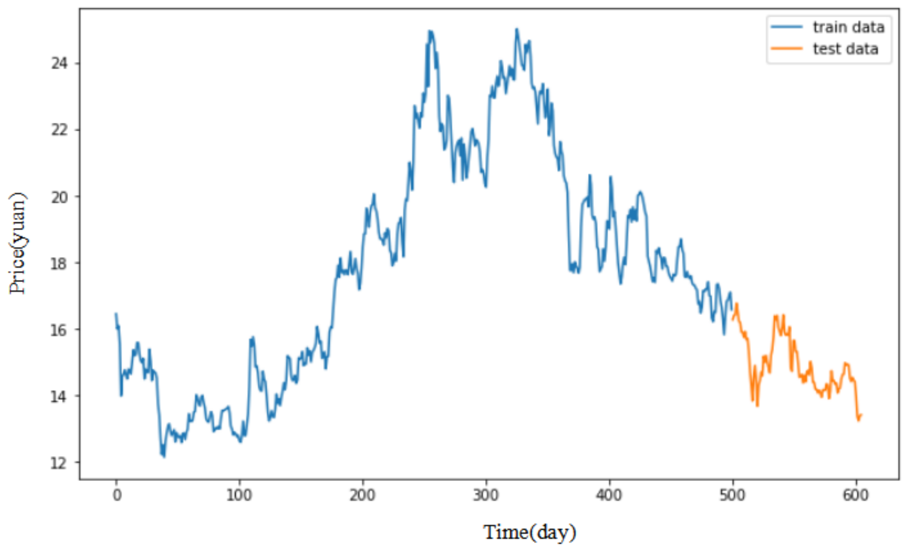


*Figure 6: Distribution of training set and test set*

ISSN 2616-5775 Vol. 5, Issue 10: 60-65, DOI: 10.25236/AJCIS.2022.051010

In Figure 6, the abscissa represents the size of the dataset, which is the trading day occupied by the training and testing sets, and the ordinate represents the stock price.

After determining the training set and the test set, the study used the CNN-LSTM model to find the RMSE difference between the predicted value and the actual value as the evaluation criterion [9], and the evaluation formula is as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2} \tag{7}$$

Where: N is the number of samples, $y_i$ is the actual value, $\hat{y}_i$ is the predicted value, the smaller the value of RMSE, the better the prediction accuracy of the model. Then get a comparison chart of the predicted value and the true value, as shown in Figure 7 below, the ordinate coordinate represents the stock price, and the abscissa coordinate represents the test set trading day.

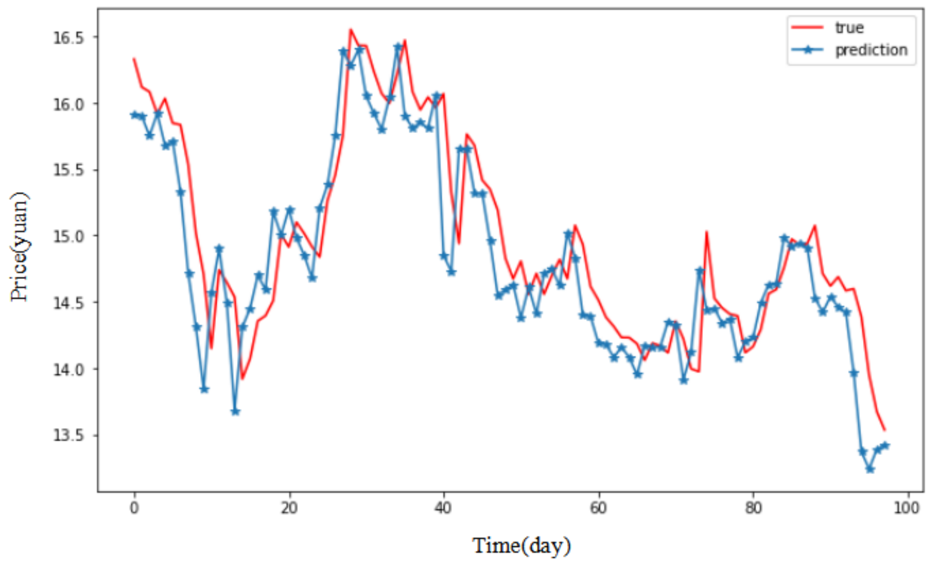


*Figure 7: Comparison of predicted and true values*

As can be seen from Figure 7, when using Ping An Bank as the dataset, the predicted value is relatively consistent with the true value, and after considering the total value of the stock, the prediction yields RMSE=0.3825, which is 58.16% lower than the LSTM model that has not been optimized by the CNN algorithm, which reflects the good effect of the CNN-LSTM model and the high accuracy improvement [10].

## 5. Conclusion

In this study, the CNN-LSTM model was used to predict the trend of stock prices, and the CNN and LSTM were first analyzed, and it was found that the convolutional neural network CNN, which is good at image processing, and the long and short memory network LSTM, which is good at processing and predicting events with long intervals and large time delays in time series, are very suitable for stock prediction. Therefore, it was decided to use the CNN and LSTM models to predict the stock price of Ping An Bank, and by calculating the RMSE and plotting the comparison chart between the predicted value and the true value, it can be found that the accuracy of the method is high. Although this study draws preliminary conclusions, there are still some problems:

1) The stock market is complex and changeable, in addition to the relevant indicator data, the factors affecting the trend of stocks are also the international situation, national policies, industry development and human intervention.

2) The relevant indicator data reflects that the stock market has a lag, and if there is an emergency, the model is difficult to predict in time.

3) The ubiquity of the model has yet to be tested on more datasets. It is also necessary to continuously

ISSN 2616-5775 Vol. 5, Issue 10: 60-65, DOI: 10.25236/AJCIS.2022.051010

study and optimize all aspects of the model in order to further improve the accuracy and speed of the model and achieve more accurate stock forecasts.

## References

*[1] Wang D, Wang X P, Yang C D. A Study of Stock Forecasting Based on LSTM Model of Principal Component Analysis [J]. Journal of Chongqing University of Technology (Natural Science), 2021, 35(2): 282-288.*

*[2] Liu H. Research on Chinese stock trend prediction model based on LSTM-CNN [D]. Xiamen University, 2019.*

*[3] Shen Shanshan, Li Qiumin. Short-term stock price prediction by CNN-LSTM based on attention mechanism [J]. Software, 2022,43(02):73-75.*

*[4] Zhou Zhihua. Machine Learning [J]. Aviation Port, 2018, (02): 94.*

*[5] Chen Q. Q. Studying stock price movements based on CNN-LSTM neural networks [J]. Journal of Economic Research, 2020, (02): 157-160.*

*[6] Luo Xiaoting. Research on deep learning-based stock risk prediction methods [J]. Time Finance, 2022, (02): 30-32.*

*[7] Cao Chaofan, Luo Zenan, Xie Jiaxin, Li Lu. Research on MDT-CNN-LSTM model for stock price prediction [J]. Computer Engineering and Applications, 2022, 58(05): 280-286.*

*[8] Li Chenyang. Research on stock price prediction and quantitative stock selection based on CNN-LSTM [D]. Northwestern University, 2021.*

*[9] Geng Jingjing, Liu Yumin, Li Yang, Zhao Zhexing. A CNN-LSTM based stock index prediction model [J]. Statistics and Decision Making, 2021, 37(05): 134-138.*

*[10] Luo X, Zhang JL. Stock price forecasting based on multi-timescale composite deep neural network [J]. Wuhan Finance, 2020, (09): 32-40.*