Enhancing Stock Price Prediction through Attention-BiLSTM and Investor Sentiment Analysis

[**Kangming Xu**](https://dl.acm.org/doi/abs/10.1145/3307339.3343254)**1\*,Biswajit Purkayastha2**

1 Computer Science and Engineering , Santa Clara University, CA, USA

**2** National Institute of Technology Silchar, Silchar, India

**rexcarry036@gmail.com**

Abstract: *The change of stock price is the focus of investors in the stock market, so stock price trend prediction has always been a hot topic in quantitative investment research. Traditional machine learning prediction model is difficult to deal with nonlinear, high frequency and high noise stock price time series, which makes the prediction accuracy of stock price trend low. In order to improve the forecasting accuracy, the temporal characteristics of stock price data are studied. A bidirectional long short-term memory neural network combining empirical mode decomposition (EMD), investor sentiment and attention mechanism is proposed to predict the rise and fall of stock prices. First, the empirical mode decomposition algorithm is used to extract the characteristics of stock price time series on different time scales, and the investor complex index of the text from the close of the last trading day to the opening of the next trading day is extracted by constructing the all-inclusive sentiment dictionary.The realization of a stock price trend prediction model based on Attention-BiLSTM involves combining the Bidirectional Long Short-Term Memory (BiLSTM) network with an attention mechanism. The BiLSTM processes data points from both past and future for better context understanding, while the attention mechanism selectively focuses on crucial information, improving the model's predictive accuracy in capturing and utilizing patterns in stock price movements. This sophisticated approach enhances the model's ability to forecast stock trends effectively.*

Keywords: Smart finance; Attention mechanism; Stock trend prediction; LSTM

# **INTRODUCTION**

In today's highly digitized financial markets, stock price forecasting has always been a hot topic in the financial field. With the continuous development of artificial intelligence and machine learning technologies, it has become possible to use these technologies to make stock price predictions. In this article, we will explore how to use BiLSTM (two-way Long Short Term Memory Network)[1-5] in Matlab to make stock price time series prediction. The stock market plays an important role in modern society, and any investor hopes to accurately predict the behavior of the market when buying and selling stocks, make the best decision and seek to maximize profits and avoid losses. Malkiel et al. 's efficient market hypothesis points out that the stock market is an effective information market, and if the trading data of the stock market can be effectively processed and appropriate algorithms used, the trend of its rise and fall can be predicted: however, due to the high noise of the stock price time series, its trend is non-linear, non-stable and chaotic, and accurate prediction of the stock trend is still an important task A challenging task.

1. **RELATED WORK**

From the perspective of the development process of stock forecasting, it is transitioning from the use of traditional statistical methods to neural networks and artificial intelligence methods. AutoRegressive Integrated Moving Average, autoregressive integrated moving Average, ARIMA[5] and grey prediction model are commonly used time series models in statistics, which are often used in early stock prediction. However, when traditional statistical methods are used in stock data with large amount of data and large fluctuations, they fail to achieve good prediction results[6]. With the development of big data and artificial intelligence technology, neural network, machine learning, deep learning and other technologies are widely used in the financial field. Zhang et al. used BackPropagation (BP) neural network to predict stock price and trend, and concluded that BP neural network was better than deep learning fuzzy algorithm in predicting stock price and trend. Jaiwang et al. used Support Vector MachineSVM to predict stock buying and selling points, indicating that [7]SVM has a good performance in stock prediction. Zhang et al. use Adaptive BoostingAdaBoost integrated algorithm to predict the annual return of stocks, which shows that machine learning algorithm has good performance in stock prediction. Zhang Xiao et al used Gradient Boosting Decision Tree (GBDT) to predict stock trend, and the results show that GBDT model has better prediction effect than linear regression and random forest model[8]. Shallow machine learning algorithms have simple structures, poor generalization ability, are prone to falling into local optimal solutions "and are limited when working with raw format data." Deep learning learns effective feature representations from large amounts of input data by forming highly abstract high-level features through simple but nonlinear modules. Recurrent Neural networks (RNNS), which represent deep learning, can consider short-term correlations in time series. The hidden layer does not only receive the current data. It also receives previous data information, so it can theoretically use data of any length of time: however, [9-11]RNNS are prone to gradient disappearance and gradient explosion when learning sequences, resulting in their inability to grasp nonlinear relationships over long time spans. Based on this, Hlochreiter proposed a Long Short-Term MemoryLSTM network to solve the problems existing in RNN to a large extent through the gate mechanism[4]. Chen et al. used LSTM to forecast the return rate of Chinese stock market and found that ISTM had the best forecasting effect.

**2.1 BiLSTM neural network**

LSTM neural network is a special recurrent neural network. Compared with RNN, LSTM neural network has a unique threshold management mechanism, which is composed of input gate, forget gate and output gate. It is also a deep learning model that combines features of long short-term memory networks (LSTM) and bidirectional recurrent neural networks (BRNN)[12]. LSTM is a special type of recurrent neural network (RNN) that is able to process time series data efficiently and is able to capture long-term dependencies. Bidirectional recurrent neural networks can learn from both past and future information to better understand the dynamic properties of time series data.



**Figure 1:** LSTM structure model

An LSTM cell has three doors, which are called the memory door (f door), the input door (i door), and the output door (o door). Note that the output ot of the output gate is not the final output of the LSTM cell[13]. The final output of the LSTM cell is ht and ct.

These three doors are the three yellow boxes marked sigma in the picture above. The sigmoid layer outputs a value of 0-1 indicating how much information to let through.

**2.2 BiLSTM carries out stock price prediction mechanism**

When using BiLSTM to forecast stock prices, we need to prepare the time series data of stock prices. The data typically includes information about a stock's opening, closing, high and low prices, as well as volume. We can use the data processing tools in Matlab to pre-process and feature engineering these data for input into the BiLSTM model for training. But once the data is ready, we can use the deep learning toolkit in Matlab to build and train BiLSTM models. When building a model, we need to consider the structure of the model, the number of layers, and the choice of hyperparameters[14-17]. By adjusting these parameters reasonably, we can get a prediction model that can fully exploit the features of time series data without over-fitting.

After realizing the stock trend prediction under the attention mechanism, we need to evaluate and optimize the model. We can use some common evaluation indicators such as mean square error (MSE), root mean square error (RMSE), and mean absolute error (MAE) [18]to evaluate the predictive performance of the model. At the same time, we can also optimize the performance of the model by adjusting the hyperparameters of the model, such as learning rate, batch size, etc.

# **Methodology**

In this paper, we mainly discuss how to use BiLSTM model in Matlab to predict stock price time series. By properly preparing data, building models, and training optimizations, we can use this powerful deep learning technique to more accurately predict stock prices[19]. However, it is important to note that stock prices are affected by a variety of factors, and the market can also change dramatically, so there is a certain amount of uncertainty in any forecast. It provides some reference and inspiration for readers who are interested in stock price prediction. The continuous development of deep learning technology will bring more possibilities for stock price forecasting, and we look forward to being able to do so in the future

**3.1 Attention mechanism BiLSTM model algorithm**

In the process of stock trend prediction, BiLSTM is composed of forward LSTM and backward LSTM. In the forward LSTM layer, the input time series is entered into the LSTM model in the original order. In the backward LSTM layer, the input time series is entered into the LSTM model in reverse order. This structure can extract the bidirectional relationship of the time series and connect the two LSTMS to the same output layer. Therefore, the theoretical prediction performance should be better than that of one-way LSTM, and the specific expression of BiLSTM is shown below[20].

(1)

Where σ is the activation function and Ht is the hidden layer input. By updating the forward structure and

The reverse structure takes the final input:



**Figure 2:** The algorithm combines BiLSTM structure

**3.2 Attention mechanism**

The Attention mechanism is a mechanism that simulates the Attention of the human brain to focus on and learn important information. The attention mechanism can be used to assign different weights to the hidden layer states in the end of the neural net according to the influence of each input feature on the output. The Attention mechanism is used to selectively pay attention to the influence of different input components on the prediction results, so as to change the effect of stock prediction[21]. The main principle is as follows: By keeping the intermediate output results of each step when the BiLSTM layer models the input sequence, and associating them with the values of the output sequence, the model is trained to selectively pay attention to the input components during learning, and assign higher weights to the input components with higher correlation. The specific principle is shown in Figure 3:

****

**Figure 3:**Attention mechanism diagram

The hidden layer state of BiLSTM is [h1,h2,...hn]T.

**3.3 Data analysis and results**

Use investor sentiment to add new features to the model input. Investor sentiment is to identify and summarize the emotional words in the text data, and then judge the investor sentiment expressed in the text. [22]The current methods of text analysis mainly include machine learning and sentiment analysis based on sentiment dictionary. Sentiment analysis methods based on machine learning generally build classifiers based on supervised machine learning algorithms to achieve sentiment discrimination. The classification accuracy of this method is relatively high. Therefore, emotion analysis based on emotion dictionary corresponds words in emotion dictionary with words after text segmentation to obtain negative and positive categories, and then adds words by rules to obtain text emotion tendency score. [23]However, the common emotion analysis based on emotion dictionary does not take into account the influence of modifiers in the text on emotion words. Therefore, this paper conducts emotion analysis based on emotion dictionary, constructs auxiliary weight dictionary when designing emotion dictionary, and optimizes the basic weight of words.



**Figure 4:**Training data and test data

After completing the basic prediction data model, the next step is to normalize the data, and build the model through keras. First, the input layer and Reshape layer are used to convert the input data into a specified input form, where each input data is in the form of a 7\*5\*1 2D image, with 7 representing width and 5 representing height. 1 indicates the number of channels. Then convolution operation is carried out through a 2D Conv layer, the number of filters is 64, padding is set to same to obtain the feature map of the same size, and the activation function is relu. [24-28]Downsampling is then performed by a Maxpooling, followed by a Dropout to prevent overfitting. Then, two LSTM layers are connected, and time sequence modeling is carried out from the dimension of time step. Finally, the predicted value of the next time is output through the full connection layer. The loss function of the model was selected as mean square error, and the optimization method was adam optimizer.



**Figure 5:** Data prediction effect and RMSE

From the experimental results, it can be seen that the simple CNN-LSTM model built by us can achieve good data fitting effect and achieve low RMSE.

**Conclusion**

This paper investigates the nonlinear, non-stationary, and chaotic prediction errors commonly associated with stock time series. To address these challenges, we propose an innovative AttentionBiLSTM model that integrates empirical mode decomposition (EMD) and investor sentiment analysis[29-30]. By utilizing EMD, we decompose closing prices to extract new features and identify trends across various time scales. This decomposition allows us to enhance the dataset with meaningful representations of price movements, enabling a more nuanced understanding of market dynamics. The sentiment scores, derived from textual data related to investor emotions, are computed and incorporated into the model, facilitating a comprehensive approach to stock price prediction.

In our model, we apply weight optimization techniques to effectively integrate these newly derived features into the BiLSTM framework. [31-33]The bidirectional capabilities of BiLSTM allow for robust time series feature learning, which is further enhanced by the incorporation of an Attention mechanism. This mechanism enables the model to selectively focus on the most influential components affecting prediction outcomes, thereby improving the accuracy of stock market forecasts. Our experimental results demonstrate that the proposed model significantly outperforms traditional forecasting methods by leveraging EMD[34], investor sentiment, and Attention mechanisms, thereby offering a more reliable tool for anticipating stock price movements.

Current stock forecasting research often relies heavily on basic market indicators while neglecting the critical influence of investor sentiment and behavior. [35-37]Our model addresses this gap by extracting in-depth information from fundamental market indicators and analyzing the sentiment that drives investor decision-making. By learning bidirectional time series features and applying a selective attention strategy, our approach yields enhanced prediction outcomes. Looking ahead, future research can explore the integration of additional variables, such as stock market announcements, regulatory policies, and other technical indicators, to further refine the model’s predictive accuracy and provide deeper insights for investors navigating the complexities of the stock market[38-40].

**REFERENCES**

1. Zhang, M., Yuan, B., Li, H., & Xu, K. (2024). LLM-Cloud Complete: Leveraging Cloud Computing for Efficient Large Language Model-based Code Completion. Journal of Artificial Intelligence General science (JAIGS) ISSN: 3006-4023, 5(1), 295-326.
2. Li, P., Hua, Y., Cao, Q., & Zhang, M. (2020, December). Improving the Restore Performance via Physical-Locality Middleware for Backup Systems. In Proceedings of the 21st International Middleware Conference (pp. 341-355).
3. Zhou, S., Yuan, B., Xu, K., Zhang, M., & Zheng, W. (2024). THE IMPACT OF PRICING SCHEMES ON CLOUD COMPUTING AND DISTRIBUTED SYSTEMS. Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online), 3(3), 193-205.
4. Shang, F., Zhao, F., Zhang, M., Sun, J., & Shi, J. (2024). Personalized Recommendation Systems Powered By Large Language Models: Integrating Semantic Understanding and User Preferences. International Journal of Innovative Research in Engineering and Management, 11(4), 39-49.
5. Sun, J., Wen, X., Ping, G., & Zhang, M. (2024). Application of News Analysis Based on Large Language Models in Supply Chain Risk Prediction. Journal of Computer Technology and Applied Mathematics, 1(3), 55-65.
6. Zhao, F., Zhang, M., Zhou, S., & Lou, Q. (2024). Detection of Network Security Traffic Anomalies Based on Machine Learning KNN Method. Journal of Artificial Intelligence General science (JAIGS) ISSN: 3006-4023, 1(1), 209-218.
7. Feng, Y., Qi, Y., Li, H., Wang, X., & Tian, J. (2024, July 11). Leveraging federated learning and edge computing for recommendation systems within cloud computing networks. In Proceedings of the Third International Symposium on Computer Applications and Information Systems (ISCAIS 2024) (Vol. 13210, pp. 279-287). SPIE.
8. Zhao, F.; Li, H.; Niu, K.; Shi, J.; Song, R. Application of Deep Learning-Based Intrusion Detection System (IDS) in Network Anomaly Traffic Detection. Preprints 2024, 2024070595.
9. Gong, Y., Liu, H., Li, L., Tian, J., & Li, H. (2024, February 28). Deep learning-based medical image registration algorithm: Enhancing accuracy with dense connections and channel attention mechanisms. Journal of Theory and Practice of Engineering Science, 4(02), 1-7.
10. Liu, Y., Tan, H., Cao, G., & Xu, Y. (2024). Enhancing User Engagement through Adaptive UI/UX Design: A Study on Personalized Mobile App Interfaces.
11. Huang, D., Yang, M., Wen, X., Xia, S., & Yuan, B. (2024). AI-Driven Drug Discovery: Accelerating the Development of Novel Therapeutics in Biopharmaceuticals. Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online), 3(3), 206-224.
12. Yang, M., Huang, D., Zhang, H., & Zheng, W. (2024). AI-Enabled Precision Medicine: Optimizing Treatment Strategies Through Genomic Data Analysis. Journal of Computer Technology and Applied Mathematics, 1(3), 73-84.
13. Wen, X., Shen, Q., Zheng, W., & Zhang, H. (2024). AI-Driven Solar Energy Generation and Smart Grid Integration A Holistic Approach to Enhancing Renewable Energy Efficiency. International Journal of Innovative Research in Engineering and Management, 11(4), 55-55.
14. Lou, Q. (2024). New Development of Administrative Prosecutorial Supervision with Chinese Characteristics in the New Era. Journal of Economic Theory and Business Management, 1(4), 79-88.
15. Liu, Y., Tan, H., Cao, G., & Xu, Y. (2024). Enhancing User Engagement through Adaptive UI/UX Design: A Study on Personalized Mobile App Interfaces.
16. Xu, H., Li, S., Niu, K., & Ping, G. (2024). Utilizing Deep Learning to Detect Fraud in Financial Transactions and Tax Reporting. Journal of Economic Theory and Business Management, 1(4), 61-71.
17. Li, S., Xu, H., Lu, T., Cao, G., & Zhang, X. (2024). Emerging Technologies in Finance: Revolutionizing Investment Strategies and Tax Management in the Digital Era. Management Journal for Advanced Research, 4(4), 35-49.
18. Shi J, Shang F, Zhou S, et al. Applications of Quantum Machine Learning in Large-Scale E-commerce Recommendation Systems: Enhancing Efficiency and Accuracy[J]. Journal of Industrial Engineering and Applied Science, 2024, 2(4): 90-103.
19. Wang, S., Zheng, H., Wen, X., & Fu, S. (2024). DISTRIBUTED HIGH-PERFORMANCE COMPUTING METHODS FOR ACCELERATING DEEP LEARNING TRAINING. Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online), 3(3), 108-126.
20. Lei, H., Wang, B., Shui, Z., Yang, P., & Liang, P. (2024). Automated Lane Change Behavior Prediction and Environmental Perception Based on SLAM Technology. arXiv preprint arXiv:2404.04492.
21. Wang, B., Zheng, H., Qian, K., Zhan, X., & Wang, J. (2024). Edge computing and AI-driven intelligent traffic monitoring and optimization. Applied and Computational Engineering, 77, 225-230.
22. Xu, Y., Liu, Y., Xu, H., & Tan, H. (2024). AI-Driven UX/UI Design: Empirical Research and Applications in FinTech. International Journal of Innovative Research in Computer Science & Technology, 12(4), 99-109.
23. Li, H., Wang, S. X., Shang, F., Niu, K., & Song, R. (2024). Applications of Large Language Models in Cloud Computing: An Empirical Study Using Real-world Data. International Journal of Innovative Research in Computer Science & Technology, 12(4), 59-69.
24. Xu, L., Gong, C., Jiang, G., & Yang, H. (2024, June). A STUDY ON PERSONALIZED WEB PAGE RECOMMENDATION BASED ON NATURAL LANGUAGE PROCESSING AND ITS IMPACT ON USER BROWSING BEHAVIOR. In The 24th International scientific and practical conference “Technologies of scientists and implementation of modern methods”(June 18–21, 2024) Copenhagen, Denmark. International Science Group. 2024. 431 p. (p. 317).
25. Ping, G., Wang, S. X., Zhao, F., Wang, Z., & Zhang, X. (2024). Blockchain Based Reverse Logistics Data Tracking: An Innovative Approach to Enhance E-Waste Recycling Efficiency.
26. Xu, H., Niu, K., Lu, T., & Li, S. (2024). Leveraging artificial intelligence for enhanced risk management in financial services: Current applications and future prospects. Engineering Science & Technology Journal, 5(8), 2402-2426.
27. Shi, Y., Shang, F., Xu, Z., & Zhou, S. (2024). Emotion-Driven Deep Learning Recommendation Systems: Mining Preferences from User Reviews and Predicting Scores. Journal of Artificial Intelligence and Development, 3(1), 40-46.
28. Xu, Qiming, et al. "Artificial Intelligence in Risk Protection for Financial Payment Systems." The 24th International scientific and practical conference “Technologies of scientists and implementation of modern methods”(June 18–21, 2024) Copenhagen, Denmark. International Science Group. 2024. 431 p.. 2024.
29. Wang, Shikai, Kangming Xu, and Zhipeng Ling. "Deep Learning-Based Chip Power Prediction and Optimization: An Intelligent EDA Approach." International Journal of Innovative Research in Computer Science & Technology 12.4 (2024): 77-87.
30. Ping, G., Zhu, M., Ling, Z., & Niu, K. (2024). Research on Optimizing Logistics Transportation Routes Using AI Large Models. Applied Science and Engineering Journal for Advanced Research, 3(4), 14-27.
31. Shang, F., Shi, J., Shi, Y., & Zhou, S. (2024). Enhancing E-Commerce Recommendation Systems with Deep Learning-based Sentiment Analysis of User Reviews. International Journal of Engineering and Management Research, 14(4), 19-34.
32. Huang, S., Diao, S., Zhao, H., & Xu, L. (2024, June). The Contribution of Federated Learning to AI Development. In The 24th International scientific and practical conference “Technologies of scientists and implementation of modern methods”(June 18–21, 2024) Copenhagen, Denmark. International Science Group. 2024. 431 p. (p. 358).
33. Zhan, X., Shi, C., Li, L., Xu, K., & Zheng, H. (2024). Aspect category sentiment analysis based on multiple attention mechanisms and pre-trained models. Applied and Computational Engineering, 71, 21-26.
34. Liang, P., Song, B., Zhan, X., Chen, Z., & Yuan, J. (2024). Automating the training and deployment of models in MLOps by integrating systems with machine learning. Applied and Computational Engineering, 67, 1-7.
35. Li, J., Wang, Y., Xu, C., Liu, S., Dai, J., & Lan, K. (2024). Bioplastic derived from corn stover: Life cycle assessment and artificial intelligence-based analysis of uncertainty and variability. Science of The Total Environment, 174349.
36. Xiao, J., Wang, J., Bao, W., Deng, T., & Bi, S. (2024). Application progress of natural language processing technology in financial research. Financial Engineering and Risk Management, 7(3), 155-161.
37. Wu, B., Gong, Y., Zheng, H., Zhang, Y., Huang, J., & Xu, J. (2024). Enterprise cloud resource optimization and management based on cloud operations. Applied and Computational Engineering, 67, 8-14.
38. Zhang, Y., Liu, B., Gong, Y., Huang, J., Xu, J., & Wan, W. (2024). Application of machine learning optimization in cloud computing resource scheduling and management. Applied and Computational Engineering, 64, 9-14.
39. Yu, K., Bao, Q., Xu, H., Cao, G., & Xia, S. (2024). An Extreme Learning Machine Stock Price Prediction Algorithm Based on the Optimisation of the Crown Porcupine Optimisation Algorithm with an Adaptive Bandwidth Kernel Function Density Estimation Algorithm.
40. Li A, Zhuang S, Yang T, Lu W, Xu J. Optimization of logistics cargo tracking and transportation efficiency based on data science deep learning models. Applied and Computational Engineering. 2024 Jul 8;69:71-7.