IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received April 28, 2021, accepted May 11, 2021, date of publication May 14, 2021, date of current version May 24, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3080459

A Prediction Model of CNN-TLSTM for USD/CNY Exchange Rate Prediction

JINGYANG WANG[1], XIAOXIAO WANG[1], JIAZHENG LI[1], AND HAIYAO WANG[2]
[1]School of Information Science and Engineering, Hebei University of Science and Technology, Shijiazhuang 050018, China
[2]School of Mechatronics, Xiamen Ocean Vocational College, Xiamen 361100, China

Corresponding author: Haiyao Wang (haiyaowang_xmoc@163.com)

This work was supported in part by the Innovation Foundation for Postgraduate of Hebei Province under Grant CXZZSS2021104, in part by the Natural Science Foundation of Hebei Province under Grant ZD2018236, and in part by the Foundation of Hebei University of Science and Technology under Grant 2019-ZDB02.

ABSTRACT The Exchange rate affects the economic development of various countries. To grasp the information of the changeable exchange rate in time, it is necessary to predict the exchange rate price. This paper proposes the CNN-TLSTM model to predict the United States Dollar/Chinese yuan (USD/CNY) exchange rate closing price of the next trading day. The model consists of two parts, namely convolutional neural networks(CNN)and tanh long short-term memory (TLSTM). The function of CNN is to extract feature factors from the input data. TLSTM is used to receive the output data of CNN for prediction, and finally obtain the prediction result. TLSTM is a new model proposed in this paper to improve the internal structure of the long short-term memory (LSTM). Its advantage is to change the range of the output value of the input gate, retain more data features, and prevent the output value of the input gate in LSTM from being overfitting. This paper selects the exchange rate of USD/CNY data and some stock data for each trading day from January 2, 2006, to October 30, 2020, as the experimental data. To prove the effectiveness of the CNN-TLSTM prediction model, the model is compared with multilayer perceptron (MLP), CNN, recurrent neural network (RNN), LSTM, and CNN-LSTM models. mean absolute percentage error (MAPE), mean square error (MSE), and R-squared (R^2) are used for comparative analysis. The experimental results show that the CNN-TLSTM model has the best predictive effect on the USD/CNY exchange rate closing price of the next trading day.

INDEX TERMS CNN, TLSTM, USD/CNY, exchange rate, prediction.

I. INTRODUCTION

The exchange rate issue has always been a hot topic in international financial research. As an important role in the economic market, the exchange rate has a certain influence on all major markets [1]. The foreign exchange market is a core component of every country's financial market, and exchange rate stability is crucial to macroeconomic stability [2], [3]. In economic life, the exchange rate is not only closely related to the country's macro-control and economic policies but also closely linked to the stability of the international economy [4]. Especially for emerging market countries, large exchange rate fluctuations in a short period will have serious consequences for the financial market and even the macro-economics of the entire country [5], [6]. The gradual acceleration of CNY internationalization also further increases the frequency and range of CNY exchange rate fluctuations. Therefore, in-depth research on the behavioral characteristics of the USD/CNY exchange rate to improve the accuracy of exchange rate prediction can help export companies and individual investors avoid foreign exchange risks. Exchange rate prediction also provides a reference for the central bank and related management departments [7].

Exchange rate fluctuations are affected by many factors. There is a non-linear relationship among the influencing factors, which complicates the problem of predicting the exchange rate closing price of the next trading day [8]. Therefore, accurately predict the closing price of the exchange rate requires many considerations, such as many influencing factors, complex non-linear relationships, and the data characteristics of the exchange rate itself [9]. This paper proposes a prediction model for the USD/CNY exchange rate based on CNN-TLSTM, which uses the exchange rate of USD/CNY data and some stock data to predict the USD/CNY exchange

The associate editor coordinating the review of this manuscript and approving it for publication was Hiu Yung Wong.

rate closing price of the next trading day. First, the CNN extracts the characteristics of the opening price of USD/CNY exchange rate, the highest price of USD/CNY exchange rate, the lowest price of USD/CNY exchange rate, the closing price of USD/CNY exchange rate, the closing price of Nasdaq Index, the closing price of Dow Jones Industrial Average, the closing price of Shanghai Composite Index and the closing price of Hang Seng Index. Then, TLSTM is used to calculate the time series data consisting of these data to get the closing price of the next trading day. This paper selects daily transaction data from January 2, 2006, to October 30, 2020, as the experimental data. At the same time, the CNN-TLSTM model is experimentally compared with five machine learning models of MLP, CNN, RNN, LSTM, and CNN-LSTM. The results show that the prediction accuracy of the CNN-TLSTM model is the highest. CNN-TLSTM shows important application value in the prediction exchange rate, lays a foundation for subsequent investment strategies. The innovations and contributions of this paper mainly include the following:

1) Based on the research of LSTM, the internal structure of the LSTM model is improved, and a new model called the TLSTM model is proposed. TLSTM model is to introduce a 1-tanh function after the input gate of the LSTM model. The range of output value of the input gate is changed to preserve as much important features of the input data as possible.
2) According to the time series and correlation of the exchange rate of USD/CNY data and stock data. The CNN-TLSTM model is proposed to predict the USD/CNY exchange rate closing price of the next trading day.
3) The CNN-TLSTM model proposed in this paper is compared with five prediction exchange rate models. The experimental results show that the accuracy and efficiency of the CNN-TLSTM model are more suitable for the prediction exchange rate.

The rest of this paper is organized as follows: Section II. presents the research of time series prediction in recent years; Section III. describes the principle of CNN-TLSTM proposed in this paper; Section IV. introduces the experimental environment, dataset, data preprocessing, experimental parameter settings, and experimental results and analyzed; Section V. summarizes the work of this paper.

II. LITERATURE REVIEW

Because of the importance of exchange rate prediction, many scholars have focused on the study of exchange rate prediction. In the past few years, different techniques have been proposed to predict the foreign exchange market [10]. Because the CNY exchange rate fluctuation is affected by many factors, and there is a nonlinear and complex relationship between these influencing factors, it is difficult for traditional statistical methods to reveal their inherent laws. With the continuous development of data mining technology and deep learning theory, machine learning algorithms have achieved excellent performance in the financial field [11], [12]. In recent years, a large number of scholars at home and abroad have used deep learning methods for financial prediction research. The deep learning method is proved to be more effective than traditional linear models (autoregressive integrated moving average (ARIMA), autoregressive moving average (ARMA), *et al.*) and nonlinear models (support vector regression (SVR), MLP, *et al.*). In 2013, Devi used the ARIMA model to predict the Niftymidcap50 index, the results showed that the prediction error of the ARIMA model was within the acceptable range, but there were better methods to be discovered [13]. In 2014, Rout thought that when using the ARMA model to study exchange rate series, the original exchange rate data needs to be trained by differential evolution algorithm first. This could eliminate the redundancy of the original data, thereby improving the prediction performance of the ARMA model [14]. In 2015, Nayak used MLP to predict the stock market index, but the effect was mediocre [15]. In 2016, Persio used MLP, CNN, and LSTM to predict the closing price of the Standard & Poor's 500 Index. The results showed neural network can predict financial time series data [16]. In 2017, Qonita used ARIMA to predict the exchange rate of the rupiah against USD, and the results showed that the ARIMA model was a feasible method to predict the rupiah against USD [17]. In 2018, Khashei used ARIMA/MLP model in the comparative study of stock price prediction, believed that ARIMA/MLP model cloud obtain more accurate results overall [18]. In 2019, Fu used SVR to predict the CNY exchange rate. The results showed that the proposed evolutionary SVR was a promising method for predicting the CNY exchange rate [19]. In 2019, Zhelev proposed to apply LSTM to financial market time series [20]. In 2019, Liao used RNN, SVM, and other machine learning techniques to predict time series data in the financial sector. He predicted the USD/ CNY exchange rate as an example, and the results show that RNN has the best results [21]. In 2020, Li used the BP neural network to predict the exchange rate trend of the USD/CNY. The prediction results showed that the USD/CNY index showed an upward trend in the short term [22]. In 2020, Islam used the GRU-LSTM hybrid network to predict exchange rate, compared it with LSTM, gate recurrent unit (GRU), and simple moving average (SMA). This model was proved to be superior to other models [23]. In 2021, Hu used CNN, LSTM, RNN, and other deep learning prediction models to study the stock and foreign exchange market prediction. Through MAPE, mean absolute error (MAE), and other performance indicators to show the results, it is concluded that the research on the financial field based on the deep learning method is gradually increasing [24].

III. MODEL

A. CNN-TLSTM

To better predict the USD/CNY exchange rate closing price of the next trading day, this paper builds a CNN-TLSTM

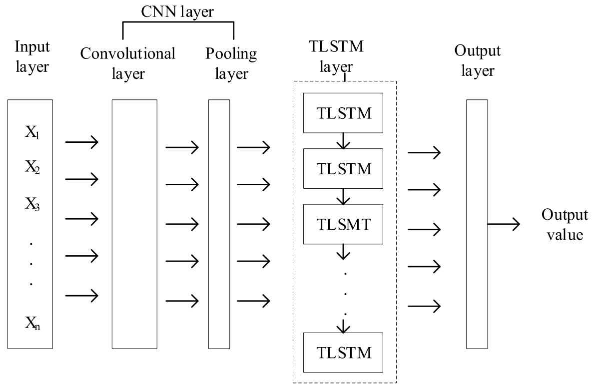


FIGURE 1. **CNN-TLSTM model structure diagram.**

prediction model. The structure diagram of the CNN-TLSTM model is shown in Fig. 1.

The model is divided into four layers: input layer, CNN layer (convolution layer and pooling layer), TLSTM layer, and output layer. The input layer is used to input the relevant data for the prediction of the exchange rate of USD/CNY. The convolutional layer in the CNN layer is used for feature extraction of input data, and the pooling layer reduces the dimension of the features extracted by the convolutional layer. The TLSTM layer performs time series prediction on the input data after feature extraction. The output layer outputs the output value of the TLSTM layer and completes the prediction of the USD/CNY exchange rate closing price of the next trading day.

B. CNN

CNN is a network model proposed by LECUN in 1998 [25]. In recent years, CNN has been widely used, which can efficiently and accurately complete prediction tasks. In the prediction of the artificial neural network, feature extraction requires humans to define features. This method of extracting features cannot effectively represent features. CNN is the first to do image recognition and classification. It has good performance in feature extraction, which can effectively solve the shortcomings of artificial neural network in feature extraction.

The results of feature extraction of data by CNN directly affect the accuracy of the predicted values. CNN is an efficient deep learning model. Its hierarchical structure is mainly composed of the input layer, convolutional layer, pooling layer, and fully connected layer. Each layer plays a very important role [26] [27]. The convolutional layer is the core of CNN. The convolution operation is mainly used for feature extraction, and the convolutional layer is used for convolution operation [28]. The pooling layer is generally after the convolutional layer because after the convolution is done, the data dimension is still large when the convolution kernel is small. To reduce the dimension of data, a pooling operation is carried out to effectively reduce the network parameters and prevent overfitting. The fully connected layer functions as a ‘‘classifier’’ in the entire convolutional neural network.

C. LSTM

LSTM is a special RNN. With the increase of training time and the number of network layers, the original RNN has a short-term memory. LSTM uses gates to control the memory process, which can effectively solve the problem of gradient disappearance and gradient explosion during long sequence training [29], [30]. After experimental verification, it is found that LSTM is good at dealing with multi-variable or multi-input problems. The characteristics of LSTM help to solve the time series prediction problem.

The LSTM architecture is shown in Fig. 2.

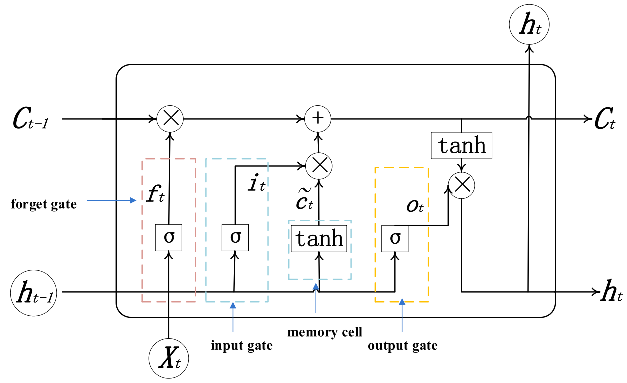


FIGURE 2. **LSTM architecture diagram.**

LSTM mainly includes three gates (the forget gate, the input gate, the output gate) and the memory cell [31]. The memory cell can control the transfer of information to the next moment. The forget gate is mainly to selectively forget the input data from the previous node [32]. The input gate is a selective memory of the input data at this stage, with more records for the important part and fewer records for the non-important part. The output gate determines what is output as the current state [33].

D. TLSTM

TLSTM is a new model proposed in this paper to improve the internal structure of the LSTM model. The TLSTM is designed to prevent the output value of the input gate in LSTM from overfitting. TLSTM introduces the 1-tanh function after the input gate of LSTM to change the value range of the output value of the input gate to ensure that the important features of the input data are retained as much as possible. The structure of TLSTM is shown in Fig. 3. When the input gate of LSTM is activated using the Sigmoid function, the output values close to 0 are discarded, and the output values close to 1 are completely preserved and transferred. The output range of the input gate is [0, 1]. When the 1-tanh function is introduced, the output value close to 1 will become close to 0.25, the output value close to 0 will become 1, and the output value close to the middle of 0-1 will become 0.5 on both sides. The output value range of the gate becomes [0.25, 1]. Transforming the data into more distinct intervals helps capture the correlation between time series data.

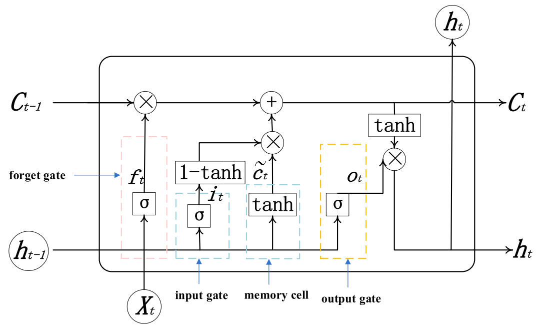


FIGURE 3. TLSTM architecture diagram.

TABLE 1. Experimental software and hardware environment.

Experimental Environment	Environmental Constraints
Hardware Environment	Processor: i5-1135G7 Radeon R4,8Compute Cores 4.2GHz Memory: 16GB Hard Disk Capacity: 512G Graphics Card: NVIDIA MX450
Software Environments	Operating System: Windows 10.0 64bit Python: 3.7.6 Keras: 2.1.0 TensorFlow: 1.14.0

The calculation formulas for TLSTM are as follows:

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

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

$$o_t = \sigma\left(W_o\, [h_{t-1}, x_t] + b_o\right) \tag{3}$$

$$\tilde{C}_t = tanh\left(W_c \cdot [h_{t-1}, x_t] + b_c\right) \tag{4}$$

$$t_t = 1 - tanh\,(i_t) \tag{5}$$

$$C_t = f_t * C_{t-1} + t_t * \tilde{C}_t \tag{6}$$

$$h_t = o_t * tanh\,(C_t) \tag{7}$$

where W_f is the weight of forget gate, b_f is the bias of forget gate, W_i is the weight of input gate, b_i is the bias of input gate, W_c is the weight of candidate cell, b_c is the bias of candidate memory cell, W_o is the weight of output gate, b_o is the bias of output gate.

The calculation processes of TLSTM as follow: Firstly, the output data of the previous TLSTM unit and the input data of the current moment enter the forget gate, the input gate, the output gate, and the candidate memory cell, the corresponding output value is obtained. Secondly, according to the output value of the input gate, the 1-tanh function is used to transform the output value of the input gate. Thirdly, according to the output value of the forget gate, the transformed output value of the input gate, the candidate memory cell state, and the memory cell state of the previous TLSTM, the current memory cell state is calculated. Finally, the output value of the TLSTM is calculated based on the current memory cell state and the output value of the output gate.

IV. EXPERIMENT

To prove the effectiveness of the CNN-TLSTM prediction model, it is also compared with five machine learning models: MLP, CNN, RNN, LSTM, and CNN-LSTM.

A. EXPERIMENTAL ENVIRONMENT

All model experiments in this paper are carried out in the same experimental environment, the hardware and software environments are shown in Table 1.

B. DATASET

The data used in this paper are all from the Wind database. The daily trading data of USD/CNY exchange rate, Nasdaq Index, Dow Jones Industrial Average, Shanghai Composite Index, and Hang Seng Index from January 2, 2006, to October 30, 2020, are used as experimental data. Where, the data of USD/CNY exchange rate includes opening price, highest price, lowest price, and closing price. This paper uses only the daily closing prices of the Nasdaq Index, Dow-Jones Industrial Average, Shanghai Composite Index, and Hang Seng Index as data content. Partial sample data are shown in Table 2.

Where, Open represents the opening price of the USD/CNY exchange rate (Yuan), High represents the highest price of the USD/CNY exchange rate (Yuan), Low represents the lowest price of the USD/CNY exchange rate (Yuan), Close represents the closing price of the USD/CNY exchange rate (Yuan), Sh Represents the closing price of the Shanghai Composite Index (Yuan), Ixic represents the closing price of the Nasdaq Index (Yuan), His represents the closing price of the Hang Seng Index (Yuan), and Dji represents the closing price of the Dow Jones Industrial Average (Yuan).

The data from January 2, 2006, to December 31, 2019, are selected as the training set. The data from January 1, 2020, to October 30, 2020, are selected as the test set.

C. DATA PREPROCESSING

The non-stationary data is not suitable for fitting the neural network model. In 2020, Livieris used Augmented Dickey–Fuller (ADF) test to test the stationarity of the original training data [34]. So ADF test is used to test the stationarity of the original training data in this paper. Table 3 shows the results of the original training data ADF test. By considering the t-statistics (t-stat) and the associated p values. The null hypothesis H_0 "the levels possess a unit root and are non-stationary" is accepted. Therefore, the original data is non-stationary.

To make the data stationary, it is suitable for fitting neural network. This paper uses a feasible region to preprocess original training data. The feasible region calculation formula is as follows:

$$y_n = x_n - x_{n-1} \tag{8}$$

where y_n is the value after the difference, x_n is the original value, x_{n-1} is the previous data.

TABLE 2. Partial sample data.

Date	Open	High	Low	Close	Sh	Ixic	Hsi	Dji
2006/01/04	8.0662	8.067	8.0655	8.067	1180.963	2263.46	15200.06	10880.15
2006/01/05	8.063	8.0654	8.063	8.0637	1197.269	2276.87	15271.13	10882.15
2006/01/06	8.0648	8.0648	8.063	8.063	1209.422	2305.62	15344.44	10959.31
2006/01/09	8.0639	8.0639	8.062	8.0621	1215.668	2318.69	15547.43	11011.9
2006/01/10	8.0674	8.0674	8.062	8.062	1220.618	2320.32	15569.91	11011.58

TABLE 3. ADF test of the original training data.

Series	Open	High	Low	Close	Sh	Ixic	Hsi	Dji
t stat.	-1.0534	-1.1171	-1.1171	-1.0726	-1.0775	3.0413	-1.5536	0.9340
p value	0.7333	0.7082	0.7082	0.7259	0.7240	1.0000	0.5069	0.9935

TABLE 4. ADF test after data preprocessing.

Series	Open	High	Low	Close	Sh	Ixic	Hsi	Dji
t stat.	-12.9752	-10.6443	-10.9645	-11.9556	-12.3046	-12.6582	-12.4792	-13.9898
p value	4.7847e-24	4.8408e-19	8.1687e-20	4.2177e-22	7.3214e-23	1.3225e-23	3.1194e-23	4.0363e-26

The results of the ADF test after preprocessing are shown in Table 4. The null hypothesis H_0 is rejected. Therefore, the preprocessed data is stationary, and it is suitable for neural network training.

D. EXPERIMENTAL PARAMETER SETTINGS

The optimal parameters for each layer of the CNN-TLSTM model are shown in Table 5. The main parameters of the convolution layer include filter, kernel_size, activation function, padding. The main parameters of the pooling layer are pool size and padding. The main parameters of TLSTM are the number of hidden units and the activation function. The time_step, optimizer, learning rate, loss function, epochs, and batch_size are parameters for training found by the grid search method.

TABLE 5. Optimal parameters table for each layer of CNN-TLSTM model.

Parameters	Value
Convolution Layer Filters	60
Convolution Layer Kernel_size	2
Convolution Layer Activation Function	Sigmoid
Convolution Layer Padding	Valid
Pooling Layer Pool_size	1
Pooling Layer Padding	Valid
Number of TLSTM Layer Hidden Units	60
TLSTM Layer Activation Function	Tanh
Time_step	5
Optimizer	Adam
Learning Rate	0.001
Loss Function	MAE
Epochs	50
Batch_size	60

The parameter settings are explained as follows:

The filters of the convolution layer are the dimensionality of the output space. Kernel_size of convolution layer is the size of the convolution kernel when CNN convolutes. The activation function of the convolution layer is used to introduce non-linear factors. Common activation functions are Relu, tanh, Sigmoid. Padding of the convolution layer is the edge filling method for convolution operation. The pool size of the pooling layer is the size of the max pooling window. Padding of the pooling layer is the edge filling method for pooling operation. The number of TLSTM layer hidden units is the dimensionality of the output space. The activation function of the TLSTM layer is used to introduce non-linear factors. Common activation functions are Relu, tanh, Sigmoid. The time step is the length of the time series. Optimizer is used to update parameters and optimize objective functions. The learning rate determines whether the neural network can converge to the global minimum. The loss function is obtained in the forward propagation calculation, and it is also the starting point of reverse propagation. Epochs is the number of times required for training. Batch_size is used to define the number of samples to be processed before updating internal model parameters.

The model structure diagram of CNN-TLSTM after setting the parameters is shown in Fig. 4. Firstly, the training set data are input into the model. Secondly, the convolution layer extracts the features of the data. Through the pooling layer, the output data are obtained. Thirdly, the output data enter the TLSTM layer for training. Finally, the trained output data enter the output layer to obtain the output value.

E. EXPERIMENTAL RESULTS AND ANALYSIS

To prove the high performance and high accuracy of CNN-TLSTM, MLP, CNN, RNN, LSTM, CNN-LSTM and CNN-TLSTM are respectively used to predict the USD/CNY exchange rate closing price of the next trading day under the

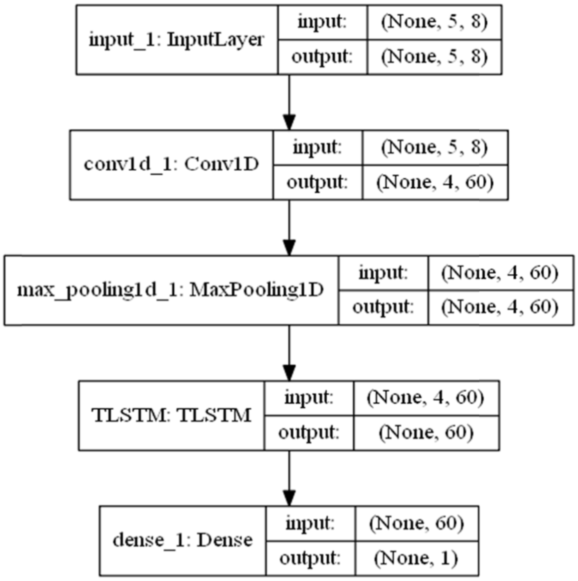


FIGURE 4. **CNN-TLSTM model structure diagram.**

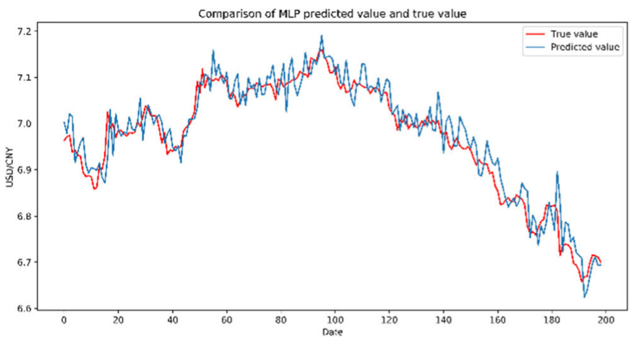


FIGURE 5. **Comparison of MLP predicted value and true value.**

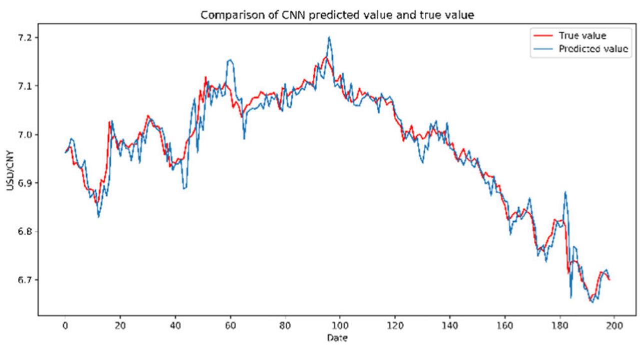


FIGURE 6. **Comparison of CNN predicted value and true value.**

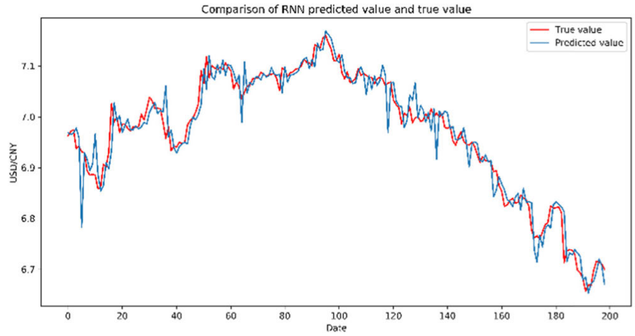


FIGURE 7. **Comparison of RNN predicted value and true value.**

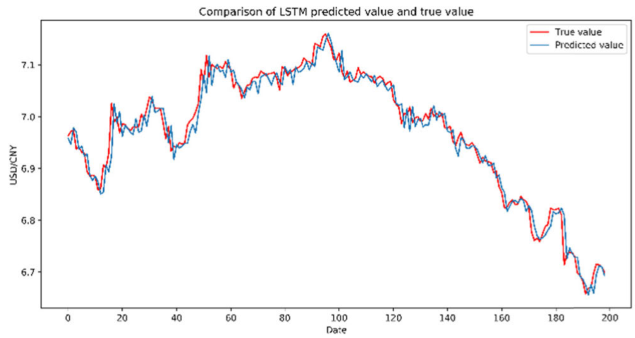


FIGURE 8. **Comparison of LSTM predicted value and true value.**

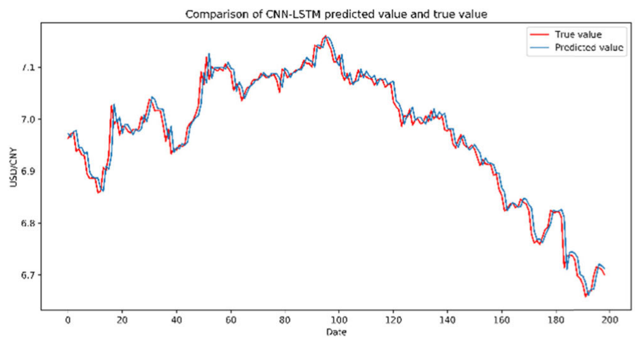


FIGURE 9. **Comparison of CNN-LSTM predicted value and true value.**

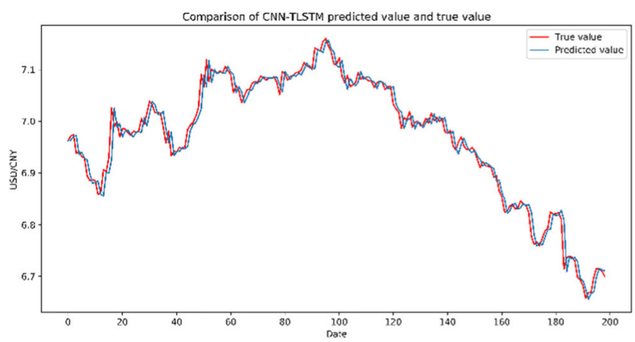


FIGURE 10. **Comparison of CNN-TLSTM predicted value and true value.**

same operating environment, the same training set and the same test set. Firstly, the six models are trained 50 times with the divided training set data, and the optimal training model is saved. Then the optimal model is used to predict the test set. Fig. 5-10 show the comparison between the true value and the predicted value of the six models.

From the comparison figure of the six models, it can be seen that the CNN-TLSTM model has the highest degree of fitting between the true value and the predicted value, and the MLP model has the lowest degree of fitting. On the whole, the fitting degree of the CNN model is better than that of the MLP model. Compared with the RNN model, the CNN model not fitting degree well in the middle part. Compared with the LSTM model, the fitting degree of the RNN model in individual values is worse than the LSTM model, and the LSTM model is slightly better. Compared CNN-LSTM model with the LSTM model, the CNN-LSTM model has a better fitting degree in the middle part. Compared with the CNN-TLSTM model, the fitting degree of the CNN-TLSTM model near the peak is better than that of the CNN-LSTM model. The true value of the CNN-TLSTM model almost coincides with the predicted value. Therefore, the overall prediction effect of the CNN-TLSTM model is the best.

To compare the prediction performance of models more clearly, MAPE, MSE, R^2 are used to evaluate and compare the prediction results. The comparison results of the six models' evaluation index are shown in Table 6. The comparison of the six models of MAPE is shown in Fig. 10, the comparison of MSE is shown in Fig. 11, and the comparison of R^2 is shown in Fig. 12.

TABLE 6. Comparison of five methods evaluation indexes.

Model	MAPE	MSE	R^2
MLP	0.40141	0.00125	0.96543
CNN	0.30956	0.00090	0.97527
RNN	0.29090	0.00084	0.97693
LSTM	0.22448	0.00047	0.98694
CNN-LSTM	0.20728	0.00042	0.98845
CNN-TLSTM	0.18945	0.00038	0.98950

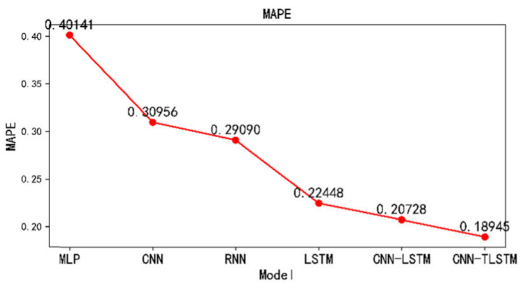


FIGURE 11. MAPE comparison diagram of six prediction models.

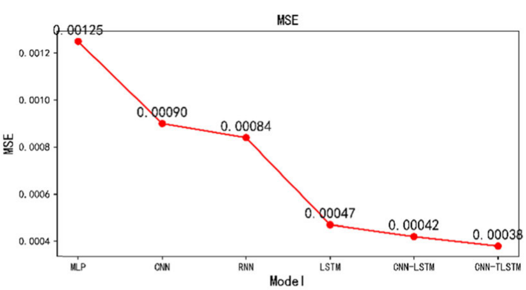


FIGURE 12. MSE comparison diagram of six prediction models.

The MAPE calculation formula is as follows:

$$MAPE = \frac{100\%}{n} \sum_{i=1}^{n} \left| \frac{\hat{y}_i - y_i}{y_i} \right| \tag{9}$$

where $\hat{y}_i$ is the predicted value and y_i is the true value. The value range of MAPE is $[0, +\infty]$. The smaller the value of MAPE, the better accuracy of the prediction model.

The MSE calculation formula is as follows:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2 \tag{10}$$

where $\hat{y}_i$ is the predicted value and y_i is the true value. The closer the MSE is to 0, the better the model prediction.

The R^2 calculation formula is as follows:

$$R^2 = 1 - \frac{\left(\sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2\right)/n}{\left(\sum_{i=1}^{n} \left(\bar{y}_i - \hat{y}_i\right)^2\right)/n} \tag{11}$$

where $\hat{y}_i$ is the predicted value, y_i is the true value, and $\bar{y}_i$ is the average value. The value range of R^2 is 0 to 1. The closer it is to 1, the better the performance.

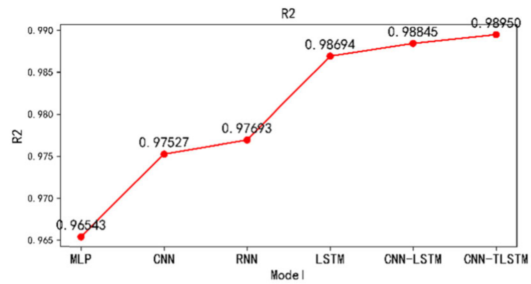


FIGURE 13. R^2 comparison diagram of six prediction models.

From Table 4 and Fig. 11-13, the following conclusions can be drawn:

Both MAPE and MSE of CNN-TLSTM are the smallest, R^2 is nearest to 1., MAPE and MSE of MLP are both the largest, and R^2 is the smallest. The order of MAPE from high to low is MLP, CNN, RNN, LSTM, CNN-LSTM, CNN-TLSTM. The order of MSE from high to low is MLP, CNN, RNN, LSTM, CNN-LSTM, CNN-TLSTM. The order of R^2 from high to low is CNN-TLSTM, CNN-LSTM, LSTM, RNN, CNN, MLP. MLP has the least predictive accuracy because MLP is shallow learning. Compared with CNN, both MAPE and MSE of RNN decreases, R^2 increases, MAPE decreases from 0.30956 to 0.29090, MSE decreases from 0.00090 to 0.00084, R^2 increases from 0.97527 to 0.97693. It shows that RNN has better prediction accuracy than CNN. Because CNN are mainly used for image recognition and classification. Compared with RNN, LSTM decreases MAPE by 22.83%, MSE decreases by 44.05%, R^2 increases by 1.02%. The results show that LSTM has a higher prediction accuracy than RNN, can effectively avoid the gradient disappearance and gradient explosion caused by RNN, more suitable for time series prediction. Compared with LSTM, the MAPE of CNN-LSTM decreases from 0.22448 to 0.20728, decreases by 7.66%, MSE decreases by 10.64%

from 0.00047 to 0.00042, R^2 increases by 0.15% from 0.98694 to 0.98845. The results show that when CNN is introduced into the LSTM model for feature extraction, the accuracy of the prediction model can be improved. Compared with CNN-LSTM, the MAPE of CNN-TLSTM decreases from 0.20728 to 0.18945, MSE decreases by 9.52%, from 0.00042 to 0.00038, R^2 increases by 0.10%, from 0.98845 to 0.98950. The results show that the TLSTM model introduces the 1-tanh function after the input gate of the LSTM model, which preserves the important features of the input data and improves the prediction accuracy.

From the comparison of the three evaluation indexes, CNN-TLSTM is better than the other five models. CNN-TLSTM shows that by changing the value range of the output value of the input gate, and it is ensured that the important features of the input data are retained as much as possible so that the prediction accuracy of the CNN-TLSTM model is higher. Therefore, the CNN-TLSTM model proposed in this paper is superior to the other five comparison models in terms of both fitting degree and error. CNN-TLSTM model can well predict the USD/CNY exchange rate closing price of the next trading day.

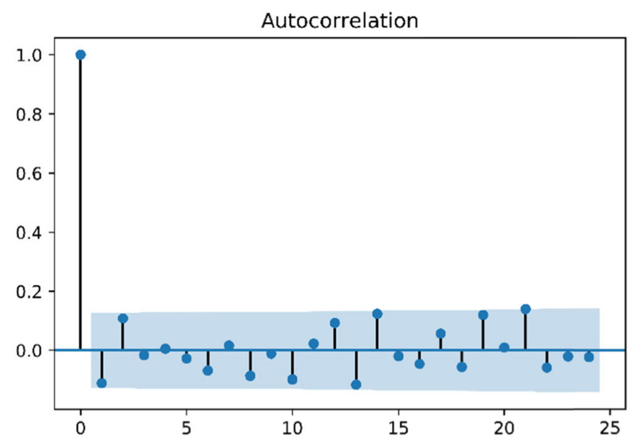


FIGURE 14. ACF plot of residuals of CNN-TLSTM.

Livieris used Auto-Correlation Function (ACF) Plot to examine for auto-correlation in the residuals. To prove the reliability of the CNN-TLSTM prediction model, Auto-Correlation Function (ACF) Plot is used to examine for auto-correlation in the residuals which are the error between the predicted value and the true value. The ACF plot is shown in Fig. 14. In the ACF plot, the confidence limits are denoted with a blue region and are constructed assuming that the residuals follow a Gaussian probability distribution. As shown in Fig.14 the residuals of CNN–TLSTM exist no autocorrelation. This implies that the model is reliable.

V. CONCLUSION

According to the research status of foreign exchange and financial time series data at home and abroad. MLP, CNN, and LSTM models have some disadvantages, such as low prediction accuracy and single model structure. Therefore, the CNN-TLSTM model is proposed. The model includes two parts, namely CNN and TLSTM. The function of CNN is to extract feature factors of the input data, and TLSTM is used to receive the output data of CNN for prediction. TLSTM is a new model proposed in this paper to improve the internal structure of the LSTM model. The 1-tanh function is introduced after the input gate of LSTM to change the value range of the output value of the input gate. The important characteristics of the input data are retained as much as possible to prevent the output value of the input gate in LSTM from overfitting. This paper selects the opening price of USD/CNY exchange rate, the highest price of USD/CNY exchange rate, the lowest price of USD/CNY exchange rate, the closing price of USD/CNY exchange rate, the closing price of Nasdaq Index, the closing price of Dow Jones Industrial Average, the closing price of Shanghai Composite Index, and the closing price of Hang Seng Index from January 2, 2006 to October 30, 2020, are used as experimental data to predict the USD/CNY exchange rate closing price of the next trading day. CNN-TLSTM model compares with MLP, RNN, CNN, LSTM, and CNN-LSTM model. The experimental results show that the CNN-TLSTM model has good predictive ability for complex nonlinear problems, and can predict the USD/CNY exchange rate closing price of the next trading day.

The future research will mainly have two aspects:

1) In the improved scheme of LSTM, it is considered to change the output range of forget gate and output gate.
2) In the study of exchange rate prediction models, due to many factors affecting the closing price of the exchange rate, this paper only considers the influence of historical price data of USD/CNY exchange rate and some stock data. However, the global impact of COVID-19 and related national policies have not become input variables.

REFERENCES

[1] M. Liao, "RMB exchange rate prediction based on ARIMA and BP neural network fusion model," *World Sci. Res. J.*, vol. 6, no. 4, pp. 252–262, Apr. 2020.
[2] M. M. Panda, S. N. Panda, and P. K. Pattnaik, "Exchange rate prediction using ANN and deep learning methodologies: A systematic review," in *Proc. Indo-Taiwan ICAN*, Rajpura, India, 2020, pp. 86–90.
[3] P. Arsi, O. Somantri, U. Hasanah, T. Astuti, M. Imron, and R. Waluyo, "A proposed model of neural network for rupiah exchange rate's prediction," in *Proc. ICITISEE*, Yogyakarta, Indonesia, 2018, pp. 90–93.
[4] W. Cao, W. Zhu, W. Wang, Y. Demazeau, and C. Zhang, "A deep coupled LSTM approach for USD/CNY exchange rate forecasting," *IEEE Intell. Syst.*, vol. 35, no. 2, pp. 43–53, Mar. 2020.
[5] C. Hu, K. Zheng, and L. Liu, "Layered exchange rate prediction model based on LSTM," in *Proc. ICMAI*, New York, NY, USA, 2020, pp. 124–127.
[6] M. M. Panda, S. N. Panda, and P. K. Pattnaik, "Multi currency exchange rate prediction using convolutional neural network," *Mater. Today*, vol. 43, pp. 1–12, Jan. 2021, doi: 10.1016/j.matpr.2020.11.317.
[7] M. Xu and W. Li, "Research on exchange rate forecasting model based on ARIMA model and artificial neural network model," in *Proc. 2nd Int. Conf. Mater. Sci., Machinery Energy Eng. (MSMEE)*, Dalian, China, 2017, pp. 1215–1220.
[8] S. Ye, "RMB exchange rate forecast approach based on BP neural network," *Phys. Procedia*, vol. 33, pp. 287–293, Dec. 2012.

[9] D. Zhang, X. Wang, L. Gao, and Y. Gong, "Predict and analyze exchange rate fluctuations accordingly based on quantile regression model and K-nearest neighbor," *J. Phys., Conf. Ser.*, vol. 1813, no. 1, Feb. 2021, Art. no. 012016.

[10] Y. Qu and X. Zhao, "Application of LSTM neural network in forecasting foreign exchange price," *J. Phys., Conf. Ser.*, vol. 1237, Jun. 2019, Art. no. 042036.

[11] L. Ji, Y. Zou, K. He, and B. Zhu, "Carbon futures price forecasting based with ARIMA-CNN-LSTM model," *Procedia Comput. Sci.*, vol. 162, pp. 33–38, Jan. 2019.

[12] Y. Fang, B. Guan, S. Wu, and S. Heravi, "Optimal forecast combination based on ensemble empirical mode decomposition for agricultural commodity futures prices," *J. Forecasting*, vol. 39, no. 6, pp. 877–886, Sep. 2020.

[13] B. U. Devi, D. Sundar, and P. Alli, "An effective time series analysis for stock trend prediction using ARIMA model for Nifly midcap-50," *Int. J. Data Mining Knowl. Manage. Process*, vol. 3, no. 1, pp. 65–78, Jan. 2013.

[14] M. Rout, B. Majhi, R. Majhi, and G. Panda, "Forecasting of currency exchange rates using an adaptive ARMA model with differential evolution based training," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 26, no. 1, pp. 7–18, Jan. 2014.

[15] S. C. Nayak, B. B. Misra, and H. S. Behera, "Artificial chemical reaction optimization of neural networks for efficient prediction of stock market indices," *Ain Shams Eng. J.*, vol. 8, no. 3, pp. 371–390, Sep. 2017.

[16] L. Di Persio and O. Honchar, "Artificial neural networks architectures for stock price prediction: Comparisons and applications," *Int. J. Circuits, Syst. Signal Process.*, vol. 10, pp. 403–413, Jan. 2016.

[17] A. Qonita, A. G. Pertiwi, and T. Widiyaningtyas, "Prediction of rupiah against US dollar by using ARIMA," in *Proc. EECSI*, Yogyakarta, Indonesia, 2017, pp. 1–5.

[18] M. Khashei and Z. Hajirahimi, "A comparative study of series arima/mlp hybrid models for stock price forecasting," *Commun. Statist.-Simul. Comput.*, vol. 48, no. 9, pp. 1–16, May 2018.

[19] S. Fu, Y. Li, S. Sun, and H. Li, "Evolutionary support vector machine for RMB exchange rate forecasting," *Phys. A, Stat. Mech. Appl.*, vol. 521, pp. 692–704, May 2019.

[20] S. Zhelev and D. R. Avresky, "Using LSTM neural network for time series predictions in financial markets," in *Proc. NCA*, Cambridge, MA, USA, 2019, pp. 1–5.

[21] R. Liao, P. Boonyakunakorn, N. Harnpornchai, and S. Sriboonchitta, "Forecasting the exchange rate for USD to RMB using RNN and SVM," *J. Phys., Conf. Ser.*, vol. 1616, Aug. 2020, Art. no. 012050.

[22] L. Li, F. Pan, and C. Wang, "Prediction analysis of USD-CNY trend based on BP neural network," in *Proc. ICBAIE*, Fuzhou, China, 2020, pp. 48–52.

[23] M. S. Islam and E. Hossain, "Foreign exchange currency rate prediction using a GRU-LSTM hybrid network," *Soft Comput. Lett.*, vol. 3, Oct. 2020, Art. no. 100009.

[24] Z. Hu, Y. Zhao, and M. Khushi, "A survey of forex and stock price prediction using deep learning," *Appl. Syst. Innov.*, vol. 4, no. 1, p. 9, Feb. 2021.

[25] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Dec. 1998.

[26] A. J. Balaji, D. S. H. Ram, and B. B. Nair, "Applicability of deep learning models for stock price forecasting an empirical study on BANKEX data," *Procedia Comput. Sci.*, vol. 143, pp. 947–953, Nov. 2018.

[27] L. Sayavong, Z. Wu, and S. Chalita, "Research on stock price prediction method based on convolutional neural network," in *Proc. Int. Conf. Virtual Reality Intell. Syst. (ICVRIS)*, Hunan, China, Sep. 2019, pp. 173–176.

[28] Q. Chen, W. Zhang, and Y. Lou, "Forecasting stock prices using a hybrid deep learning model integrating attention mechanism, multi-layer perceptron, and bidirectional long-short term memory neural network," *IEEE Access*, vol. 8, pp. 117365–117376, Jun. 2020.

[29] T. Ergen and S. S. Kozat, "Unsupervised anomaly detection with LSTM neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 8, pp. 3127–3141, Aug. 2020.

[30] V. Ta, C. Liu, and D. Tadesse, "Portfolio optimization-based stock prediction using long-short term memory network in quantitative trading," *Appl. Sci.*, vol. 10, no. 2, pp. 437–456, Jan. 2020.

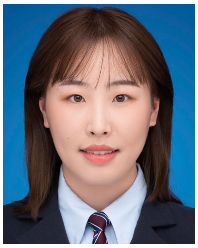
[31] D. L. Minh, A. Sadeghi-Niaraki, H. D. Huy, K. Min, and H. Moon, "Deep learning approach for short-term stock trends prediction based on two-stream gated recurrent unit network," *IEEE Access*, vol. 6, pp. 55392–55404, Sep. 2018.

[32] W. Lu, J. Li, J. Wang, and L. Qin, "A CNN-BiLSTM-AM method for stock price prediction," *Neural Comput. Appl.*, vol. 86, no. 11, pp. 1–13, Nov. 2020.

[33] I. E. Livieris, E. Pintelas, and P. Pintelas, "A CNN–LSTM model for gold price time-series forecasting," *Neural Comput. Appl.*, vol. 32, no. 23, pp. 17351–17360, Dec. 2020.

[34] I. E. Livieris, S. Stavroyiannis, E. Pintelas, and P. Pintelas, "A novel validation framework to enhance deep learning models in time-series forecasting," *Neural Comput. Appl.*, vol. 32, no. 23, pp. 17149–17167, Dec. 2020.

JINGYANG WANG received the B.Eng. degree in computer software from Lanzhou University, China, in 1995, and the M.Sc. degree in software engineering from the Beijing University of Technology, China, in 2007. He is currently a Professor with the School of Information Science and Engineering, Hebei University of Science and Technology, Shijiazhuang, Hebei, China. His research interests include machine learning, natural language processing, big data processing, and distributed computing.



XIAOXIAO WANG is currently pursuing the master's degree with the Hebei University of Science and Technology. Her research interests include machine learning and deep learning.

JIAZHENG LI is currently pursuing the master's degree with the Hebei University of Science and Technology. His research interests include machine learning and deep learning.

HAIYAO WANG received the B.Eng. degree in mechanical design and manufacturing and the M.Sc. degree in industry engineering from the Hefei University of Technology, China, in 1998 and 2009, respectively. She is currently an Associate Professor with the School of Mechatronics, Xiamen Ocean Vocational College, Xiamen, Fujian, China. Her research interests include machine learning, process optimization, and efficiency improvement.