

A Study on CNN Feature Extraction for Stock Price Prediction

Yuanhang Li*, Zhengjie Xie

Jinan University – University of Birmingham Joint Institute, Jinan University, Guangzhou 511400, China

*Corresponding author Email: liyuanhang2019051771@stu2019.jnu.edu.cn

Abstract. The return on investment for investors in the stock market is highly dependent on the investor's timing strategy, that is, the decision of what time to buy or sell a stock. A successful timing strategy requires investors to accurately identify the price movement of a company. As a result, some investment professionals have created technical analysis analytical methodologies to forecast the short-term trend of a stock. However, technical analysis approaches are prone to subjectivity, such as the selection of technical indicators and indicator periods. This essay attempts to utilize a convolution layer in deep learning to extract features as an alternative to technical indicators and to reduce subjective elements' effect on prediction bias. Several stock predictions are evaluated between a standard LSTM model and an LSTM model with convolution layers (CNN-LSTM model) in this research. The experimental results show that the CNN-LSTM model outperforms the standard LSTM model in predicting the price of certain stocks with a big market capitalization and high liquidity.

Keywords: Stock Price Prediction, Technical Analysis, Deep Learning, CNN-LSTM Model.

1. Introduction

Stocks are the most common and most important underlying investment in the financial markets. To earn from investing in the stock market, investors must precisely predict the stock market's fluctuations. In the past, stock market investment experts utilized fundamental research, which included examination of macroeconomic factors and the financial position of the companies. This type of analysis is unsuitable for short-term investments since the short-term stock market is volatile and the reasons for volatility are largely unrelated to the company's value. To evaluate short-term movements in the stock market, investment experts have developed the technique of technical analysis.

Technical analysis is the study of the market's behavior to identify market trends and track their cyclical changes in order to make trading decisions in stocks and all financial derivatives. Technical analysis assumes that market behavior incorporates and assimilates all available information, that prices move cyclically, and that the past repeats itself[1]. In technical analysis, it is frequently important to make judgments regarding the patterns of stock Candlestick charts (K-charts), which are extremely subjective and frequently ineffective for novice investors.

In the past decade, with the rapid growth of deep learning, methodologies based on deep learning have been gradually applied to stock trend prediction (such as BP neural network, RNN, LSTM, etc.). Among them, LSTM models are the most accurate[2]. But models for single structural neural networks depend more on choosing features than on choosing the type of model. In order to solve the problem of difficult subjective selection of features, this paper refers to the analysis idea of technical analysis, using a convolution layer to extract features from several stocks that meet the conditions of technical analysis (large market capitalization, high turnover, long listing time, etc.), and then combine them with LSTM model for prediction. In addition, the research also explores the influence of different time periods on the model's effectiveness.

2. Methodology

2.1 Data Acquisition

The subject of this study was China Merchants Bank, the most heavily weighted company in the CSI 300 index. This stock's financial data from 2010-01-01 to 2017-12-31 were picked from Tushare as the data set, with the last 100 days as the test set and the remaining data as the training set. The features of the stock are displayed in table1. In addition, Ping An Bank and Kweichow Moutai were utilized for further testing[3][4].

Table.1. The features of the stock

| Name | Definition/Implication |
|------------------|---------------------------------------|
| Open/Close Price | Nominal daily open/close price |
| High/Low Price | Nominal daily highest/lowest price |
| Trading volume | Daily trading volume |
| Trading amount | Daily trading amount |
| Close-open | Daily increasing/decreasing degree |
| High-low | Highest price - Lowest price in a day |

2.2 Introduction of Technical Analysis Method

Practitioners utilize several technical indicators, but they generally fall into two broad categories: trend followers and countertrend indicators. This section will explain Moving Averages (MA) and Exponential Moving Averages (EMA), the most well-known trend followers[5].

(1) Moving Averages

Avramov has proved that the distances between short- and long-run moving averages of prices can predict cross-sectional stock returns well[6].The most widely used moving average (MA) is the n-day simple MA given by:

$$MA_{t,n} = \frac{1}{n} \sum_{i=t-n+1}^t P_i = (P_t + P_{t-1} + \dots + P_{t-n+1})/n \quad (1)$$

$MA_{t,n}$ is the simple n-day moving average at period t and P_i is the closing price for period i. When the short-term period moving average crosses above the long-term period moving average, it can be regarded as a buy signal, that is, the stock price will rise in the short term.

(2) Exponential Moving Average

An Exponential Moving Average (EMA) is a moving average (MA) that gives greater weight and importance to the nearest data point[7]. For the series $\{x_n\}$ define its exponential moving average $EMA_N(x_n)$ with period N up to the nth term as:

$$EMA_N(x_n) = \frac{2}{N+1} \sum_{k=0}^{\infty} \left(\frac{N-1}{N+1}\right)^k x_{n-k} = \frac{2x_n + (N-1)EMA_N(x_{n-1})}{N+1} \quad (2)$$

Since there is no data prior to x_1 , we add the definition $x_0 = x_{-1} = x_{-2} = \dots = x_1$ to the expression. Consequently, $EMA_N(x_1)$ equals x_1 . The defining equation for the EMA weighted average reveals its characteristics. The EMA indicator reduces the weighting factor for each day's price in an exponentially proportional manner. The closer a time period is to the present, the greater its significance. This indicates that the EMA function has a greater weighting for recent prices and is more responsive to recent price fluctuations. EMA is therefore more informative than MA.

2.3 Model

(1) LSTM model

The LSTM model is a variant of the RNN model. By introducing forgetting gates, memory gates, and update gates, the LSTM model solves the problem of gradient disappearance of the RNN model in dealing with long time series tasks, so the LSTM model is more suitable for dealing with long time series of stock data[8].

The experiments were conducted using windows of size 5, 10, 15, 20(Fig.1). Then using these data to predict the close price of next trading date[9], and use RMSE as the evaluation criteria.

$$RMSE(X, h) = \sqrt{\frac{1}{n} * \sum_{i=1}^n (h(x)^{(i)} - y^{(i)})^2} \tag{3}$$

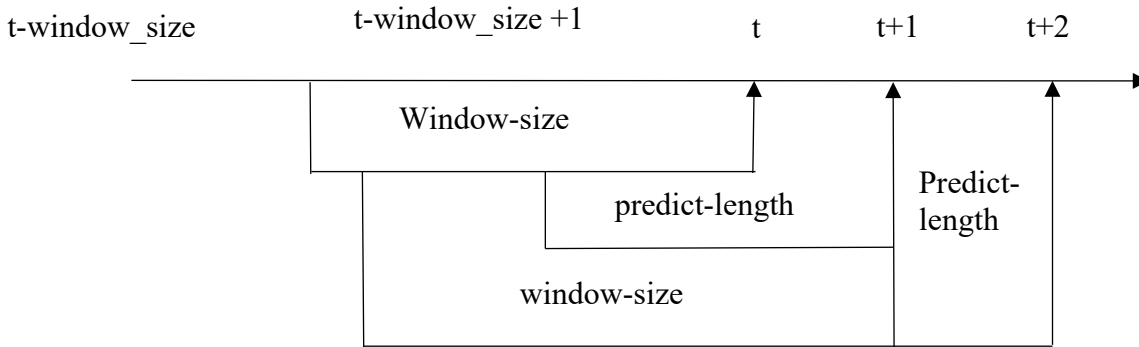


Figure 1. Sliding windows splitting data

Depending on the number of inputs, the LSTM model in this paper contains one LSTM layer with 64 LSTM cells, one dropout layer which can prevent over-fitting of the model, and two dense(linear) layers, as shown in Fig. 2.

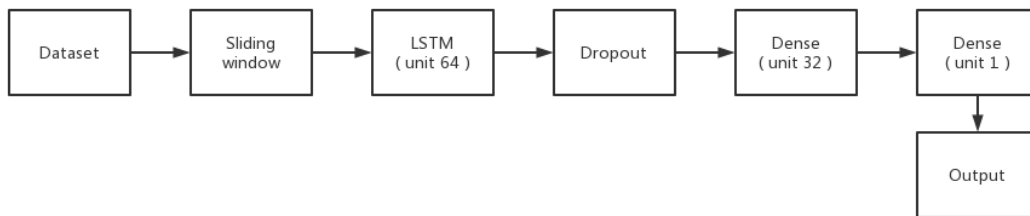


Figure 2. The layout of the LSTM model

(2) CNN-LSTM model

The purpose of CNN is to extract features from things in a certain model, and later classify, recognize, predict or decide on that thing based on the features, etc. CNN’s are commonly utilized in image recognition due to the presence of convolutional layers. And in technical analysis, analysts must predict the next movement of a stock based on its technical indicator charts (such as the MA and EMA) for the previous N days. This paper takes into account this strategy by treating (window size, feature number) as a 2-D image and applying convolutional layers to extract additional stock features to replace artificially calculated technical indicators and reduce the prediction bias caused by subjective analysis[10].

The CNN-LSTM model in this thesis consists of a CNN layer (consisting of a convolution layer and a mean pooling layer) added to the LSTM model described previously, as depicted in the figure below. The kernel size in the convolutional layer is set to (3, 1) and the pool size in the averaging layer is set to (2, 1), alluding to the mean-like analysis technique used in technical analysis.

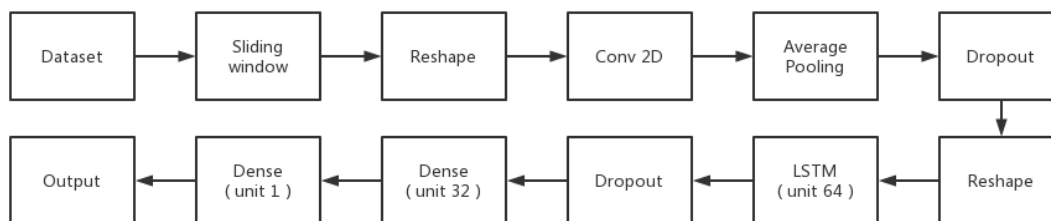


Figure 3. The layout of the CNN-LSTM model

3. Results analysis

According to China Merchants Bank's test results, the LSTM model performs optimally with a window size of 15 and an RMSE of 0.481. The CNN-LSTM model outperforms the LSTM model in various time windows, with the best performance in the 20-day window and an RMSE of 0.378, which improves the prediction effect by approximately 20% compared to the LSTM optimal window. This demonstrates that the use of CNN to extract features is effective.

Table 2. Test results of China Merchants Bank

| | | 5 | 10 | 15 | 20 |
|----------------------|----------|-------|-------|-------|-------|
| China Merchants Bank | LSTM | 0.754 | 0.606 | 0.481 | 0.602 |
| | CNN-LSTM | 0.702 | 0.532 | 0.423 | 0.378 |

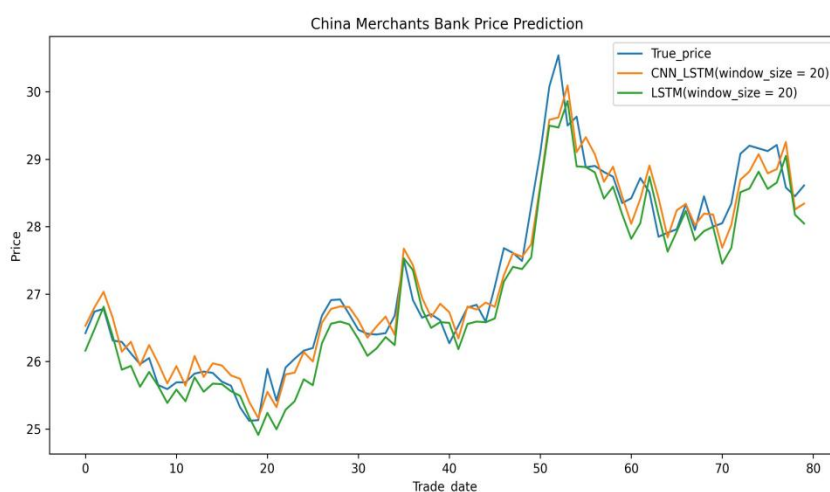


Figure 4. China Merchant Bank Price Prediction

To further test the validity of the CNN layer, two companies with a relatively large market capitalization in the CSI 300, Ping An Bank and Kweichow Moutai, were randomly selected to test the model and the results are shown in Table 3.

Table 3. Results of Ping An Bank and Kweichow Moutai

| | | 5 | 10 | 15 | 20 |
|-----------------|----------|-------|-------|-------|-------|
| Ping An Bank | LSTM | 0.157 | 0.148 | 0.156 | 0.156 |
| | CNN-LSTM | 0.153 | 0.145 | 0.128 | 0.127 |
| Kweichow Moutai | LSTM | 0.434 | 0.653 | 0.354 | 0.610 |
| | CNN-LSTM | 0.149 | 0.128 | 0.889 | 0.789 |

In Ping An Bank's test findings, the CNN-LSTM model outperforms the LSTM model in all four time periods. In Kweichow Moutai's test results, the window size had a greater impact on the CNN-

LSTM model's performance. This is likely because Kweichow Moutai's short-term fluctuations were greater, rendering the usage of a longer time frame to extract features inefficient. Nonetheless, under ideal window size, the CNN-LSTM model outperforms the LSTM model.

Combining the test results of the three companies, the optimal time window size for the LSTM model is 15, while the effect of the CNN-LSTM model is more dependent on the window size. However, as long as the optimal window size is determined, the CNN-LSTM model will outperform the LSTM model, proving that CNN is an effective tool for extracting features.

4. Conclusions

This paper aims to apply a unique deep learning technique to the prediction of stock prices. Specifically, the CNN-LSTM model suggested in this article is built on a simple LSTM model concerning the concept of technical stock analysis and adds a convolutional layer to extract information from previous trading days for improved prediction outcomes. The CNN-LSTM model has been shown to have good prediction results for the three selected stocks, and the results have been significantly enhanced compared to those of the standard LSTM model.

However, in the stock market, technical analysis does not apply to all stocks at all times. A stock must have been listed for an extended period of time to ensure that the pattern of recent swings has occurred in the past. This implies that if the previous image does not depict the present state of the stock's movement, the CNN layer's retrieved features may be invalid. This is where the model's shortcomings become apparent.

References

- [1] Edwards, R.D., Magee, J., & Bassetti, W.H.C. (2018). *Technical Analysis of Stock Trends* (11th ed.). CRC Press. <https://doi.org/10.4324/9781315115719>
- [2] Roondiwala, M., Patel, H., & Varma, S. (2017). Predicting stock prices using LSTM. *International Journal of Science and Research (IJSR)*, 6(4), 1754-1756.
- [3] K. Chen, Y. Zhou and F. Dai, "A LSTM-based method for stock returns prediction: A case study of China stock market," *2015 IEEE International Conference on Big Data (Big Data)*, 2015, pp. 2823-2824, doi: 10.1109/BigData.2015.7364089.
- [4] Cho, C. H., Lee, G. Y., Tsai, Y. L., & Lan, K. C. (2019, December). Toward stock price prediction using deep learning. In *Proceedings of the 12th IEEE/ACM International Conference on Utility and Cloud Computing Companion* (pp. 133-135).
- [5] Wing-Keung Wong, Meher Manzur & Boon-Kiat Chew (2003) How rewarding is technical analysis? Evidence from Singapore stock market, *Applied Financial Economics*, 13:7, 543-551, DOI: 10.1080/0960310022000020906
- [6] Avramov, D., Kaplanski, G., & Subrahmanyam, A. (2018). Stock return predictability: New evidence from moving averages of prices and firm fundamentals. Available at SSRN 3111334.
- [7] Chen, Z. (2022, March). The Study of Using Machine Learning Algorithms to Construct Portfolio Formation. In *2022 7th International Conference on Financial Innovation and Economic Development (ICFIED 2022)* (pp. 2105-2109). Atlantis Press.
- [8] Liu, T., Ma, X., Li, S., Li, X., & Zhang, C. (2022). A stock price prediction method based on meta-learning and variational mode decomposition. *Knowledge-Based Systems*, 252, 109324.
- [9] Sezer, O. B., & Ozbayoglu, A. M. (2018). Algorithmic financial trading with deep convolutional neural networks: Time series to image conversion approach. *Applied Soft Computing*, 70, 525-538.
- [10] Kirkpatrick II, F. C. D., & Julie, R. (2017). Moving averages. *CMT Level II 2017: Theory and Analysis*.
- [11] CSI 300 index https://tushare.pro/document/2?doc_id=27