The 6th International Conference on Economic Management and Green Development (ICEMGD 2022)
DOI: 10.54254/2754-1169/3/2022823

Review of Stock Price Predicting Method Based on LSTM

Huizi Qian[1,a,*]

[1]*Department of Industrial Economics, University of Chinese Academy of Social Sciences, 102445, Beijing, China*
a. qianhuizish@163.com
**corresponding author*

Abstract: Stock market forecasting is a challenging field for investors to make profits in the financial market. Investors need to understand that financial markets are more unstable and affected by many external factors. Time series analysis of daily stock data and the establishment of prediction model are very complex. The development of stock market forecasting technology is changing with each passing day and deep learning method is more and more used in finance field. This paper review the stock predicting method based on LSTM from the year 2015 to 2022.

Keywords: Stock Price, Predicting, LSTM

1. Introduction

The stock price fluctuates and is influenced by a variety of complex factors such as social development, economic and financial situation, the listed company's life cycle and operating conditions, policies as well as news will all have an impact on the index. The stock market allows investors and traders to invest and is a key indicator of the country's economics so that it interacts with various elements such as traders' expectations, economic conditions, and political events. Furthermore, as communication technology improve, the rapid data processing of these events leads to the rapid change of stock prices. A stable and flexible stock market is also highly desirable by investment banks, financial institutions, individual investors, and stockbrokers in the financial field. Consequently, several academics have used fundamental analysis, technical analysis, time series prediction, and machine learning methods to enable stock market predictability. In addition, lots of companies have devised new methods for analyzing financial data and making investment decisions.

The development of big data, especially in the field of deep learning, would help investors analyze stock data. LSTM(Long Short Term Memory) is one of the artificial recurrent neural networks which is widely used nowadays and has been proven to have high accuracy in modeling. Although there are many stock predicting methods such as GBDT(Gradient Boosting Decision Tree) and ARMA(Auto Regressive Moving Average), LSTM is one of the artificial recurrent neural networks that is widely used today and has been proven to achieve high accuracy. LSTM has a feedback connection, unlike a standard feed forward neural network, it can handle not only a single data point, but also the entire data sequence as well. It contains an input gate, an output gate, and a forgetting gate to make up a typical LSTM unit. The LSTM network is highly suited for classification, processing, and prediction based on time series data since there is an undetermined duration between critical occurrences in time series. The purpose of developing LSTM is to address the issue of disappearing gradients that might

DOI: 10.54254/2754-1169/3/2022823

occur when training traditional RNN. Although there are a lot of researchers contribute themselves into this field, especially in recent years, however, there is very few papers which write reviews to make comparison and conclusion among those ideas and experiments as well as connect them together from 2015. This paper mainly fill in the research gap to review the stock predicting method based on LSTM from the year 2015 to 2022.

2. Methodologies

In 2015, Chen et al. proposed an LSTM model to anticipate Chinese market returns. Their LSTM model boosted stock return prediction accuracy from 14.3 percent to 27.2 percent when compared to the random prediction technique. In the Chinese stock market, their efforts revealed the predictive power of LSTM [1].

In 2017, Zhuge et al. proposed a model including an emotion analysis model and a long-term LSTM time series learning model. Using Shanghai Composite Index data and emotional data as input factors could greatly enhance prediction accuracy, according to the results of the experiments [2]. Zhao et al. carefully allocated weight to the data based on the temporal proximity of the data to be forecasted using the time weight function. The term "stock trend" was defined formally, with references to financial theory and best practice. The LSTM network was tweaked to detect potential time dependencies in data, outperforming previous models and allowing it to be applied to additional stock indexes. They defined the trend with 83.91 percent accuracy in the CSI300 index test [3]. Hansson, M. used the returns of three stock indexes to forecast financial time series. Among these were the S&P500 index in the United States, the BOVESPA50 index in Brazil, and the OMX30 index in Sweden. The output of the LSTM network was found to be very similar to the output of the traditional time series model [4].

In 2018, Liu et al. used an LSTM recurrent neural network to filter, extract eigenvalues, and evaluate stock data by constructing a stock trading prediction model based on stock time features and the LSTM neural network technique [5]. Baek et al. proposed a new stock market index prediction data augmentation approach using the framework. An over fitting prevention LSTM module and a prediction LSTM module were included in the framework. The performance of the model was evaluated using two sample stock market data sets: the S&P500 index and the Korean composite stock price index 200. The results showed that the model was extremely accurate in its predictions [6]. Li et al. introduced a new multi input LSTM model that could extract important information from low correlation factors while eliminating detrimental noise by adding additional input gates controlled by compelling factors called as mainstream. They also incorporated some new variables to boost prediction accuracy, such as the values of other connected equities. The results of experiments using Chinese stock market data showed that this strategy is more effective than other methods [7].

In 2019, Naik et al. used an LSTM-based recursive neural network to predict future stock returns. It use past stock returns and forecast future market returns. Using data from the National Stock Exchange of India, RNN with LSTM was utilized to keep the most recent stock information rather than the most recent relevant stock information (NSE) [8]. Li et al. developed a deep neural network based on LSTM that had three components: the LSTM, the Vader model, and the differential privacy (DP) mechanism. The suggested DP-LSTM technique had the potential to reduce prediction error while also boosting resilience. A significant number of experiments on S&P500 index stocks demonstrated that the proposed DP-LSTM increases average prediction outcomes by 0.32 percent, while the market index S&P500 prediction on MSE improved by 65.79 percent [9]. Eapen et al. suggested a new deep learning model that incorporated several convolutional neural network channels as well as two-way long-term and short-term memory units. The model improved prediction performance by 9 percent for the single channel deep learning model and by more than 6 times for the support vector regression model on the S&P500 data set. They showed multiple deep learning

DOI: 10.54254/2754-1169/3/2022823

models based on varied CNN kernel widths and the number of bidirectional LSTM units, suggesting better prediction accuracy while reducing over fitting[10]. Qian et al. examined stock time series data, used the LSTM neural network technique to predict stock data under different stationary settings, and statistically analyzed multiple experimental data sets. In addition, the LSTM algorithm was introduced to compare with the ARIMA method. The LSTM neural network prediction algorithm had great prediction accuracy and was unaffected by the stability response, according to a large number of experimental data [11]. Borovkova et al. suggested a set of LSTM neural networks for stock forecasting that used a large number of technical analysis indicators as input. The proposed integration is online and was proportionally weighted depending on a single model's recent performance, allowing them to deal with probable nonstationarity in a novel way. The model's performance was evaluated using the area under the receiver operating characteristic curve. They compared their model to LASSO and ridge logistic classifiers and tested its predicting ability on many large-cap equities in the United States. The model outperformed the benchmark model as well as the equal weight integration model [12].

In 2020, Moghar and Hamiche developed a model that used recurrent neural networks, specifically LSTM, to forecast future stock market value. The results of their algorithm for projecting the future worth of Google and NKE assets were positive [13]. Yadav et al. created an LSTM model and it was tweaked by comparing the stateless and stateful models as well as altering the number of hidden layers [14]. Mehtab et al. built four regression models based on deep learning, LSTM networks, and a new forward verification approach to improve the prediction ability of our prediction framework using the NIFTY50 index values of the national stock exchange in India (NSE). The super parameters of the LSTM model were adjusted using grid search technology to ensure that the verification loss remained steady as the epoch number rose and that verification accuracy convergence was achieved [15]. Lu et al. utilized CNN to extract features from data. Using the retrieved characteristic data, LSTM was utilized to predict the stock price. CNN-LSTM had the highest prediction accuracy and can deliver credible stock price predictions, according to the testing data [16]. Zou and Qu created and implemented the most advanced deep learning sequence models LSTM, stacked LSTM, and attention based LSTM, as well as the standard ARIMA model, to predict the next day's stock price [17]. Ding and Qin suggested a multi valued association network model based on the LSTM deep recursive neural network. It was able to predict many numbers at once, with an average accuracy of above 95 percent [18]. Jin et al. utilized empirical mode decomposition to gradually breakdown the complicated sequence of stock prices. They also proposed LSTM because it allowed them to employ a memory function to evaluate the relationship between time series data. They tweaked it and included an attention mechanism to focus on the most important data. The results of the experiments showed that the enhanced LSTM model might improve prediction accuracy while also reducing time delay. Investors' emotional tendencies were useful in enhancing prediction outcomes and the design of the model could increase the predictability of inventory series. Besides, the attention mechanism could help LSTM retrieve specific information and current task targets from huge data [19].

In 2021, Wu et al. suggested a new framework structure that merged CNN and LSTM. Stock sequence predicted by convolution LSTM was given to this new approach. It created a sequence array of historical data and leading indicators, utilized the array as the CNN framework's input pictures, extracted some feature vectors through convolution layer and pool layer, and selected ten stocks from the US and Taiwan as experimental data [20]. Lin et al. employed a hybrid model that combines LSTM and completely integrated empirical mode decomposition with adaptive noise to forecast the stock index prices of the S&P500 and the CSI300 index [21]. Ceemdan broke down the original data into numerous IMFs and a residual [21]. Gao et al. developed a novel model to improve stock forecasting. They used deep learning LASSO and principal component analysis to minimize the

DOI: 10.54254/2754-1169/3/2022823

dimensionality of the retrieved stock price's various affecting elements [22]. They also compared the stock market prediction performance of LSTM and GRU under a variety of conditions [22].

In 2022, Zhang et al. adjusted the prediction range of the LSTM network model to 1-10 days [23]. After data standardization and other pre-processing steps, they push the data to the built LSTM network model and use training and testing to determine the best settings for epoch, batch, dropout, optimizer, and other parameters. In terms of short-term prediction, results showed that the LSTM network model did not exceed linear regression, extreme gradient enhancement, and moving average [23]. Shi et al. suggested a CNN-LSTM-XGBOOST hybrid model to predict stock prices. The created model which combined a time series model with an attention mechanism, a long short memory network, and an XGBOOST regressor with a nonlinear connection to increase prediction accuracy [24].

3. Result

This paper summarizes the various literature, data, models, results for predicting stock prices using LSTM. Table 1 compares 24 LSTM models for stock forecasting that were recently deployed. It brings together multiple indexes or stocks from different countries. Generally, index data is more widely used than a single stock data, specifically, the S&P500 index in the United States and the CSI300 index in China are the most often used data. The results' correctness may be measured by Accuracy, and the errors can be measured with MSE, RMSE, MAE, MAPE, R2 and so forth. LSTM is used with other models and algorithms to improve accuracy. Different data sets, time periods, indexes, stocks, models, and approaches will decide the outcomes. As it is shown in the table, the over all accuracy of predicting is from 0.143 to 0.99582651, while the range of RMSE is 0.02004 to 97.3054. The model of Li, X. et al. achieved the highest accuracy, which is 0.99582651. Specifically, the combined algorithms of LSTM perform better than the single model, e.g., CNN or Deep learning models.

4. Conclusion

This paper has reviewed 24 papers on predicting stock prices using LSTM in detail since 2015. Data, models, and results are analyzed to evaluate existing stock models. According to the data, the CSI300 index is used by more than 5 papers in China, the S&P500 index in the United States is used by 6 papers, and the NIFTY50 index in India is used by 2 papers. To improve prediction performance, the majority of them employ a combination of LSTM and other models or algorithms. The most common models are attention-based LSTM and LSTM with RNN. Many articles used Accuracy, MSE, RMSE, MAE, MAPE to measure the accuracy. Nevertheless, each method has advantages and disadvantages, and no model can adapt to every stock market while maintaining high accuracy. There will be many more improvements to models and algorithms in the future. The main contribution of this paper is summarizing different methods, algorithms, models of LSTM in predicting stock price from 2015 to 2022. It can be concluded that LSTM can perform very high accuracy in prediction of stock price when it is combined with other algorithms such as CNN or Deep learning algorithms.

DOI: 10.54254/2754-1169/3/2022823

Table1: Recent Research of Stock Predicting Methods using LSTM.

Author	Data	Model	Result
Chen, K. et al.[1]	SSE Index in China	LSTM	Stock Returns Prediction Accuracy:14.3%~27.2%
Zhuge, Q. et al.[2]	Stocks in China	LSTM	MSE:0.000141~0.0143
Zhao, Z. et al.[3]	CSI300 index in China	LSTM	Accuracy:83.91%
Hansson, M[4]	S&P500 index in US , Bovespa50 index in Brazilian, OMX30 index in Swedish	LSTM,Deep LSTM,Softmax LSTM,Softmax deep LSTM	Accuracy:0.5000~0.5530
Liu, S. et al.[5]	CSI300 Index in China	LSTM	Accuracy:0.66-0.78
Baek, Y. et al.[6]	S&P500 index in US , KOSPI200 Index in Korea	LSTM	S&P500: MSE:54.1%,MAPE: 35.5%,MAE:32.7%; KOSPI200: MSE:48%,MAPE: 23.9%,MAE:32.7%;
Li, H. et al.[7]	CSI300 Index in China	Multi-Input LSTM	MSE:0.996(Min),1.012(Avg)
Naik, N. et al.[8]	NSE,India	RNN with LSTM	MAE:0.1645~0.1956 RMSE:23.54~25.9
Li, X. et al.[9]	S&P500 index in US	DP-LSTM	Accuracy: 0.99582651
Eapen, J. et al.[10]	S&P500 index in US	CNN,LSTM	Mean Test Score:0.000281317~0.000420501
Qian, F. et al.[11]	CSI300 index，Shenzhen200 ,Shenzhen300 index in China	LSTM	Test MAE: 0.01648~0.02227,Test RMSE:0.02004~0.02759

DOI: 10.54254/2754-1169/3/2022823

Table 1: (continued).

Author	Data	Model	Result
Borovkova, S. et al.[12]	US large-cap stocks	Stacked-LSTM	AUC Score:0.5132~0.5362
Moghar, A. et al.[13]	NKE and GOOGL	RNN with LSTM	Loss(100 epochs): GOOGL 4.97E-04; NKE:8.74E-04
Yadav, A. et al.[14]	Stocks from NIFTY50 in India	Stateful and stateless LSTMs	RMSE:5.818488~113.44813
Mehtab, S. et al.[15]	NIFTY50 in India	LSTM regression,Encoder decoder LSTM regression	RMSE/Mean: 0.0311~0.1711
Lu, W. et al.[16]	Stocks in China	CNN-LSTM	MAE:27.564,RMSE: 39.688,R2:0.9646
Zou, Z. et al.[17]	Stocks from S&P500 in US	LSTM,Attention-LSTM,Stacked-LSTM	MSE:0.002~0.163
Ding, G. et al.[18]	Stocks in China	LSTM	Accuracy: over 95%
Jin, Z. et al.[19]	AAPL	LSTM,S_LSTM,S_AM_LSTM,S_EMDAM_LSTM	Accuracy: 0.6012~0.7056 MAPE:1.65%~4.58% MAE: 2.396121~7.031646 RMSE:3.196534~8.712122 R2:0.832031~0.977388
Wu, J. M. T. et al.[20]	Ten stocks in U.S and Taiwan	CNN and LSTM	Accuracy:0.514~0.951

DOI: 10.54254/2754-1169/3/2022823

Table 1: (continued).

Author	Data	Model	Result
Lin, Y. et al.[21]	CSI300 index in China and S&P500 in US	CEEMDAN-LSTM	S&P 500: MAE:16.3634, MAPE:0.0060, MSE:436.0208, RMSE:20.8811; CSI 300: MAE:18.9512, MAPE:0.0055, MSE:615.0082, RMSE:24.8008;

DOI: 10.54254/2754-1169/3/2022823

Table 1: (continued).

Author	Data	Model	Result
Gao, Y. et al.[22]	Shanghai Composite Index in China	LASSO-LSTM,PCA-LSTM,LASSO-GRU	LASSO-LSTM: Test MSE:733.8773~9468.3418, Test RMSE:27.0902~97.3054, Test MAE:18.6517~93.4085; PCA-LSTM: Test MSE:3843.0637~100281.7656, Test RMSE:61.9924~316.6730, Test MAE:42.8904~261.6820; LASSO-GRU: Test MSE:753.7004~1196.4927, Test RMSE:27.4536~34.5904, Test MAE:19.1470~24.4835; PCA-GRU: Test MSE:6050.9692~34108.7148, Test RMSE:77.7880~184.6854, Test MAE:61.3247~166.6717;

DOI: 10.54254/2754-1169/3/2022823

Table 1: (continued).

Zhang, R. et al.[23]	Vanguard Total Stock Market Index Fund ETF in China	LSTM	RMSE:1.750 MAPE:0.633%
Shi, Z. et al.[24]	Stocks in China	SL-LSTM ML-LSTM BiLSTM CNN-BiLSTM ACNN-BiLSTM	MSE:0.00020~0.00045, RMSE:0.01424~0.02282,MAE:0.01126~0.01960,R2: 0.79434~0.88342

References

[1] *Chen, K., Zhou, Y., & Dai, F. (2015, October). A LSTM-based method for stock returns prediction: A case study of China stock market. In 2015 IEEE international conference on big data (big data) (pp. 2823-2824). IEEE.*

[2] *Zhuge, Q., Xu, L., & Zhang, G. (2017). LSTM Neural Network with Emotional Analysis for prediction of stock price. Engineering letters, 25(2).*

[3] *Zhao, Z., Rao, R., Tu, S., & Shi, J. (2017, November). Time-weighted LSTM model with redefined labeling for stock trend prediction. In 2017 IEEE 29th international conference on tools with artificial intelligence (ICTAI) (pp. 1210-1217). IEEE.*

[4] *Hansson, M. (2017). On stock return prediction with LSTM networks.*

[5] *Liu, S., Liao, G., & Ding, Y. (2018, May). Stock transaction prediction modeling and analysis based on LSTM. In 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA) (pp. 2787-2790). IEEE.*

[6] *Baek, Y., & Kim, H. Y. (2018). ModAugNet: A new forecasting framework for stock market index value with an overfitting prevention LSTM module and a prediction LSTM module. Expert Systems with Applications, 113, 457-480.*

[7] *Li, H., Shen, Y., & Zhu, Y. (2018, November). Stock price prediction using attention-based multi-input LSTM. In Asian conference on machine learning (pp. 454-469). PMLR.*

[8] *Naik, N., & Mohan, B. R. (2019, May). Study of stock return predictions using recurrent neural networks with LSTM. In International conference on engineering applications of neural networks (pp. 453-459). Springer, Cham.*

[9] *Li, X., Li, Y., Yang, H., Yang, L., & Liu, X. Y. (2019). DP-LSTM: Differential privacy-inspired LSTM for stock prediction using financial news. arXiv preprint arXiv:1912.10806.*

[10] *Eapen, J., Bein, D., & Verma, A. (2019, January). Novel deep learning model with CNN and bi-directional LSTM for improved stock market index prediction. In 2019 IEEE 9th annual computing and communication workshop and conference (CCWC) (pp. 0264-0270). IEEE.*

[11] *Qian, F., & Chen, X. (2019, April). Stock prediction based on LSTM under different stability. In 2019 IEEE 4th International Conference on Cloud Computing and Big Data Analysis (ICCCBDA) (pp. 483-486). IEEE.*

[12] *Borovkova, S., & Tsiamas, I. (2019). An ensemble of LSTM neural networks for high‐frequency stock market classification. Journal of Forecasting, 38(6), 600-619.*

[13] *Moghar, A., & Hamiche, M. (2020). Stock market prediction using LSTM recurrent neural network. Procedia Computer Science, 170, 1168-1173.*

[14] *Yadav, A., Jha, C. K., & Sharan, A. (2020). Optimizing LSTM for time series prediction in Indian stock market. Procedia Computer Science, 167, 2091-2100.*

[15] *Mehtab, S., Sen, J., & Dutta, A. (2020, October). Stock price prediction using machine learning and LSTM-based deep learning models. In Symposium on Machine Learning and Metaheuristics Algorithms, and Applications (pp. 88-106). Springer, Singapore.*

[16] *Lu, W., Li, J., Li, Y., Sun, A., & Wang, J. (2020). A CNN-LSTM-based model to forecast stock prices. Complexity, 2020.*

[17] *Zou, Z., & Qu, Z. (2020). Using LSTM in Stock prediction and Quantitative Trading. CS230: Deep Learning, Winter.*

[18] *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.*

[19] *Jin, Z., Yang, Y., & Liu, Y. (2020). Stock closing price prediction based on sentiment analysis and LSTM. Neural Computing and Applications, 32(13), 9713-9729.*

DOI: 10.54254/2754-1169/3/2022823

[20] *Wu, J. M. T., Li, Z., Herencsar, N., Vo, B., & Lin, J. C. W. (2021). A graph-based CNN-LSTM stock price prediction algorithm with leading indicators. Multimedia Systems, 1-20.*

[21] *Lin, Y., Yan, Y., Xu, J., Liao, Y., & Ma, F. (2021). Forecasting stock index price using the CEEMDAN-LSTM model. The North American Journal of Economics and Finance, 57, 101421.*

[22] *Gao, Y., Wang, R., & Zhou, E. (2021). Stock Prediction Based on Optimized LSTM and GRU Models. Scientific Programming, 2021.*

[23] *Zhang, R. (2022, March). LSTM-based Stock Prediction Modeling and Analysis. In 2022 7th International Conference on Financial Innovation and Economic Development (ICFIED 2022) (pp. 2537-2542). Atlantis Press.*

[24] *Shi, Z., Hu, Y., Mo, G., & Wu, J. (2022). Attention-based CNN-LSTM and XGBoost hybrid model for stock prediction. arXiv preprint arXiv:2204.02623.*