Dean&Francis
ISSN 2959-6130

FE005682

Predicting the stock prices of listed companies based on the LSTM model of recurrent neural network

Jiayue Lai

Abstract:
This experiment uses the LSTM model and establishes a two-layer LSTM structure to enhance the learning ability of long-term dependencies. The Dropout layer can prevent the model from overfitting historical data during training. Normalization, time series sliding window technology, and 80-20 training-test split method are used for data preprocessing. MAE, RMSE, and MAPE are used for evaluation to compare the predicted value with the actual value.

Keywords: LSTM, RNN, KFC dataset, error analysis, MAE, RMSE, MAPE.

1. Introduction

This is a serious problem. Stock price prediction has many complexities and is easily affected by market sentiment and some economic indicators as well as external events. Most time series methods, cannot capture its nonlinear complex relationships [1]. With the continued advancement of deep learning, RNN and its variant LSTM are able to learn long-term dependencies, opening up huge possibilities for financial forecasting.

Typically, researchers use the predictive power of a single-layer LSTM model to predict stock prices to analyze market trends and volatility. Although such effects are stable, there are long-term dependencies that cannot be handled and long-term trends of stocks cannot be mined [11]. Then, the performance of LSTM on long-term dependencies is evaluated.

This article will explore the accuracy of LSTM in stock price prediction and use KFC Dataset.csv, a stock dataset of KFC, for some tests. This dataset covers historical stock data and key features from 2000 to 2024. It helps to provide a more comprehensive picture of long-term market fluctuations. This article uses LSTM to establish a two-layer LSTM structure. The first layer of LSTM passes time series features to the second layer of LSTM to enhance the learning ability of long-term dependencies. And add a Dropout layer to avoid the model from overfitting historical data during training [12]. In order to complete the research in this article, data preprocessing methods such as normalization, time series sliding window technology, and 80-20 training-test split were used. The mean square error (MSE) loss function was optimized for both LSTM and RNN models, and the gradient was updated using Adam. Then, the stock price predicted by the model is compared with the actual stock price, and the performance is evaluated using metrics such as MAE, RMSE, and MAPE.

2. Research Methods

LSTM is a special type of RNN that is specifically designed to solve the “long sequence dependency

ISSN 2959-6130

problem" that ordinary RNNs cannot effectively handle. LSTM uses a "gating mechanism" to control the flow of information, remembering important information for a long time and discarding irrelevant information [4]. In the LSTM structure, "memory units" and "three gating mechanisms" are used to control the direction of message.

In the memory unit, the core is the "cell state", which is like a "conveyor belt" that can directly transmit information, allowing the model to remember useful information for a long time [5]. Its update formula is as follows:

$$C_t = f_t * C_{t-1} + i_t * \widetilde{C_t} \tag{1}$$

Where C_t is the cell state at the current time step, C_{t-1} is previous cell state, f_t is the control value of the forget gate, i_t is the input gate and $\widetilde{C_t}$ is the new information.

In the forget gate, the formula that determines how much past information the current unit should forget is as follows:

$$f_t = \sigma(W_i * [h_{t-1}, x_t] + b_f) \tag{2}$$

If the f_t value is close to 0, the LSTM will discard the information; if it is close to 1, the information is retained.

In the input gate, the formula for determining how much new information to store at the current moment is as follows:

$$\begin{aligned} i_t &= \sigma(W_i * [h_{t-1}, x_t] + b_i) \\ \widetilde{C_t} &= \tanh(W_C * [h_{t-1}, x_t] + b_C) \end{aligned} \tag{3}$$

i_t is the control input ratio and $\widetilde{C_t}$ is the newly calculated information.

The advantages of LSTM in applications are as follows:

Ÿ Solve the gradient vanishing problem: remember long-term dependencies.

Ÿ Capture long sequence information: better than ordinary RNN in processing long sequences [6].

Ÿ High flexibility: combined with CNN and Transformer to handle different tasks [7].

3. Evaluation

This experiment uses the KFC stock price dataset (KFC Dataset.csv), which comprehensively records the changes in KFC stock prices over time. The dataset covers 24 years of stock market data from 2000 to 2024. The time span is relatively long, and it can effectively analyze long-term market trends.

3.1 Model evaluation methodology

3.1.1 Data preprocessing

In order to improve the training effect of LSTM, this experiment first normalizes the data to between 0 and 1 to avoid the problem of calculation overflow caused by too large values [8]. Then, the time series sliding window is used to predict the stock price of the next day using the stock price data of the past 60 days, so that the model can learn short-term trends. Finally, the data set is divided into 80% training set and 20% test set. The predictive ability is verified by the test set to ensure that the model can fully learn historical data.

3.1.2 LSTM and RNN model training

In order to compare the performance of different neural networks in stock price prediction, this experiment constructed two models, LSTM and RNN, for testing and evaluation.

LSTM uses a two-layer LSTM network, each layer contains 50 neurons, and uses a Dropout layer to prevent overfitting, and finally uses a connection layer to output the predicted value. The first layer of LSTM outputs the hidden state of all time steps of the entire time series, not just the last time state, which allows it to learn deeper time series features. The second layer LSTM only outputs the final prediction. Compared with single-layer LSTM, double-layer LSTM can mine more rich feature information, thereby improving prediction accuracy.

RNN uses a similar architecture, but uses SimpleRNN instead of LSTM, which makes it easy to compare the effects of traditional recurrent neural networks and LSTM in time series prediction. Both models use MSE as the loss function and use the Adam optimizer for training. After training, this experiment uses the test set data for prediction and denormalizes the prediction results to achieve comparative analysis with the actual stock price.

3.2 Comparison of LSTM and RNN prediction results

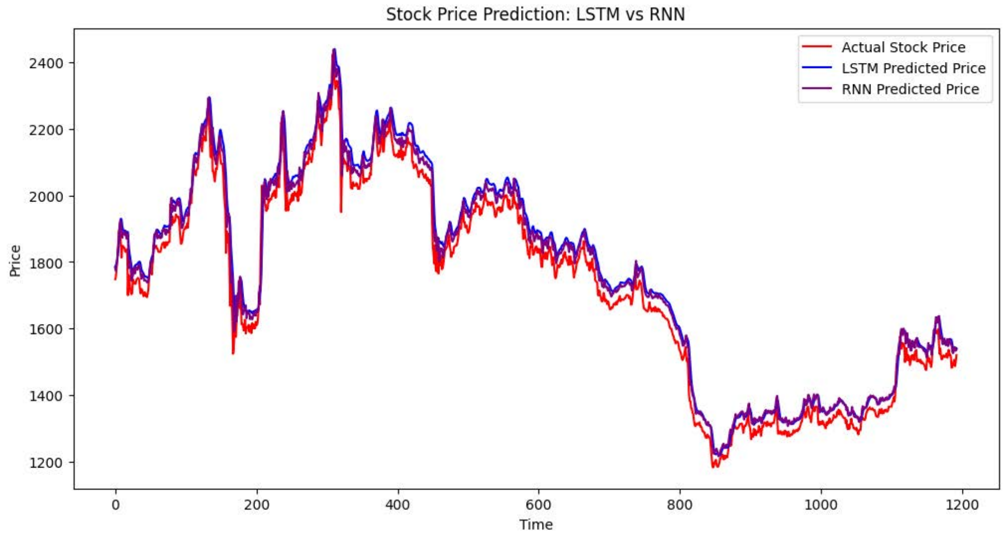

Fig. 1. LSTM vs RNN.(Picture credit: Original)

In Fig.1., the LSTM (blue) prediction curve is closer to the true value (red), indicating that it can better capture the fluctuation trend of stock prices. RNN (purple) can also follow the overall trend, but the deviation is large at some turning points, which may be because RNN cannot capture the dependencies of long time series well. According to the prediction error analysis, LSTM has a smaller error. Because LSTM uses a gating mechanism, it can more effectively remember the long-term dependencies of historical data, so that at most 19 points, the predicted value closely follows the actual stock price trend. However, RNN has a larger error. Since RNN cannot effectively store long-term memory, it may cause large deviations in long-term trend predictions, especially when stock prices fluctuate violently.

3.3 Error Analysis

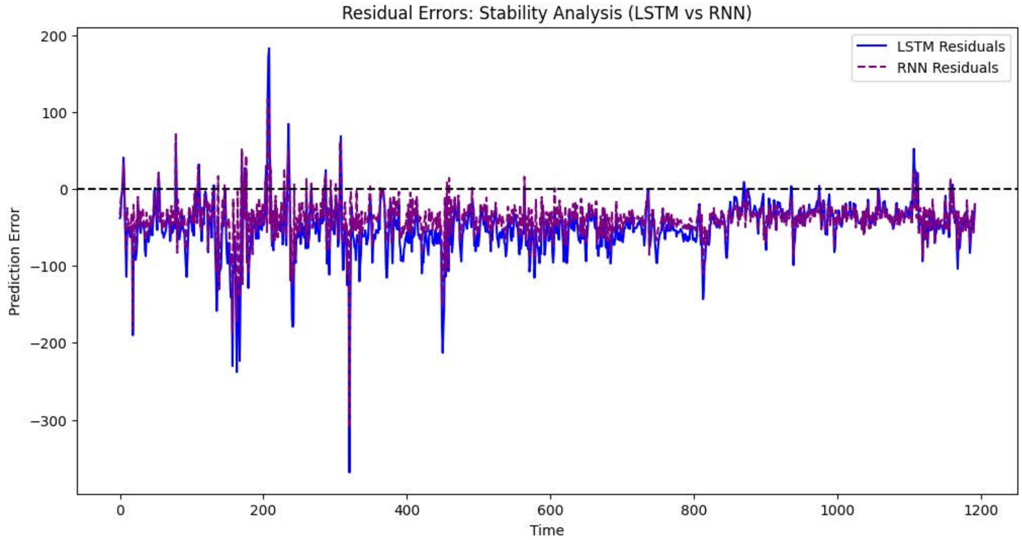

Fig. 2. Error stability analysis (LSTM vs RNN).

(Picture credit: Original)

In Fig.2, the prediction residuals of LSTM and RNN fluctuate around 0, but the error of LSTM at multiple time points is relatively small, while RNN has larger deviations at some points. This shows that LSTM is more stable than RNN in long-term series dependency learning and can better capture market trends.

After evaluating LSTM, the following indicator results

ISSN 2959-6130

were obtained: The MAE is 23.97 and the RMSE is 34.68, indicating that the average error between the model prediction value and the actual value is small and has good accuracy. At the same time, MAPE is only 1.42%, indicating that the model's prediction error ratio is very low and the overall prediction effect is relatively accurate.

3.4 Comparison of prediction results

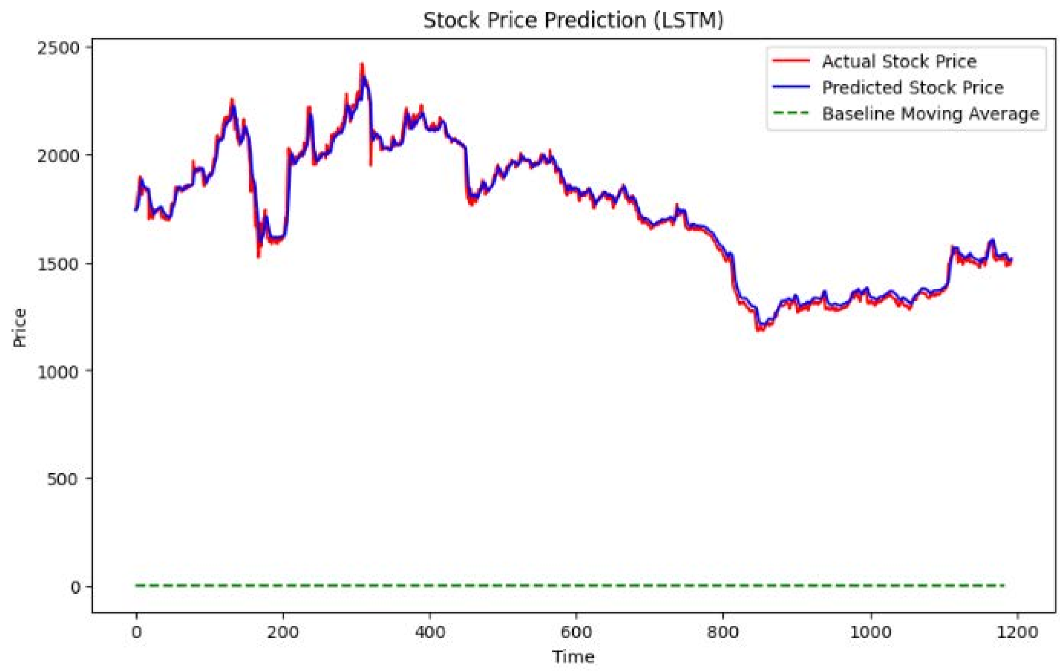

Fig. 3. LSTM prediction results.

(Picture credit: Original)

From Fig.3., the predicted stock price (blue) of the LSTM is highly consistent with the actual stock price (red), and can track the market trend well. The moving average baseline model (green dotted line) can not adapt to the dynamic changes of market prices because it only considers the historical mean, which shows that LSTM is more practical in the stock prediction task. In addition, LSTM can quickly adjust the predicted value at the turning point of trend change (such as market rise or fall).

3.5 Experimental conclusion

The experimental results show that LSTM is more suitable for time series prediction, can effectively handle long-term dependencies, and better capture market trends, and has better prediction effects than RNN.

4. Conclusions

This study uses LSTM and RNN to predict stock price time series, and uses the moving average method as a comparison baseline. The experimental findings indicate that, within the stock-prediction implementing schemes, LSTM performs better than the other models of RNN types and then the moving-average baseline models, with features such as accuracy, error stability, etc., obviously favoring LSTM.

4.1 Research limitations

Although LSTM performed well in this experiment, it still has limitations:

Ÿ The ability to predict sudden market events is limited, and LSTM may still have large errors.

Ÿ Long-term prediction errors accumulate, because LSTM relies on historical data for prediction. In the long-term prediction process, small errors will gradually accumulate, resulting in a decrease in long-term prediction accuracy.

4.2 Future research directions

This experiment provides the following ideas to further improve the performance of LSTM in stock prediction:

Ÿ Combine CNN-LSTM and use CNN to extract short-term market features to improve sensitivity to market fluctuations.

Ÿ Introduce macroeconomic variables (such as CPI, GDP, interest rates) to enhance the model's predictive ability and enable it to better adapt to changes in the market environment.

In summary, this study verifies the feasibility of LSTM in forecasting. Compared with RNN and moving average baselines, LSTM has higher forecasting accuracy, better trend fitting ability, and stronger market adaptability. However, in order to further improve the generalization ability, it is still necessary to combine multiple models and external market factors to improve the practicality of LSTM in actual trading scenarios.

References

1. Patel, J., Patel, M., & Darji, M. (2018). Stock price prediction using RNN and LSTM. *Journal of Emerging Technologies and Innovative Research*, *5*(11), 1069-1079.
2. Ding, G., & Qin, L. (2020). Study on the prediction of stock price based on the associated network model of LSTM. *International Journal of Machine Learning and Cybernetics*, *11*(6), 1307-1317.
3. Selvin, S., Vinayakumar, R., Gopalakrishnan, E. A., Menon, V. K., & Soman, K. P. (2017, September). Stock price prediction using LSTM, RNN and CNN-sliding window model. In *2017 international conference on advances in computing, communications and informatics (icacci)* (pp. 1643-1647). IEEE.
4. Zhuge, Q., Xu, L., & Zhang, G. (2017). LSTM Neural Network with Emotional Analysis for prediction of stock price. *Engineering letters*, *25*(2).
5. Pahlawan, M. R., Riksakomara, E., Tyasnurita, R., Muklason, A., Mahananto, F., & Vinarti, R. A. (2021). Stock price forecast of macro-economic factor using recurrent neural network. *IAES International Journal of Artificial Intelligence*, *10*(1), 74.
6. Ho, M. K., Darman, H., & Musa, S. (2021, July). Stock price prediction using ARIMA, neural network and LSTM models. In *Journal of Physics: Conference Series* (Vol. 1988, No. 1, p. 012041). IOP Publishing.
7. Pramod, B. S., & Pm, M. S. (2020). Stock price prediction using LSTM. *Test Engineering and Management*, *83*(3), 5246-5251.
8. Minami, S. (2018). Predicting equity price with corporate action events using LSTM-RNN. *Journal of Mathematical Finance*, *8*(1), 58-63.
9. Kavinnilaa, J., Hemalatha, E., Jacob, M. S., & Dhanalakshmi, R. (2021, July). Stock price prediction based on LSTM deep learning model. In *2021 International Conference on System, Computation, Automation and Networking (ICSCAN)* (pp. 1-4). IEEE.
10. Wu, S., Liu, Y., Zou, Z., & Weng, T. H. (2022). S_I_LSTM: stock price prediction based on multiple data sources and sentiment analysis. *Connection Science*, *34*(1), 44-62.
11. Pawar, K., Jalem, R. S., & Tiwari, V. (2019). Stock market price prediction using LSTM RNN. In *Emerging Trends in Expert Applications and Security: Proceedings of ICETEAS 2018* (pp. 493-503). Springer Singapore.
12. Dey, P., Hossain, E., Hossain, M. I., Chowdhury, M. A., Alam, M. S., Hossain, M. S., & Andersson, K. (2021). Comparative analysis of recurrent neural networks in stock price prediction for different frequency domains. *Algorithms*, *14*(8), 251.