

Research on Stock Price Prediction Based on CNN-LSTM Combined Model

Shangzhi Gao

Dalian Minzu University, Dalian, Liaoning, 116650, China

Keywords: CNN-LSTM; Deep Learning; Predictive model

Abstract: It is difficult to forecast time series data such as stock prices because of the complexity of the internal structure of the stock price system and the diversity of external factors that determine the complexity of the stock market and the uncertainty of the stock price forecasting task. In addition, the impact of different factors on stock prices may be linear or non-linear. Therefore, this paper proposes a combined CNN-LSTM model with good results and generalization ability, which can fully exploit the advantages of each model in stock price prediction and improve the accuracy and stability of prediction. This research can improve the effectiveness of stock price prediction by constructing suitable model structures, parameter configurations and optimization algorithms to provide investors and financial practitioners with a powerful reference for decision-making. An attention mechanism is also introduced in order to improve the prediction accuracy of the CNN-LSTM model, and the results confirm the efficacy of the proposed approach.

1. Introduction

With the rise of machine learning in the area of financial investment and the rise of artificial intelligence, machine learning methods have been used by many researchers to study market changes, and trends and stock price forecasts to guide stock investment. At present, many research scholars at home and abroad have proposed various forecasting methods, such as VAR (vector autoregressive model), ARM (autoregressive sliding average model), exponential smoothing model based on statistics and probability theory, GM, SVM and ANN innovative prediction models based on non-statistical principles, the gray prediction method, the artificial neural network method, and so on. And there is a growing body of research demonstrating the usefulness of artificial neural networks for predicting stock prices.

However, with the development of computer technology, some studies have shown that artificial neural networks may not be suitable for stock prediction, so many studies started to start from deep architecture, so machine learning-based techniques were gradually applied to the problem of financial time series data, and a series of research methods have been proposed by experts and scholars in this study when studying deep learning algorithms. Wei Bao et al. first proposed to extract deep stock features into stock price prediction, adding wavelet transform to the LSTM model and orthogonal decomposition of stock price features to eliminate the results[1]. Han, Shanjie et al. Proposed TensorFlow-based deep learning model for stock price prediction[2]. Ming Li establishes a double LSTM deep learning network and adds convolution

for deep feature learning, establishes a convolution-double GRU deep learning network to predict stock prices and upward and downward trends, and then compares the experiments with the logistic model to get more accurate prediction results[3]. Zhang Zeya et al. proposed a recurrent neural network model that can improve the prediction of stock ups and downs by more than 5% compared to the SVM classifier based on price features[4]. And in 2022 W Gao et al. proposed a new model based on MFNN and IBHA to model the stock market near and analyze two data sets of the Shanghai stock market from 2015.9-2018.8, which finally showed that the new method calculated results match better with the real data and have higher prediction accuracy[5].

In summary, scholars at home and abroad have already made many contributions to the issue of stock trend forecasting, most of which are based on analysis of foreign stock markets due to the late start of the domestic stock market. However, as the domestic stock market continues to develop and the financial system continues to improve, while the domestic stock market gradually rises in the financial system, forecasting studies based on the domestic stock market will become more and more important. In this paper, a series of studies will be conducted on the short-term trend forecasting of domestic individual stocks by constructing LSTM models and CNN-LSTM models.

2. Model construction

2.1 LSTM recurrent neural network

LSTM is a special type of Recurrent Neural Network (RNN) with a memory capacity for capturing long term dependencies in sequential data[6]. The structure is illustrated in Figure 1:

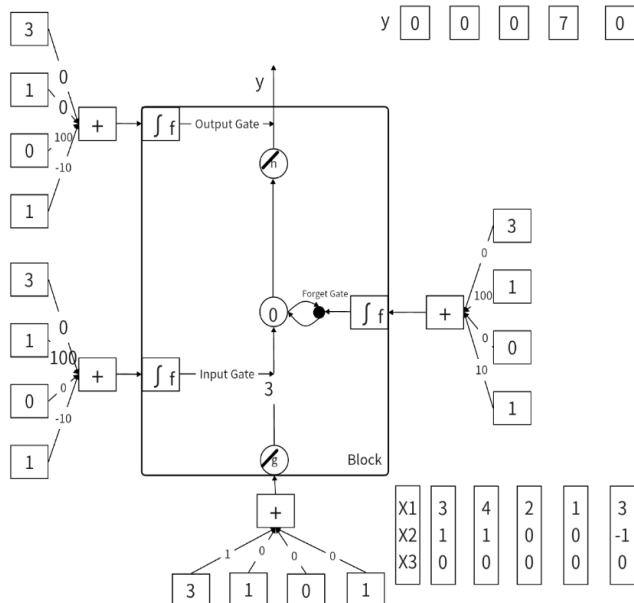


Figure 1: LSTM recurrent neural network structure

In stock price forecasting, LSTM can be used to model the time-series relationship of stock prices and thus make predictions of future price movements.

2.2 CNN Convolutional Neural Network

CNN convolutional neural networks take advantage of the fact that the input is an image and design the neurons in three dimensions: width, height, and depth (note that depth is not the depth of the neural network but is used to describe the neurons). For example, if the input image size is $32 \times 32 \times$

3 (RGB), then the input neuron will also have a dimension of $32 \times 32 \times 3$. As shown in Figure 2:

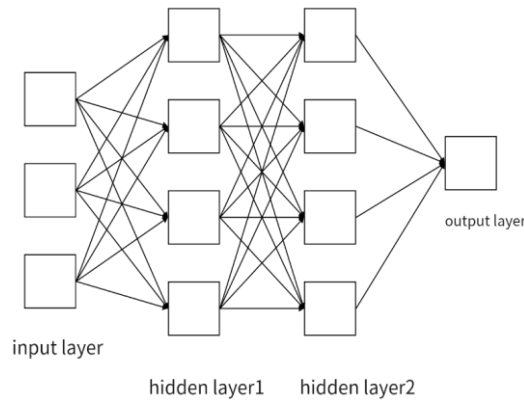


Figure 2: Structure of CNN convolutional neural network

The convolutional neural network CNN is a deep learning model with good performance in the image processing domain. The system can automatically learn features from data and is able to deal with multidimensional data such as images and time series data. CNNs can be used in stock price prediction to extract time series characteristics of stock prices such as price trends and fluctuations.

2.3 CNN-LSTM based stock price prediction model

The CNN-LSTM model combines the benefits of both CNN and LSTM, and the CNN-LSTM network model can be seen in Figure 3:

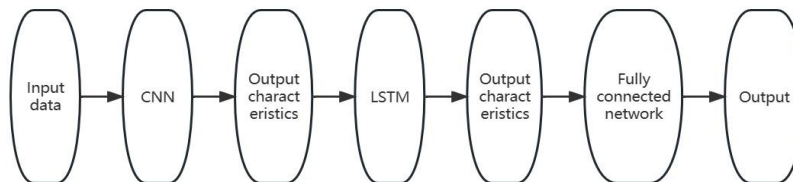


Figure 3: CNN-LSTM network model

In the current work, the first part of the CNN-LSTM model employed in this study is based upon a partial pooling layer of CNNs consisting of one convolutional layer and one max pooling layer, the raw data is fed to the CNN convolutional layer after preprocessing, and the lifetime characteristics are extracted adaptively using the convolutional kernel, the convolutional layer will iterate across the entire input dataset, and the weights of the convolutional kernel with local sequences will be convolutional with the body of the operational information of the tuple of bodies with the goal of deriving a preliminary feature matrix from them, which is far more expressive than raw sequence data (matrix).

The pooling layer that is used is the max pooling layer, which takes as input the feature matrix obtained from the previous convolutional layer, drag a pooling window over this matrix, and take the maximum value of the pooling window in each slide in order to produce a more expressive feature matrix.

Once the pooling is complete, an LSTM layer is connected to extract the correlation vectors built by the CNN into a long-term time series as the input data to the LSTM. In the convolutional layer, the data from the convolutional layer (Flatten) is propagated, Flatten is added to the model to compress their data (height, width, channel) into a one-dimensional array of channels of length height

and width, and then we can add a dense layer directly to it.

To perform feature compression operations on grouped convolutional data, multiple convolutional feature extraction framework extracted merged or merged features from the output layer, fully connected layer aggregates the learned features, the activation function using Relu. The Dropout rate is set to 0.2 prior to the final output to prevent overfitting of the model.

Typically, it is necessary to optimize the hyperparameters during model training in order to select an optimal set of hyperparameters for the model to improve prediction performance and efficiency. Empirically tuning the hyperparameters will result in a fine-tuned combination of model hyperparameters that may not be optimal, which will impact the model network's fit and ability to generalize to test data[7].

3. Example analysis

3.1 Data sources and pre-processing

3.1.1 Stock Price Influencing Factors

To compare the predictive effectiveness of LSTM and CNN-LSTM models, this paper collects stock-related data from Tushare from January 2016 to February 2022, using Ping An Bank as an example. Existing research shows that the factors affecting stock prices are complex, but can be mainly classified into opening price, closing price, high price, low price, previous closing price, up/down amount, up/down range, volume and revenue amount.

3.1.2 Data preprocessing

The construction of LSTM and CNN-LSTM prediction models requires consistent data dimensions for both the independent and dependent variables, but as the presence of missing data can affect the construction of prediction models, this paper takes the approach of pre-processing the data by removing the missing values. A total of 610 data sets from 8 May 2019 to 5 November 2021 were eventually left. We then divided the data set into a training set and a test set, where the first 500 data were used for model training and the later data were used for model testing. Specifically, we used the open-high-low-close and volume data to predict the next moment's closing price dataset. To avoid the possible impact of odd data on model training this paper uses min max scaler(). The fit function is used to normalize the data to a specified range, usually [0,1] or [-1,1]. Here is the normalization formula:

$$\frac{X_i - X_{min}}{X_{max} - X_{min}} \quad (1)$$

X_{min} (axis=0) is the row vector consisting of the smallest values in each column; X_{max} (axis=0) is the row vector consisting of the maximum values in each column; max corresponds to the maximum value of the interval to be mapped to, the default value is one; min is the minimal value of the interval to be mapped to, the default value is zero[8].

3.2 Key influences screening based on Pearson correlation analysis

To build an LSTM stock price prediction model, for example, there are many factors affecting stock prices, but not all of them are significantly correlated with stock prices, and if all of them are involved in the training model without screening, the convergence speed of the model will be reduced, affecting the robustness and generalization of the model. Therefore, this paper adopts Pearson's correlation analysis to screen the possible correlation influencing factors and exclude the correlation

factors with a significance level greater than 0.01. The final screened influencing factors are opening price, closing price, highest price, lowest price and volume.

3.3 LSTM-based stock price prediction

3.3.1 Key parameter settings

Taking the construction of a stock price forecasting model as an example, the LSTM forecasting model is used in this paper to set the key parameters in Table 1.

Table 1: Key parameters of the LSTM prediction model

Name	Parameters
Timestep	1
Batch_size	16
Input_dim	4
Hidden_dim	64
Output_dim	1
Num_layers	3
Epochs	10
Best_loss	0

3.3.2 Training and validation

Take the construction of a stock price prediction model as an example, the normalized data are used as training data for the model, and the stock price prediction values are obtained by using the inverse normalization of the data predicted by the trained model. The fitted curve of stock price predictions from the LSTM regression model to the original values is shown in Figure 4.

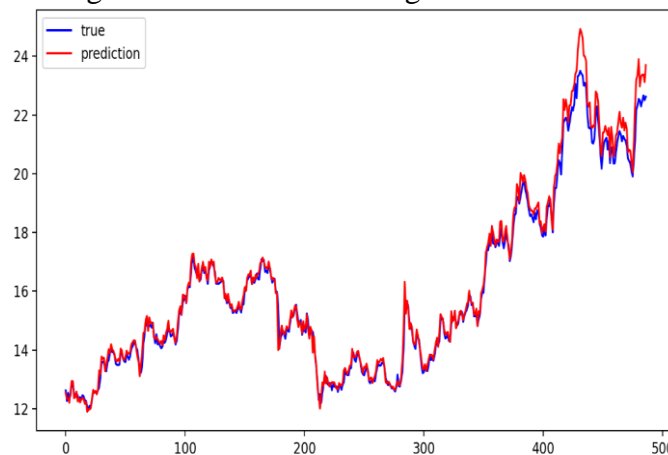


Figure 4: Fitting curves of stock price predictions from LSTM regression models to the original values

3.4 CNN-LSTM based stock price prediction

3.4.1 Key parameter settings

Taking the construction of a stock price prediction model as an example, this paper uses a CNN-LSTM prediction model to set the key parameters in Table 2.

Table 2: CNN-LSTM prediction model key parameters

Name	Parameters
Timestep	1
Batch_size	32
Feature_size	1
Hidden_size	256
Out_channels	50
Output_size	1
Num_layers	2
Epochs	10
Best_loss	0
Learning_rate	0.0003

3.4.2 Training and validation

In this paper, all the normalized data are used as training data for the model, and the predicted data from the trained model are used for inverse normalization to obtain the stock price prediction values. The fitted curve of the stock price prediction value of the CNN-LSTM regression model with the original value is shown in Figure 5.

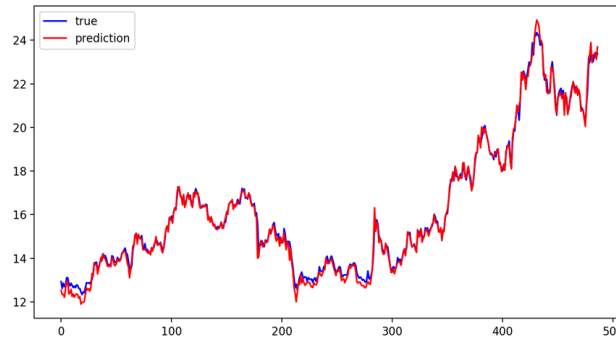


Figure 5: Fitting curves of stock price predictions to original values for CNN-LSTM regression models

3.5 Prediction results and accuracy checks

In this paper, the trained models are evaluated, firstly the prediction results need to be back-normalized and then visualized to obtain the prediction comparison of the LSTM regression model in Figure 6 and the CNN-LSTM regression model in Figure 7, and the RMSE is calculated. The final RMSE result is 0.6430091611855628.

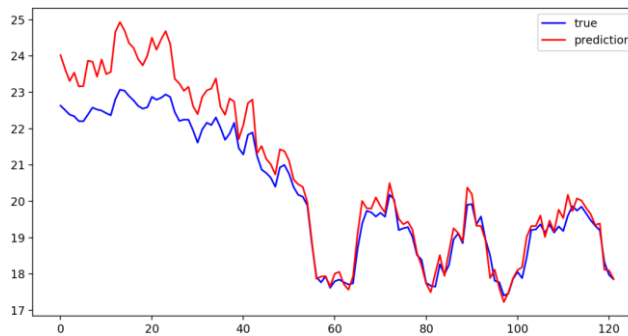


Figure 6: Comparison chart of predictions from LSTM regression models

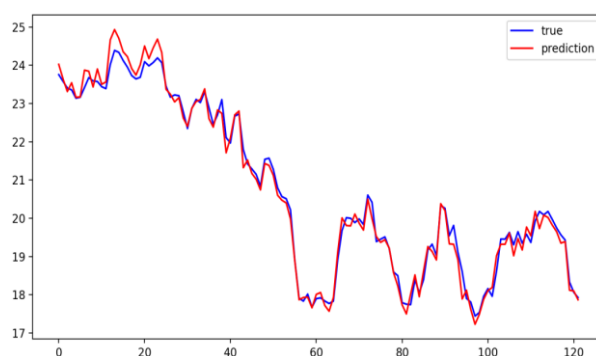


Figure 7: Prediction comparison chart for CNN-LSTM regression model

It can be seen from the experimental results that our simple CNN-LSTM model can achieve a good fit to the data and a low RMSE.

4. Conclusion

This paper analyzes and researches stock price prediction by applying deep learning method. Firstly, it is feasible to study stock price change by using neural network. A neural network model for action prediction combining CNNs and LSTM neural networks is then proposed, which can efficiently study changes in stock prices and provide insight for building a neural network model based on the characteristics of stock data. Taking Ping An Bank as an example, the prediction experiment on its share price tests the effectiveness of both LSTM and CNN-LSTM and also shows that for the CNN-LSTM model, the prediction accuracy and stability are better. However, the prediction of long-term stock data is not perfect at present. It is necessary to optimize and refine the neural network, such as improving error function and improving the accuracy of future stock price forecast.

References

- [1] Bao W, Yue J, Rao Y. A deep learning framework for financial time series using stacked autoencoders and long-short term memory [J]. *PloS one*, 2017, 12(7): e0180944.
- [2] Han Shanjie, Tan Shizhe. Design and implementation of a deep learning model for stock prediction based on Tensor Flow [J]. *Computer Applications and Software*, 2018, 35(6): 267-271.
- [3] Li M. Research on stock market prediction based on deep learning [D]. Nanjing University of Posts and Telecommunications, 2019. doi:10.27251/d.cnki.gnjdc.2019.000946.
- [4] Zhang Zeya, Huang Liming, Chen Chong et al. A stock prediction method based on news feature extraction and recurrent neural networks[J]. *Journal of Literature and Data*, 2020, 2(01):45-56.
- [5] Gao W. Modeling stock market using new hybrid intelligent method based on MFNN and IBHA[J]. *Soft Computing*, 2022, 26(15): 7317-7337.
- [6] Su XN. Data warehousing and data mining [M]. Tsinghua University Press Limited, 2006.
- [7] Wang Min, Lu Xikun, Zhou Yidi. Gun Life Prediction Model Based on Bayesian Optimization CNN-LSTM. *Integrated Ferroelectrics*, 2022.
- [8] Teng ZW, Zhou JH. Anomalous cell detection based on variational self-encoder[J]. *Mobile Communication*, 2020.