Academic Journal of Sociology and Management
Journal Home: http://ajsm.suaspress.org/ | CODEN: AJSMAS
Vol. 2, No. 6, 2024 | ISSN 3005-5040 (Print) | ISSN 3005-5059 (Online)

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

XU, Kangming [1*] PURKAYASTHA, Biswajit [2]

[1] Santa Clara University, USA
[2] Cihan University Sulaimaniya, Iraq

**** XU, Kangming is the corresponding author, E-mail: etekedahibi@outlook.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: Financial Management Methods, Smart Finance, Attention Mechanism, Stock Trend Prediction, LSTM.

Disciplines: Finance. **Subjects:** Financial Management.

DOI: https://doi.org/10.5281/zenodo.14065931 **ARK:** https://n2t.net/ark:/40704/AJSM.v2n6a03

1 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.

2 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

Academic Journal of Sociology and Management
Journal Home: http://ajsm.suaspress.org/ | CODEN: AJSMAS
Vol. 2, No. 6, 2024 | ISSN 3005-5040 (Print) | ISSN 3005-5059 (Online)

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 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.

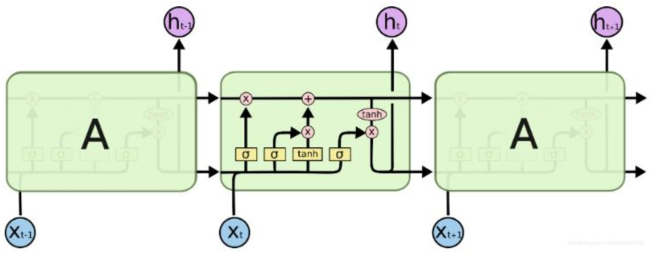

FIGURE 1. LSTM STRUCTURE MODEL

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 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]. 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) [15]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.

3 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. 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,

Academic Journal of Sociology and Management
Journal Home: http://ajsm.suaspress.org/ | CODEN: AJSMAS
Vol. 2, No. 6, 2024 | ISSN 3005-5040 (Print) | ISSN 3005-5059 (Online)

and the specific expression of BiLSTM is shown below.

$$\overrightarrow{h}_t = \sigma(\overrightarrow{W}_{xh} x_t + \overrightarrow{W}_{hh} \overrightarrow{h}_{t-1} + \overrightarrow{b}_h)$$
$$\overleftarrow{h}_t = \sigma(\overleftarrow{W}_{xh} x_t + \overleftarrow{W}_{hh} \overleftarrow{h}_{t-1} + \overleftarrow{b}_h)$$
$$H_t = \overrightarrow{W}_{xh} \overrightarrow{h} + \overleftarrow{W}_{hy} \overleftarrow{h} + b_y \quad (1)$$

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:

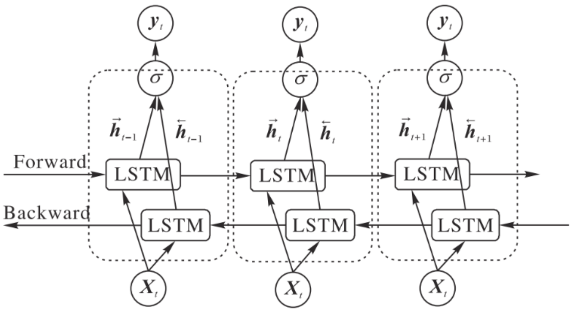

FIGURE 2. THE ALGORITHM COMBINES BILSTM STRUCTURE

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. 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:

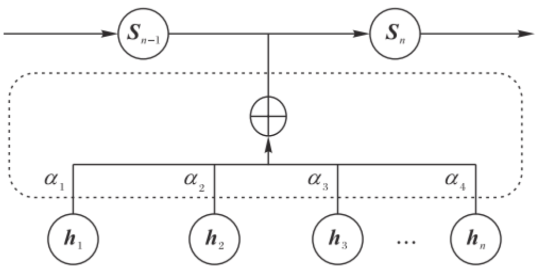

FIGURE 3. ATTENTION MECHANISM DIAGRAM

The hidden layer state of BiLSTM is $[h_1, h_2, ... h_n]^T$.

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. 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. 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.

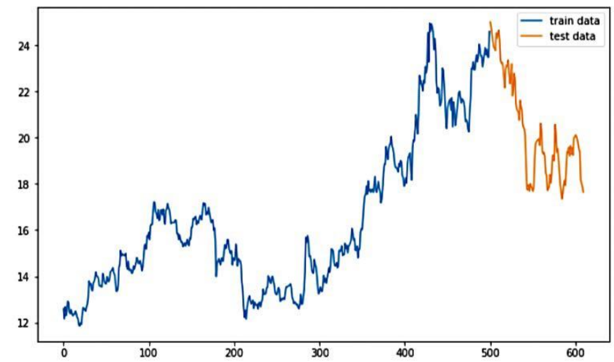

FIGURE 4. TRAINING DATA AND TEST DATA

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. 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.

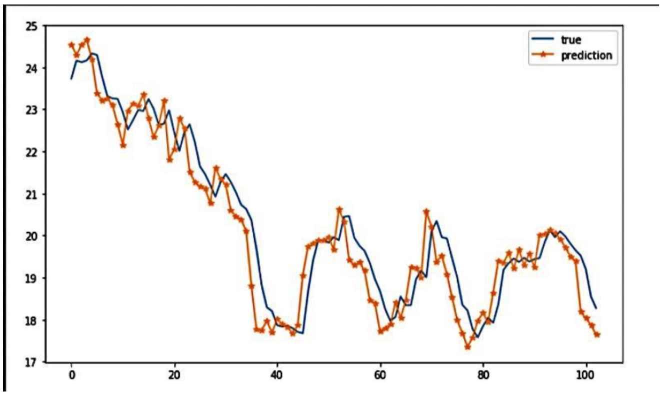

FIGURE 5. DATA PREDICTION EFFECT AND RMSE

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.

4 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. 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, 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. 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.

ACKNOWLEDGMENTS

The authors thank the editor and anonymous reviewers for their helpful comments and valuable suggestions.

FUNDING

Not applicable.

INSTITUTIONAL REVIEW BOARD STATEMENT

Not applicable.

INFORMED CONSENT STATEMENT

Not applicable.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

PUBLISHER'S NOTE

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

AUTHOR CONTRIBUTIONS

Not applicable.

ABOUT THE AUTHORS

XU, Kangming

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

PURKAYASTHA, Biswajit

Department of Computer Science, Cihan University Sulaimaniya, Sulaimaniya, Iraq.

REFERENCES

[1] Ma, X., Wang, J., Ni, X., & Shi, J. (2024). Machine Learning Approaches for Enhancing Customer Retention and Sales Forecasting in the Biopharmaceutical Industry: A Case Study. International Journal of Engineering and Management Research, 14(5), 58-75.

[2] Li, H., Sun, J., & Ke, X. (2024). AI-Driven Optimization System for Large-Scale Kubernetes Clusters: Enhancing Cloud Infrastructure Availability, Security, and Disaster Recovery. Journal of Artificial Intelligence General science (JAIGS) ISSN: 3006-4023, 2(1), 281-306.

[3] Xia, S., Wei, M., Zhu, Y., & Pu, Y. (2024). AI-Driven Intelligent Financial Analysis: Enhancing Accuracy and Efficiency in Financial Decision-Making. Journal of Economic Theory and Business Management, 1(5), 1-11.

[4] Zhang, H., Lu, T., Wang, J., & Li, L. (2024). Enhancing Facial Micro-Expression Recognition in Low-Light Conditions Using Attention-guided Deep Learning. Journal of Economic Theory and Business Management, 1(5), 12-22.

[5] Wang, J., Lu, T., Li, L., & Huang, D. (2024). Enhancing Personalized Search with AI: A Hybrid Approach Integrating Deep Learning and Cloud Computing. International Journal of Innovative Research in Computer Science & Technology, 12(5), 127-138.

[6] Che, C., Huang, Z., Li, C., Zheng, H., & Tian, X. (2024). Integrating generative ai into financial market prediction for improved decision making. arXiv preprint arXiv:2404.03523.

[7] Li, L., Zhang, Y., Wang, J., & Ke, X. (2024). Deep Learning-Based Network Traffic Anomaly Detection: A Study in IoT Environments.

[8] Cao, G., Zhang, Y., Lou, Q., & Wang, G. (2024). Optimization of High-Frequency Trading Strategies Using Deep Reinforcement Learning. Journal of Artificial Intelligence General science (JAIGS) ISSN: 3006-4023, 6(1), 230-257.

[9] Wang, G., Ni, X., Shen, Q., & Yang, M. (2024). Leveraging Large Language Models for Context-Aware Product Discovery in E-commerce Search Systems. Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online), 3(4).

[10] Li, H., Wang, G., Li, L., & Wang, J. (2024). Dynamic Resource Allocation and Energy Optimization in Cloud Data Centers Using Deep Reinforcement Learning. Journal of Artificial Intelligence General science (JAIGS) ISSN: 3006-4023, 1(1), 230-258.

[11] Che, C., Zheng, H., Huang, Z., Jiang, W., & Liu, B. (2024). Intelligent robotic control system based on computer vision technology. arXiv preprint arXiv:2404.01116.

[12] Zheng, H.; Wu, J.; Song, R.; Guo, L.; Xu, Z. Predicting Financial Enterprise Stocks and Economic Data Trends Using Machine Learning Time Series Analysis. Applied and Computational Engineering 2024, 87, 26–32.

[13] Ju, C., & Zhu, Y. (2024). Reinforcement Learning‐Based Model for Enterprise Financial Asset Risk Assessment and Intelligent Decision‐Making.

[14] Huang, D., Yang, M., & Zheng, W. (2024). Integrating AI and Deep Learning for Efficient Drug Discovery and Target Identification.

[15] Yang, M., Huang, D., & Zhan, X. (2024). Federated Learning for Privacy-Preserving Medical Data Sharing in Drug Development.