Stock Price Prediction based on CNN-LSTM Model in the PyTorch Environment

Weidong Xu[1,*]

[1] *Shanghai University of Electric Power*
[*]*Corresponding author. Email: alvinxu1399@outlook.com*

ABSTRACT
The stock market, as the main financing channel for listed companies and the most accessible wealth creation opportunity for investors, has always attracted attention from all walks of life. With the evolution of the technology, deep learning has started to play a very important role in forecasting stock price. Based on in-depth research on CNN and LSTM, this paper builds a CNN-LSTM stock price prediction model in PyTorch environment, and takes the data from the A-share market, choosing Shanghai Composite Index for a total of ten years from January 2012 to December 2021 as the experimental object, then verifying the feasibility of this joint model in the field of stock price forecasting, while comparing with the predicted values obtained using CNN and LSTM alone. The result confirms that the CNN-LSTM joint model performs well.

Keywords: *Stock price prediction, PyTorch, CNN, LSTM.*

1. INTRODUCTION

As the most important financing channel for listed companies and the most accessible opportunity for investors to create wealth, the stock market has always attracted attention from all walks of life. Nowadays, the majority of investors in the stock market are "retail investors". In the investment process, the biggest problem that retail investors faced with is their lack of relevant investment knowledge and the "information asymmetry" exists. To obtain valuable information, retail investors always have to pay a high price, so that they often rely on "grass news" which is cost-free rather than scientific technical analysis, resulting in the phenomenon of "blind investment" and "following suit". Therefore, stock price prediction has become particularly important. It is the emergence of stock price forecasting that can help retail investors to obtain effective data at a low cost in a market that lacks transparency, discover stocks and industries that may rise, and to some extent even reduce noise in the stock market.

In recent years, we are gradually entering the era of big data, neural network models have begun to be used in the field of stock price prediction. Stock price, as a proxy for time series data, time series models have become the mainstream to predict them. As a new type of recurrent neural network model, LSTM can solve the problem of gradient disappearance very well because of its good selectivity, memory and internal influence of time series. Since the CNN model was proposed, it has been mostly used in image feature extraction and face recognition. However, the advantage that CNN can extract the abstract feature vector which reflects the data information from the given training data set also do well in processing and learning data, so CNN may also have a good effect in the field of stock prediction.

Based on the in-depth study of CNN and LSTM, in order to further improve the stock prediction accuracy, this paper builds a joint stock price prediction model of CNN-LSTM in the PyTorch environment. It is hoped that the empirical study of this model in the field of stock price forecasting can broaden the research perspective and enrich the content of stock price forecasting based on neural network technology. At the same time, I also hope that it can provide a more accurate prediction reference for "retail investors" in real life.

2. LITERATURE REVIEW

Before the neural network technology became the mainstream, scholars all over the world regarded historical stock data as time series, and were keen to establish time series models to fit stock price changes to predict the future share price trends, such as linear regression methods like ARIMA, ARCH and GARCH. Adebiyi et al. extracted the influencing factors of stock

Y. Jiang et al. (Eds.): ICEDBC 2022, AEBMR 225, pp. 1272–1276, 2022.
https://doi.org/10.2991/978-94-6463-036-7_188

prices and used the ARIMA model to forecast stock prices in the future, and the outcome confirmed that the ARIMA model performed well in short-term.[1] Xu Feng selected China Southern Airlines and China Eastern Airlines as the research objects, building a complete model for prediction of two stock through the autoregressive process. The study found that the GARCH model had short-term memory in predicting stock prices.[2] It was concluded that the method of linear regression was only suitable for short-term stock price forecasting, while the long-term forecasting performance was poor.

With the advancement of science and technology, machine learning and deep learning have begun to be favored by scholars. Compared with linear regression models, neural network models have strong potential in further improving the accuracy of stock price forecasting, because they can extract nonlinear features. Nowadays, most of the current research is devoted to exploring how to form a mixed model according to the different characteristics of each model, find the optimal model collocation, and complement each other. Xu Yuemei et al. used the CNN-BiLSTM model combined with financial news sentiment analysis to predict stock prices, confirming that it does perform well in long-term forecasting.[3] Yu Z et al. firstly utilized LLE, after obtaining the processed data, then implanting them into the BP neural network, and found that the prediction accuracyof LLE-BP was better in comparison with single BP, PCA-BP and ARIMA.[4]

PyTorch, as a concise, efficient and fast Python open source machine learning library, is seldom used in the existing research on price forecasting. Therefore, this experiment attempts to use the CNN-LSTM model in the PyTorch environment to realize stock price prediction.

3. STOCK PRICE PREDICTION BASED ON CNN-LSTM

3.1. Technical Rout And Research Method

This experiment used PyTorch as the framework of the neural network, used the Python language as the code implementation of the network, and selected the data information of Shanghai Stock Index stocks from January 2012 to December 2021 using one of the most influential websites in China, "NetEase Finance". A total of 2431 samples were chosen to establish a dataset, implanting them into CNN to obtain output data, and then placed output data into the LSTM model to obtain predicted values. Through this progress to realize empirical research in the field of stock price prediction.

In this experiment, the comparative analysis method was adopted to compare the trend map of the CNN-LSTM forecasting outcomes with the actual outcomes, and at the same time compared it with results predicted by the CNN and LSTM models alone, which proved the superiority and efficiency of the hybrid model.

3.2. Stock Data Feature Extraction And Preprocessing

The prediction of the closing price of the Shanghai Stock Exchange Index was the purpose of this experiment, and it was particularly important to decide what kind of characteristics of the stock should be selected as the input factor. In this experiment, the basic trading indicators that are most closely related to stock price fluctuations were selected, namely opening price, closing price, high price, low price and volume.

After obtaining the dataset, the first 70% was classified as the training set, and the last 30% was classified as the test set. Since each trading indicator of a stock has different dimensions and units, this will make it impossible to compare the data and finally affect the experimental results. So it is necessary to use data normalization to eliminate the dimensional influence among indictors. In this experiment, the Z-score standardization method was used to perform dimensionless processing on each feature factor, which was beneficial to accelerate the training speed and increase prediction accuracy of the model.

In the PyTorch environment, DataLoader was used in this experiment as the uploading tool, which could convert the preprocessed dataset into an iterator, output a predefined batch_size number of images per iteration, and then use shuffle random numbers to shuffle the data in the batch to avoid overfitting.

3.3. Construction And Training Of The CNN

After the pre-commissioning of the convolutional neural network in this experiment, since the convolution kernel extracted features in two dimensions, the two-dimensional convolution Conv2D function was used to construct the convolution layer. The first layer of CNN was the input layer, with a 7-day window, and 5 data of "opening price, closing price, highest price, lowest price and trading volume" as input feature, and the number of channels was set to 1. The second layer was the convolution layer. Since the convolution kernel acted on the stock factor data, it was essentially performing factor synthesis, so only one convolution layer was used in this experiment. The number of convolution kernels was set to 64, and the size of the convolution kernel was set to 3*3. At the same time, in order to keep the data size unchanged, padding was applied. After the convolution layer, ReLU was adopted as the activation function. Because ReLU is linear and unsaturated, so that the convergence speed of the ReLU activation function is faster than other activation functions. In addition, adding a BatchNormalization layer after the convolutional layer could alleviate the over-fitting phenomenon and speed up the training and convergence of the neural network. The third layer was the pooling layer, the pooling area was 2 ×2, and the stride was 1. The function of pooling was to further sample the convolved sample features. The fourth layer was the Dropout layer, and the dropout rate was set to 0.3. The function of this layer was to alleviate the overfitting phenomenon of the model. So far, in the joint model, the construction of the CNN model was completed.

3.4. Construction And Training Of The LSTM

The LSTM model is a special RNN model. After effective features are extracted by CNN model, and when implanting them into LSTM, it can not only find the interdependence of the data in the time series data, but also automatically detect the best mode suitable for the relevant data. Mainly thanks to the three gates of LSTM: input gate, forget gate and output gate. In practical application, it is necessary to adjust the parameters of LSTM according to different situations to ensure the optimal results.

In this experiment, after the CNN output the data, it was then inputing into the LSTM layer. After continuous debugging, a total of two LSTM layers were set in this experiment. The neuron nodes of the first layer of LSTM layer were set to 128, and the neuron nodes of the second layer were set to 64. In terms of hyperparameter adjustment, the learning rate was set to 0.0001, the number of iterations (epochs) was 30, and the ReLU function was still adopted for activation, and finally outputed the data.

3.5. Loss Value Assessment

The final loss value of this experiment used MSE for loss evaluation. In this paper, Adam was the optimizer chosen to use, and the 0.001 was set as learning rate to calculate the optimized loss value. The loss diagram of the CNN-LSTM joint model was shown in Figure 1.

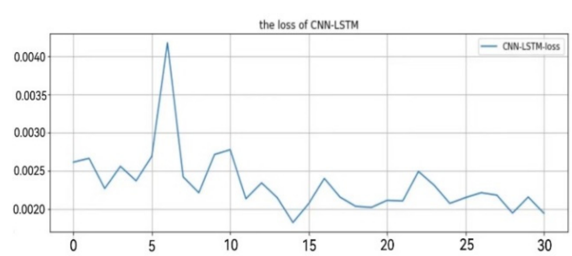

Figure 1 MSE loss value of CNN-LSTM model.

In the process of building the joint model, this experiment also realized the prediction of stock price by CNN and LSTM model separately in the PyTorch environment. The fitted MSE loss graph standing for CNN predicting result was shown in Figure 2, and the fitted MSE loss graph standing for LSTM predicting result was shown in Figure 3. The result proved that the loss value of the CNN-LSTM joint model was significantly lower than the loss value when CNN and LSTM were used alone.

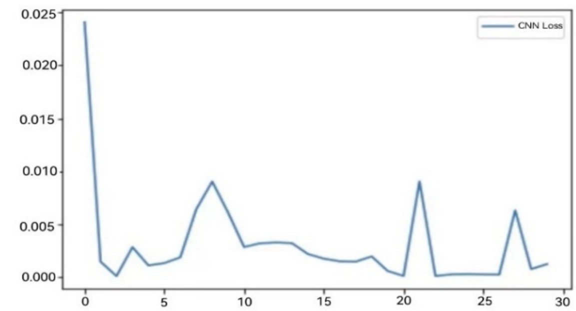

Figure 2 MSE loss value of CNN model.

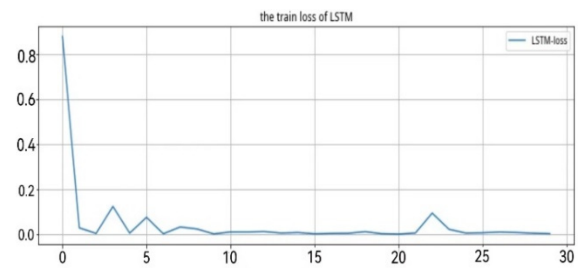

Figure 3 MSE loss value of LSTM model.

3.6. Analysis Of Research Results

Conclusions drawn from the changes between the predicted results and the real values in Figure 4 were listed as follows: Firstly, the stock price trend predicted by the CNN-LSTM joint model can be almost consistent with the real price trend, and the error results are relatively small. Secondly, although the real stock price fluctuates in a wide range and with sharp fluctuations, the joint model of CNN-LSTM can well capture a large amount of mutation information in the real stock price fluctuation, and can better predict the price trend of these mutations, which reflects the robustness of the CNN-LSTM joint model.

Although the CNN-LSTM joint model reflects better prediction results, it still has some flaws. It can be seen from the trend chart that there will be a certain lag between the predicted trend and the real trend, and it is difficult for prediction value to respond immediately. This is also the evidence that the Chinese stock market is a weakly efficient market.

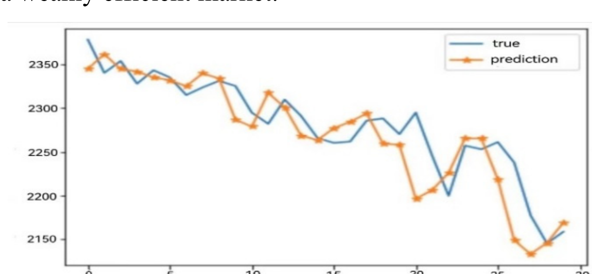

Figure 4 Trend comparison between predicted value of CNN-LSTM and true value.

In order to further verify the excellent prediction ability of the CNN-LSTM joint model, a comparison experiment was conducted with the forecasting outcome of the CNN and the LSTM model alone. The comparison results of CNN and the real value are shown in Figure 5. The comparison results of LSTM and the real value are shown in Figure 6.

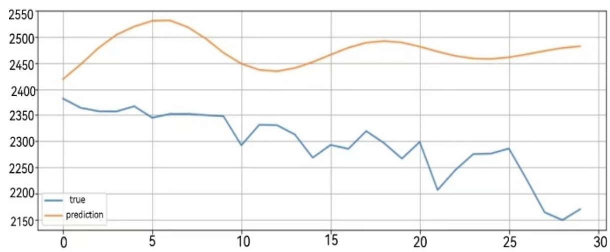

Figure 5 Trend comparison between predicted value of CNN and true value.

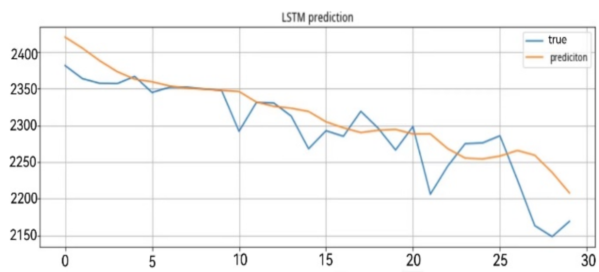

Figure 6 Trend comparison between predicted value of LSTM and true value.

From the comparison of Figure 4, Figure 5, and Figure 6, it can be seen that the prediction result based on the CNN-LSTM joint model is the best, indicating that CNN-LSTM can achieve good results in stock price prediction. The prediction ability of the LSTM model is second. It can be seen from the comparison that it can roughly predict the trend, but the ability to capture mutation information is not as strong as that of the CNN-LSTM model. The CNN model shows the worst prediction result, whose trend curve changes slightly, and the predicted curve is too smooth, which makes it almost impossible to accurately predict the stock price. Therefore, the predicted stock price deviates greatly from the real stock price.

4. CONCLUSION

This experiment realized the prediction of stock prices with the CNN-LSTM joint model in the PyTorch environment. Through its performance on the Shanghai Composite Index dataset, it can be concluded that when making stock predictions, the model can not only show the overall trend of stock prices, but also show subtle changes, finally achieving good prediction results.

However, at the same time, there are still some shortcomings in this experiment. For example, this paper only uses five basic trading indicators of stocks as feature input. Subsequent experiments can consider adding more stock features, such as technical indicators, in order to achieve better prediction effect. Secondly, due to the existence of many influencing factors in the stock market, only numerical data used can not fully capture fluctuations. In the future experimental process, stock public opinion analysis or financial news evaluation analysis will also be considered to extract emotional factors, and combined with CNN-LSTM model for prediction, I believe there will be better results.

REFERENCES

[1] Ayodele Adebiyi A, Aderemi O, Adewumi Charles K.Ayo. (2014) Stock price prediction using the ARIMA model. 2014 UK Sim-AMSS 16th International Conference on Computer Modelling and Simulation, Cambridge: IEEE, 106-112.

[2] Xu, F. (2006) GARCH Model for Stock Price Prediction. Statistics and Decision,18:107-109.

[3] Xu, Y. M., Wang, Z. H., Wu, Z. X. (2021) A CNN-BiLSTM based Multi-feature Integration Model for Stock Trend Prediction. Data Analysis and Knowledge Discovery, 5(7): 126-137.

[4] Yu Z, Qin L, Chen Y. (2020) Stock price forecasting based on LLE-BP neural network model. Physica A: Statistical Mechanics and its Application, 553: 124197.