Stock Price Prediction Research Based on CNN-LSTM

Jingjuan Zhang [1,*,†], Wingho Chan [2,†], Yiwei Lin [3,†]

[1] College of Information and Electrical Engineering, China Agricultural University, Beijing, China, 100083

[2] College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China, 518000

[3] Maynooth International Engineering College, Fuzhou University, Fuzhou, China, 350108

* Corresponding Author Email: zhangjingjuan@cau.edu.cn

† These authors contributed equally.

Abstract. In recent years, neural network technology has been widely used in the field of stock forecasting, especially in the aspect of LSTM, which has become a research hotspot because of its unique advantages in dealing with time series problems in stock forecasting. By analyzing and summarizing the recent research literature based on LSTM and its extended model, this paper aims to deeply explore the effect and potential limitations of LSTM in capturing the dynamic and nonlinear characteristics of the stock market. By constructing different combination models combining LSTM and CNN, the paper uses normalization to preprocess the split data and realize the generalization prediction of stock price. Compared with the experimental results, a suitable combination model of LSTM and CNN is found. This hybrid model can significantly improve the accuracy and stability of prediction, and also shows the advantages in multi-input data processing. The results of this study not only provide a new technical means for the direction of stock prediction, but also explore a new direction for the further application of deep learning technology in the financial field.

Keywords: LSTM, CNN, CNN-LSTM hybrid model, Stock Price Prediction.

1. Introduction

Neural networks have been widely used in stock price prediction and have demonstrated high accuracy in forecasting in recent years. Researchers have combined traditional neural networks, convolutional neural networks (CNN), long short-term memory networks (LSTM), and other models, as well as introducing advanced technologies such as attention mechanisms and graph neural networks, to construct complex hybrid models that effectively capture the nonlinear features and dynamic changes in financial data. In addition, the combination of traditional financial theory and neural network models has further improved prediction accuracy. These innovations have not only enriched the research paradigm of stock prediction but also provided more scientific and reliable decision support tools for the financial industry.

With regard to the use of neural network for stock price prediction, Zhang Sulin et al effectively improved the model's ability to capture nonlinear features by combining CNN and LSTM [1]. Zhu Zijing et al. introduced ABC algorithm to enhance the robustness of the model to noisy data [2]. Further, Xiao Tiantian's K-mean-LSTM model realizes personalized prediction of different stocks through clustering [3]. Pan Tang et al. 's multi-input LSTM model integrates data from different markets to improve the comprehensiveness of prediction [4]. In addition, researchers have also introduced advanced technologies such as attention mechanisms and graph neural networks into the model, further improving the expressive power of the model. Moreover, Ma Yuanyuan et al.'s research shows that the introduction of investor sentiment data into the model can improve the prediction accuracy [5]. In addition, Wang Jian et al., Chen Menglong et al., Cui Ting et al., applied natural language processing technology to stock prediction and fully mined the information in text data [6-8]. At the same time, the WD-CNN-LSTM model proposed by Cao Yugui et al., by introducing width learning, effectively processes high-dimensional and nonlinear time series data, and can be trained quickly and in parallel [9]. By combining random forest and LSTM, Li Hao et al.

not only improved the generalization ability of the model, but also automatically screened out features that had a greater impact on the prediction results [10]. By using CEEMDAN and LSTM-Attention networks, Sun Chenyu et al. can effectively remove data noise, extract more pure features, and capture the relationship between features to improve the interpretation of the model [11].

Combing through the research on stock price prediction based on LSTM network in recent years in combination with specific literature systems, we find that it only discusses the stacking comparison of the number of network layers, and does not pay attention to the unidirectional distribution of the network. Although relatively mature research results have been achieved, there is still a large space for research. Through in-depth analysis of various LSTM models and their combinations with CNN, this paper discusses the performance of different combination models in capturing the complex dynamics of the stock market, focusing on the impact of different model structures and parameter configurations on the forecasting performance, aiming to provide guidance for researchers and promote the development of the field of stock forecasting.

2. Establishment of Model

2.1. The basic structure and funamental of LSTM

LSTM model is a kind of neural network model proposed by S. ochreiter et al in 1997. Through a complex gating mechanism, as shown in Figure 1, it can better control the mixed ratio of historical semantics and current semantics, ensure the long-term maintenance of gradient transmission, and effectively alleviate the problem of gradient disappearance and gradient explosion in the standard RNN model, which is widely used in the processing and analysis of time series data.

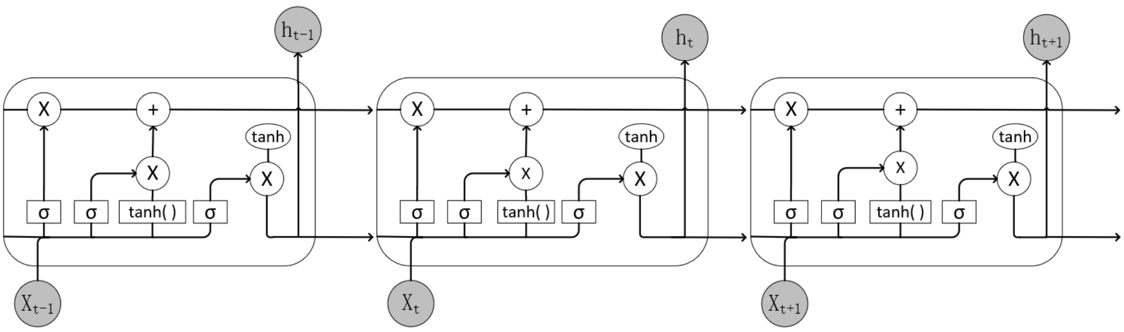

Figure 1. Calculation diagram of LSTM

The LSTM model memory cell is composed of three parts, namely, the forgetting gate, the input gate and the output gate, as shown in Figure 2.

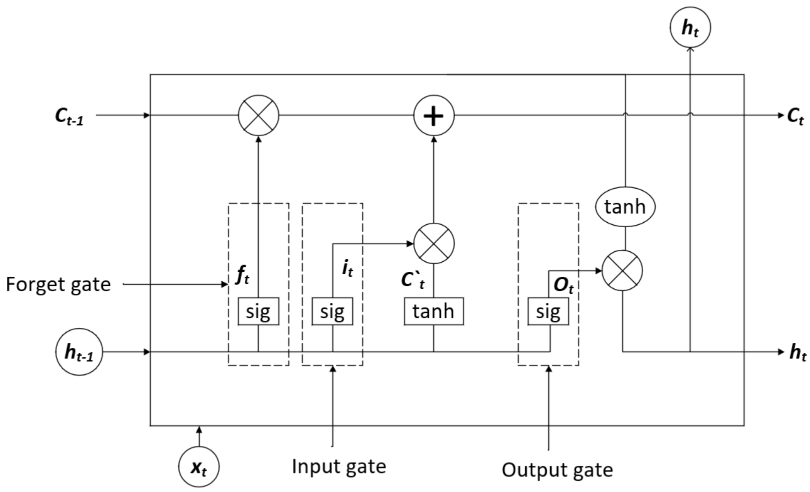

Figure 2. LSTM neuronal gating mechanism

The forward computation process of the LSTM model is as follows:

(1) Calculate the forget gate proportion coefficient ft by inputting the output value from the previous time step and the current input value into the forget gate, and then compute the output value of the forget gate.

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

Wherein $\sigma(\cdot)$ signifies the sigmoid activation function. W_f denotes the weight matrix, b_f represents the bias vector.

(2) The output value from the previous time step and the current input value are fed into the input gate, whereupon, through computational processing, the output value and the candidate cell state along with its proportional coefficient for the input gate are obtained.

$$\begin{cases} \dot{C}_t = tanh\,([[h_{t-1}, x_t]W_C + b_C]) \\ i_t = \sigma([h_{t-1}, x_t]W_i + b_i) \end{cases} \tag{2}$$

The activation function here is replaced with $\tanh(\cdot)$, where W_i, W_C represents the weight matrices and b_i denotes the bias vector.

(3) The historical unit state and the current unit state are respectively scaled by the coefficient f_t and i_t, and then summed together to obtain the unit state at the current moment.

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \dot{C}_t \tag{3}$$

The value of C_t ranges from 0 to 1.

(4) Calculate the proportional coefficient O_t and output feature h_t outward according to O_t.

$$\begin{cases} O_t = \sigma([h_{t-1}, x_t]W_O + b_O \\ h_t = O_t \times tanh(C_t) \end{cases} \tag{4}$$

W_O represents the weight matrices and b_O denotes the bias vector.

2.2. LSTM model combined with CNN model

2.2.1 Introduction of Convolutional Neural Networks

CNN has a feedforward structure and are widely used in tasks such as image classification, object detection and semantic segmentation. In this paper, CNN is applied to the prediction of time series combined with LSTM.

CNN mainly consists of the following parts: the convolution layer, the pool layer, the dropout layer, the full connection layer, and the activation function.

(1) Convolution layer. The convolutional layer is the core part of CNN and is used to extract local features from the input time series data. Through the use of multiple convolution kernels, different time series features can be captured. At the same time, the parameter sharing of convolutional kernel makes the network have translation invariance, which enhances the robustness to the signal.

(2) Pooling layer. Used to reduce the size of feature maps and retain important features in time series data.

(3) dropout layer. Enhance the generalization ability of the network and prevent overfitting in the training phase.

(4) Fully connected layer. By learning the correlation of different positions in the sequence, global feature information is obtained to provide more accurate predictions.

(5) Activation function. The function of the activation function is to introduce nonlinear characteristics and enhance the expression ability of the network. In this experiment, ReLU activation function is used to eliminate negative values, introduce sparsity, solve the problem of disappearing gradient, accelerate the training process and improve the learning ability of the network.

2.2.2 Model network structure combining CNN and LSTM

Since stock price data, as time series data, is characterized by high latitude and nonlinearity, the data is feature extracted using Convolutional Neural Networks (CNN) and the time series data can be selectively learned using Long Short Term Memory Networks (LSTM). In this paper, we combine these two models to realize stock price prediction using CNN composite LSTM.

The model mainly consists of input layer, convolutional layer, pooling layer, dropout layer, LSTM layer and output layer. Among them, the LSTM layer will be experimented using a single-layer LSTM layer, a two-layer LSTM layer, a single-layer bidirectional LSTM layer and a bidirectional two-layer LSTM layer, respectively.

Taking the two-layer bidirectional LSTM network model as an example, the specific network layer structure is shown in Table 1 below.

Table 1. Hierarchical structure of a two-layer bidirectional LSTM network

Layer (type)	Input Shape	Output Shape	Param
Conv1D-1	[[23643, 4, 6]]	[23643, 8, 6]	104
MaxPool1D-1	[[23643, 8, 6]]	[23643, 8, 2]	0
Dropout-1	[[23643, 8, 2]]	[23643, 8, 2]	0
Conv1D-2	[[23643, 8, 2]]	[23643, 16, 2]	400
MaxPool1D-2	[[23643, 16, 2]]	[23643, 16, 1]	0
Dropout-2	[[23643, 16, 1]]	[23643, 16, 1]	0
LSTM-1	[[23643, 1, 16]]	[[23643, 1, 128], [[2, 23643, 64], [2, 23643, 64]]]	41,984
Dropout-3	[[23643, 1, 128]]	[[23643, 1, 128]]	0
LSTM-2	[[23643, 1, 128]]	[[23643, 1, 128], [[2, 23643, 64], [2, 23643, 64]]]	99,328
Flatten-1	[[23643, 1, 128]]	[23643, 128]	0
Linear-1	[[23643, 128]]	[23643, 16]	2064
Linear-2	[[23643, 16]]	[23643, 1]	17

3. Construction and resolution of models

3.1. Research dataset

The research dataset we use is the USDCNY stock price data from 2019 to 2020, which consists of 26,277 pieces of data. We get the dataset from AI Studio and each data contains five fields: stock code, opening price, closing price, and the lowest and highest prices of the trading segment.

3.2. Data preprocessing

Normalize the data before inputting it into the model network, and split the dataset into groups of data every 6 days. Finally, divide the training and testing datasets into a 9:1 ratio. The normalization formula is as follows:

$$X\_std = \frac{X_i - X_{min}}{X_{max} - X_{min}} \tag{5}$$

X_{max} and X_{min} are the maximum and minimum values of component X_i, respectively.

3.3. Model Implementation

The implementation of model building mainly consists of input layer, convolutional layer, pooling layer, dropout layer, LSTM layer, and output layer. The schematic diagram of the algorithm flow for model building is shown in Figure 3.

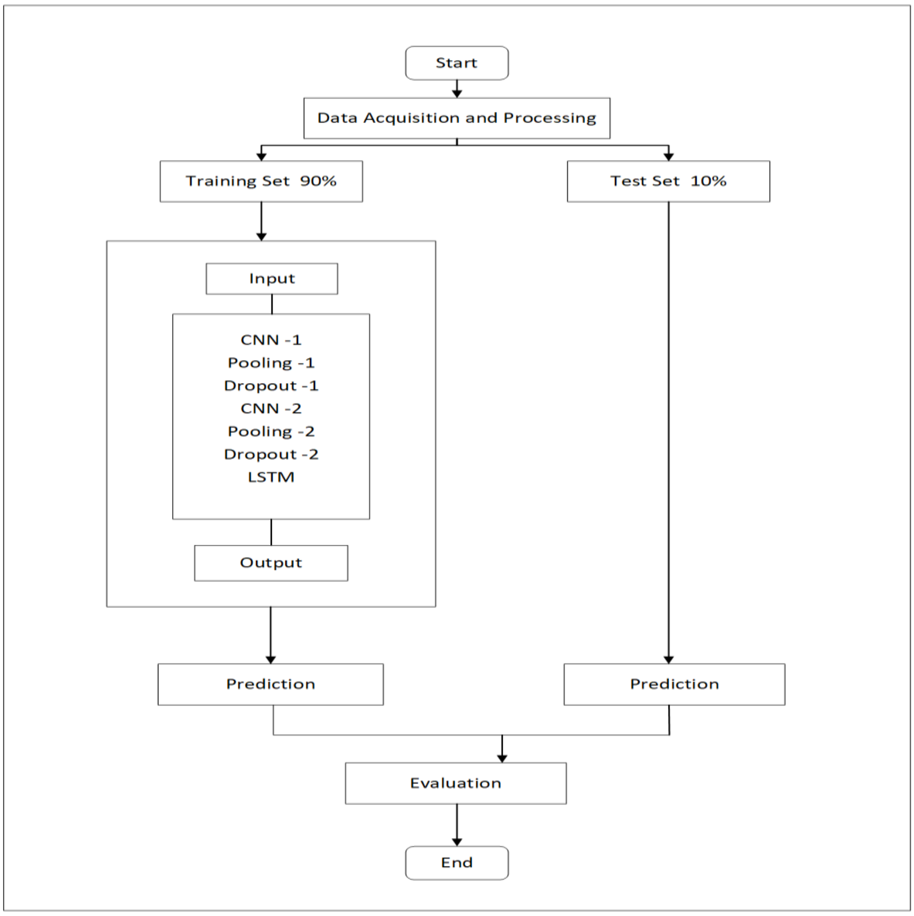

Figure 3. Schematic diagram of CNN-LSTM algorithm flow

In the model, the input layer is responsible for receiving data, which primarily consists of normalized values from collected historical stock price data samples. Each sample includes data from the past six trading days. The input data first undergoes convolution operations to extract high-level features, followed by pooling operations to reduce the spatial dimensions of the feature maps. A dropout layer is then applied to mitigate the risk of overfitting. Subsequently, the data goes through additional convolution and pooling operations to extract more sophisticated features. Finally, after a dimensionality transformation, the data is fed into a bidirectional LSTM (Long Short-Term Memory) layer to learn the correlation and temporal sequence information within the series. The output is then subjected to inverse normalization and passed through a fully connected layer to produce the final prediction.

4. Experiment and results

4.1. Evaluating metrics

The experiment selects five metrics, namely Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and the time spent on model training (Training runtime) to evaluate the performance of the model in stock price prediction. The specific calculation formulas are as follows:

$$MSE = \frac{1}{N}\sum_{n=1}^{N}(X_{pred} - X_{real})^2 \tag{6}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(X_{pred} - X_{real})^2} \tag{7}$$

$$MAE = \frac{1}{N}\sum_{n=1}^{N}\left|X_{pred} - X_{real}\right| \tag{8}$$

$$MAPE = \frac{1}{N}\sum_{n=1}^{N}\left|\frac{X_{pred} - X_{real}}{X_{real}}\right| \times 100\% \tag{9}$$

X_{pred} and X_{real} denote the predicted price and the actual price of the stock at time point n, respectively.

4.2. Configuration of comparative models

To better evaluate the predictive performance of the CNN-LSTM model, the paper constructs and predicts using single-layer and double-layer LSTM models in both unidirectional and bidirectional configurations. The predictive outcomes from these different setups are compared, with all models sharing the same hyperparameter settings.

4.3. Model solving and analysis

For all combinations of models, the Adam optimizer is selected, with identical input and test values. The learning rate is set to 0.001, the dropout ratio is 0.05, the number of hidden neurons in the models is set to 16, and the number of epochs is set to 200. Each model is trained and tested on the training and prediction sets, respectively, with the prediction results visualized. The evaluation plots are shown in Figures 4 through 11, respectively.

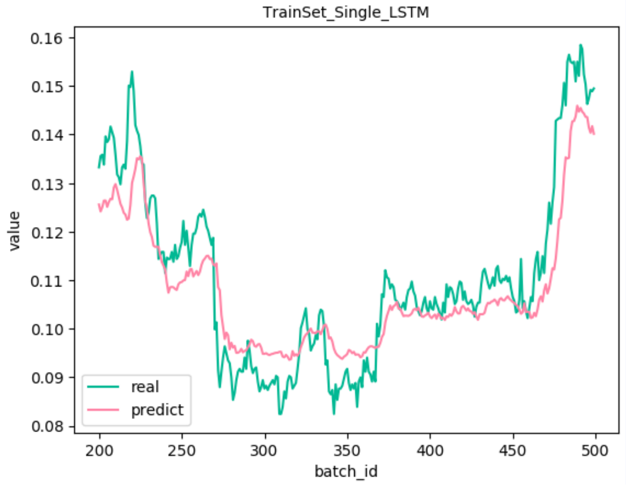

Figure 4. Train Set - Single layer Unidirectional CNN-LSTM Model Prediction

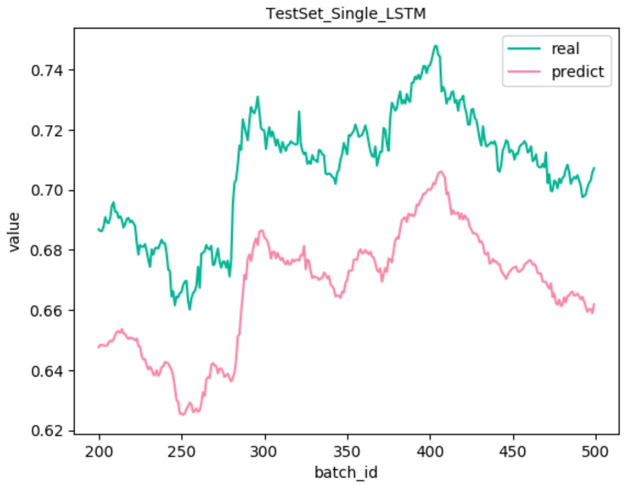

Figure 5. Test Set - Single layer Unidirectional CNN-LSTM Model Prediction

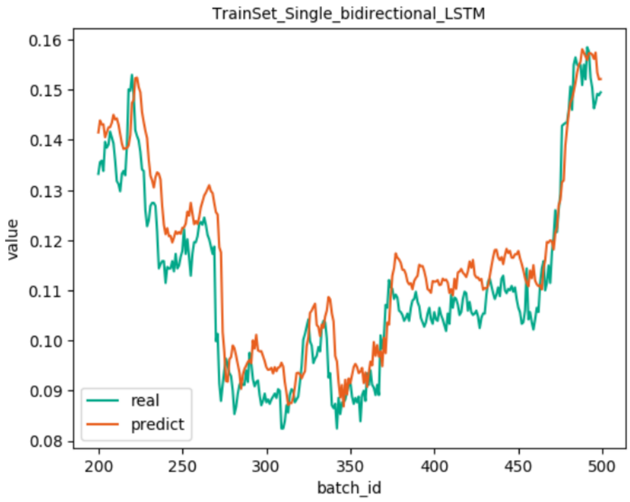

Figure 6. Train Set - Single layer Bidirectional CNN-LSTM Model Prediction

Figure 7. Test Set - Single layer Bidirectional CNN-LSTM Model Prediction

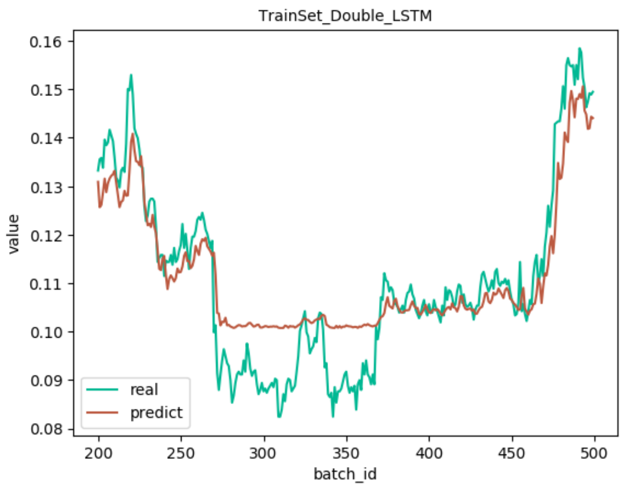

Figure 8. Train Set - Prediction of Double layered Unidirectional CNN-LSTM Model

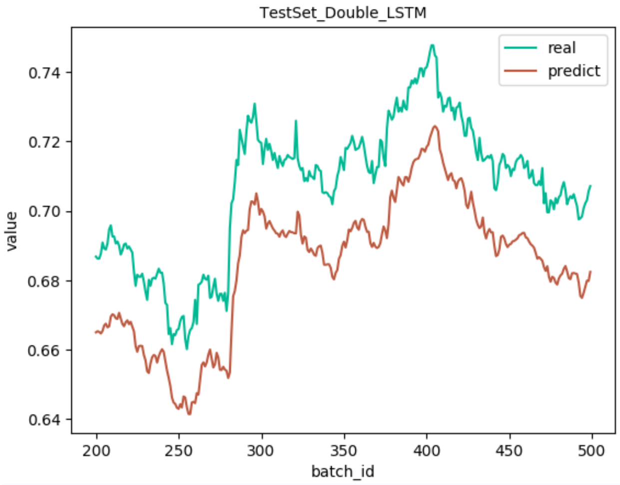

Figure 9. Test Set - Prediction of Double layered Unidirectional CNN-LSTM Model

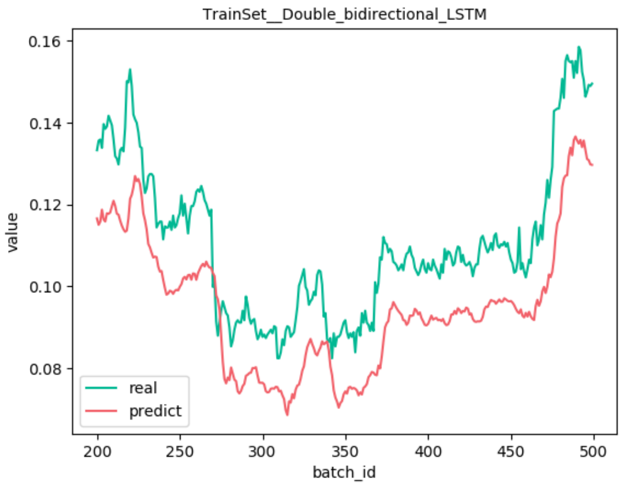

Figure 10. Train Set - Double layer Bidirectional CNN-LSTM Model Prediction

Figure 11. Test Set - Double layer Bidirectional CNN-LSTM Model Prediction

It can be observed that, on the training set, the trend of the predicted value curves for all model groups aligns with the true value curves, but only achieves a level of trend prediction rather than precise price forecasting. The predicted values are generally lower than the actual values. On the testing set, the performance of the models in terms of prediction accuracy follows the order: Double-layer Unidirectional LSTM > Single-layer Bidirectional LSTM > Single-layer Unidirectional LSTM > Double-layer Bidirectional LSTM. This suggests that for unidirectional LSTM networks, increasing the number of layers can improve the model's predictive accuracy to some extent. Conversely, for bidirectional LSTM networks, an increase in the number of layers leads to a decline in predictive capability. Under the same conditions, the best-performing model among the various combinations is the Double-layer Unidirectional LSTM network. Specific comparisons of the evaluation metrics for each model group on both the training and testing sets are presented in Table 2 below.

Table 2. Evaluation Metrics

CNN-LSTM Model	Dataset	MAE	MSE	RMSE	MAPE	Training runtime
Single-layer Unidirectional LSTM	Train Set	0.042	0.002	0.049	0.0835	326 s
	Test Set	0.062	0.003	0.063	0.0958	
Single-layer Bidirectional LSTM	Train Set	0.033	0.001	0.037	0.0709	326 s
	Test Set	0.043	0.001	0.044	0.0652	
Double-layer Unidirectional LSTM	Train Set	0.031	0.001	0.036	0.0651	340 s
	Test Set	0.044	0.002	0.044	0.0659	
Double-layer Bidirectional LSTM	Train Set	0.136	0.0255	0.159	0.328	416 s
	Test Set	0.205	0.043	0.207	0.402	

For the three initial network models, the evaluation metrics on the testing set compared to the training set differ by less than 0.05, indicating strong generalization ability on the prediction set. Specifically, for the Double-layer Unidirectional LSTM network, across the entire dataset, the model achieves a Mean Squared Error (MSE) < 0.005, Root Mean Squared Error (RMSE) < 0.05, Mean Absolute Error (MAE) < 0.05, and Mean Absolute Percentage Error (MAPE) < 0.1, demonstrating high prediction accuracy and excellent predictive performance.

4.4. Experimental Conclusion

Based on the CNN-LSTM composite model, nearly 30,000 data points of USDCNY exchange rates were used for stock price prediction. From the relevant evaluation metrics, it can be concluded that the model exhibits good generalization capabilities. This study, building upon the traditional single-layer LSTM model, proposed a hybrid model combining CNN with LSTM and conducted a comparative evaluation of various LSTM configurations. The best-performing model for stock price prediction was found to be the Double-layer Unidirectional CNN-LSTM hybrid model. This model leverages the spatial feature extraction capabilities of CNN and the temporal analysis advantages of LSTM. Through experimental validation, it was discovered that the model demonstrated good fitting and predictive performance on both the training and testing sets, particularly on the testing set, where the model showed excellent adaptability and accuracy, maintaining overall prediction errors at a low level.

5. Conclusion

This paper primarily employs a composite neural network model combining CNN and LSTM networks to predict stock prices, dropout mechanisms are introduced between the convolutional and LSTM layers. Additionally, the paper explores the combination of LSTM layer depth and bi-directionality. Through experimental comparisons of single-layer and double-layer combinations, as well as single-directional and bi-directional LSTM-CNN combinations, the prediction accuracy and evaluation indicators of the four models are evaluated. It is found that compared with the traditional single-layer unidirectional LSTM neural network, the double-layer bi-directional LSTM-CNN model shows degradation in prediction performance, while the single-layer bi-directional CNN-LSTM model and the double-layer unidirectional CNN-LSTM model both have better generalization ability and prediction accuracy in stock price prediction. Experimental results indicate that among the four model combinations, the single-layer bi-directional CNN-LSTM neural network and the double-layer unidirectional CNN-LSTM neural network outperform other models in various evaluation indicators, with very accurate prediction performance. However, from the perspective of running efficiency, the single-layer bi-directional CNN-LSTM model performs better. In summary, the application of the single-layer bi-directional CNN-LSTM model in stock price prediction can help optimize the prediction performance of the single LSTM neural network model, and it is more practical than the double-layer LSTM model in terms of running efficiency, which has practical value in real financial market analysis.

References

[1] Zhang Sulin, Song Jianan. Stock Price Prediction Based on the CNN-LSTM Model [J]. Cooperative Economics and Technology, 2023, (03): 58-62.

[2] Zhu Zijing, He Liwen. Time Series Decomposition and Forecasting Model Based on ABC-LSTM-GRU [J]. Software Engineering, 2024, 27 (03): 58-62.

[3] Xiao Tiantian. Security Stock Price Prediction Based on the K-means-LSTM Model [J]. Science Technology and Industry, 2024, 24 (03): 210-215.

[4] Tang, P., Tang, C., & Wang, K. Stock movement prediction: A multi-input LSTM approach [J]. Journal of Forecasting, 2024: 1-13.

[5] Ma Yuanyuan, Liu Yanze, Liu Chenglong, & Zhang Tianjie. (2023). A multi-perspective sentiment analysis of Chinese investors and its role in stock market prediction. Cooperative Economy & Science & Technology, (3), 58-62.

[6] Wang, J., Zhou, S., Liu, W., & Jiang, L. An ensemble model for stock index prediction based on media attention and emotional causal inference [J]. Journal of Forecasting, 2024: 1–23.

[7] Chen Menglong, Fan Cheng, Wu Zhipeng. Research on the Prediction of Stock Prices in the New Energy Vehicle Industry Based on Machine Learning Algorithms [J]. Journal of Jilin University of Commerce, 2024, 40 (01): 93-100.

[8] Cui Ting, Huang Feiran. Stock Prediction Based on Sentiment Analysis Large Model: A Predictive Model Combining GRU and ALBERT [J]. Dongyue Forum, 2024, 45 (02): 113-123.

[9] Cao Yugui, Xie Mengxing. Stock Price Prediction Analysis Based on the WD-CNN-LSTM Model [J]. Journal of North China University of Water Resources and Electric Power (Social Science Edition), 2023, 39 (05): 15-22.

[10] LI H, HUA J J, ZOU B R. Stock price prediction based on RF-LSTM combination model [J]. Journal of Henan Polytechnic University (Natural Science), 2022, 41 (1): 136-142.

[11] Sun Chenyu, Zhang Shudong. Stock Market Prediction Model Based on CEEMDAN and LSTM-Attention [J]. Computer Applications and Software, 2023, 40 (12): 119-125+146.