ISSN: 2688-9323

Deep Learning and Variational Modal Decomposition in Stock Price Prediction

Juan Liu*, Wei Huang, Pingping Kong

College of Science, Hunan Institute of Technology, Hengyang 421002, China

*Corresponding Author: Juan Liu (Email: hnhylj@hnit.edu.cn)

Abstract

This study explores stock price forecasting, a critical topic for economic stability and investor decision-making. Traditional models like ARIMA struggle with stock market complexity due to their linear assumptions. To address this, the study examines advanced methods, focusing on deep learning techniques such as CNNs and LSTMs for their predictive strengths. It proposes a hybrid model combining Variational Mode Decomposition (VMD) and Bi-directional Long Short-Term Memory Networks (BiLSTM). VMD reduces time series non-stationarity, while BiLSTM captures sequence features via bi-directional processing. Empirical results show the VMD-BiLSTM model outperforms others in RMSE, MAE, and R^2 metrics, achieving higher forecasting accuracy. Although performance decreases during extreme price fluctuations, it effectively captures main trends. This research highlights the practical value of deep learning in handling complex financial time series and provides innovative methods for stock price prediction.

Keywords

Long Short-Term Memory Networks (LSTM); Bidirectional Long Short-Term Memory Networks (BiLSTM); Variational Modal Decomposition (VMD); Stock Price Forecasting.

1. Introduction

Stock markets play a key role in the stability of the global economy, and their price forecasts have far-reaching implications for investors and the stability of national economies. Global research confirms the predictability of stock prices, which provides a way to understand the complex dynamics of economic cornerstones. At the same time, policies, economic indicators, market volatility, and investor strategies are intertwined with each other and work together to influence stock prices, providing an important analytical perspective for economic decision-makers. As early as 1976, Box, G. E P et al. authored the landmark book in which traditional econometric models such as the AR model, the MA model, and the ARIMA model were systematically described. And due to the complexity of the market, the traditional models are not able to handle these complexities effectively due to the limitations of the linearity assumptions. As a result, researchers have turned to machine learning and deep learning, such as neural networks, random forests (RF), and support vector machines (SVR), to cope with nonlinear and nonsmooth problems.

Deep learning has become a hotspot for scientific research and innovation in the field of stock prediction. 2023 Wu M T et al. combined convolutional neural network (CNN) and long short-term memory (LSTM) models, showing that this hybrid model outperforms the traditional model in prediction performance, with statistically significant prediction accuracies. Yuxu Feng et al. predicted stock prices and compared LSTM, SVR, and Adaboost, confirming the superiority of LSTM in predictive accuracy and handling time series complexity. Banerjee Soham et al. constructed a multivariate hierarchical prediction framework involving deep learning models

ISSN: 2688-9323

such as MLP, LSTM, GRU, BiLSTM, and BiGRU for a comprehensive understanding of stock price dynamics.

Hybrid models are used in problems such as traffic flow, power load forecasting, etc., and a number of scholars use the SSA algorithm, PSO algorithm, and GWO algorithm to improve the LSTM model and BiLSTM model. VMD is widely used in the field of water and electricity, with excellent prediction results. An innovative stock price prediction method is a key challenge in the field of finance; few scholars have introduced VMD into stock prediction problems, so this paper combines VMD with the BiLSTM model to predict stock prices.

2. Model Introduction

2.1. Long Short-Term Memory Network (LSTM) Model

The LSTM model is an important method for deep learning and is particularly suitable for dealing with time series tasks. As a special form of recurrent neural network, it introduces a system of gated units, including input gates, forgetting gates and output gates[1-2]. These gates control the update and output of the memory unit state through current and historical data, selectively controlling the information, forgetting the old information and updating the unit state based on the new information.The structure of the LSTM model is shown in Figure 1.

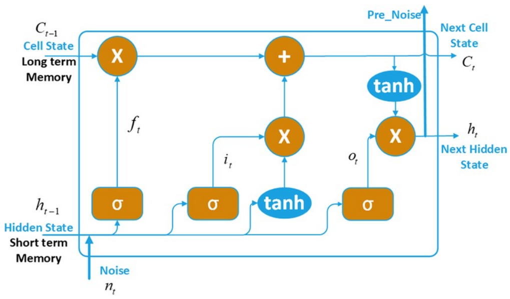

Figure 1. LSTM structure(Images from the web)

2.2. Bidirectional Long Short-Term Memory Network (BiLSTM) Modeling

The BiLSTM model, which introduces two LSTM models with bi-directional processing capability, adjusts the network weights through forward propagation and back propagation algorithms[3-5]. Combining the information from both directions provides a more comprehensive representation of the sequence features, which can better predict the stock price. The BiLSTM network structure is shown in Figure 2.

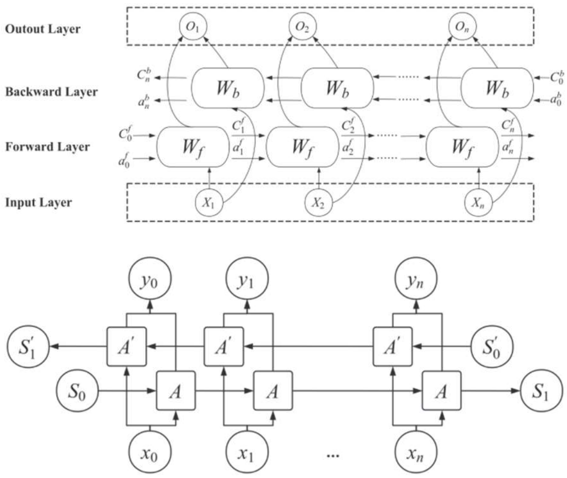

Figure 2. BiLSTM network structure diagram

ISSN: 2688-9323

2.3. Variational Modal Decomposition

VMD is an adaptive, non-recursive signal decomposition method. It is robust to anomalies and noise.

1) The variational solution process decomposes the original data into K modal components to minimize the sum of the estimated bandwidth of each modal component[6]. Its constrained variational model:

$$\min_{\{u_k\},\{\omega_k\}} = \{\sum_k \left\| \partial_t [(\delta(t) + \frac{j}{\pi t}) * u_k(t)] e^{-j\omega_k t} \right\|_2^2 \} \tag{1}$$

$$s.t \sum_{k=1}^{K} u_k(t) = y(t) \tag{2}$$

where K is the number of intrinsic modal functions (IMFs), is the set of IMFs, is the set of center frequencies, is the impulse function, is the convolution operator, is the imaginary unit, and is the original series.

1) Introducing quadratic penalty factors and Lagrange multipliers:

$$L(\{u_k\},\{\omega_k\},\lambda) = \sigma \sum_k \left\| \partial_t [(\delta(t) + \frac{j}{\pi t}) * u_k(t)] e^{-j\omega_k t} \right\|_2^2 + < \lambda(t), y(t) - \sum_k u_k(t) > \tag{3}$$

2) Iterate using the alternating multiplier direction method to find the saddle point:

$$u_k^{n+1}(\omega) = \frac{z(\omega) - \sum_{i<k} u^{n+1}(\omega) - \sum_{i>k} u^n(\omega) + \frac{\lambda^n(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k^n)^2} \tag{4}$$

$$w_k^{n+1} = \frac{\int_0^\infty w \left| u_k^{n+1}(w) \right|^2 dw}{\int_0^\infty \left| u_k^{n+1}(w) \right|^2 dw} \tag{5}$$

3) Stop updating when the following (equation 6) is satisfied.

$$\sum_{k=1}^{K} \frac{\left\| u_k^{n+1} - u_k^n \right\|_2^2}{\left\| u_k^n \right\|_2^2} < \varepsilon \tag{6}$$

2.4. Model Construction

In order to further explore the time series features, a stock price prediction model based on VMD-BiLSTM is established, and the prediction process is as follows:

1) Use the VMD method to extract the time series features of stock price data separately and reduce the non-stationarity of the time series.

2) Use BiLSTM to obtain the optimal solution, according to which the VMD-BiLSTM model is established.

3) Input the sub-series into the BiLSTM model for prediction and superimpose the prediction results of each component to get the final prediction value.

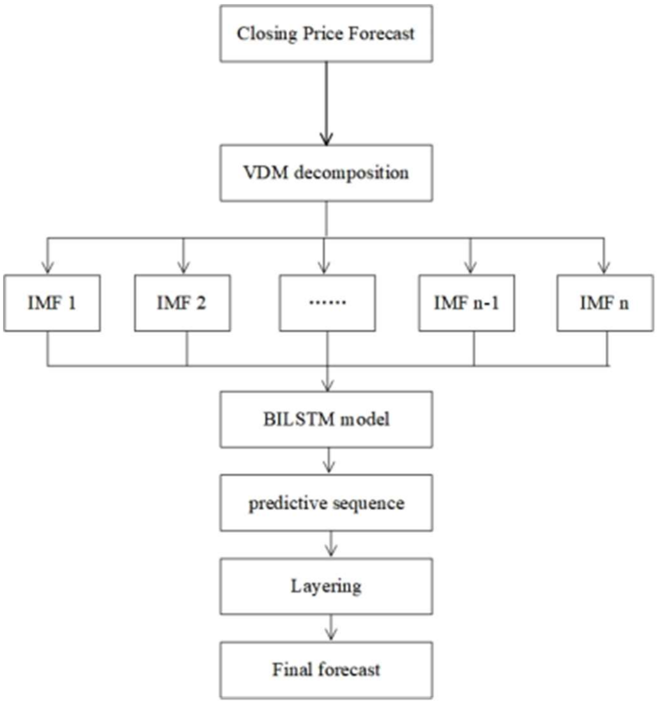

Figure 3. VMD-BILSTM algorithmic process

2.5. Evaluation Indicators

When evaluating the predictive performance of a model, we typically rely on several key statistical metrics that quantify the difference between the model's predicted values and the actual observed values. First, the Mean Absolute Error (MAE) provides an intuitive measure of error by calculating the average of the absolute values of the difference between the predicted and true values. Second, the Root Mean Square Error (RMSE) further takes into account the square of the error, which allows larger errors to be given more weight in the calculation, thus allowing for a more sensitive reflection of the model's performance on outliers. Finally, the coefficient of determination (R2) is a normalized metric that measures the ratio of variability explained by the model to the total variability of the data [7-10]. Together, these metrics form a multidimensional perspective for assessing model performance, allowing us to comprehensively understand and improve the predictive power of the model.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \tag{7}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{8}$$

$$R^2 = 1 - \frac{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \tag{9}$$

ISSN: 2688-9323

3. Experiments and Analysis of Results

3.1. Data Set and Experimental Setup

In this paper, the stock price data of Tanaka Precision Machinery (300461) from January 2, 2020, to June 28, 2024, is selected. The parameter settings of the VMD-BILSTM model for this experiment are shown in Table 1.

Table 1. Model Parameter Setting

Parameter	Parameter
Learning rate	0.001
Number of neurons in hidden layer	10
Regularization coefficient	0.0001
Time step	10

3.2. Results and Analysis

In this paper, we use the VMD method to decompose the data into six subsequences to improve the prediction accuracy.The VMD decomposition results are shown in Fig. As can be seen from the figure, the smoothness of the components obtained after the secondary decomposition by VMD is relatively good, which is conducive to improving the accuracy of the subsequent model prediction.

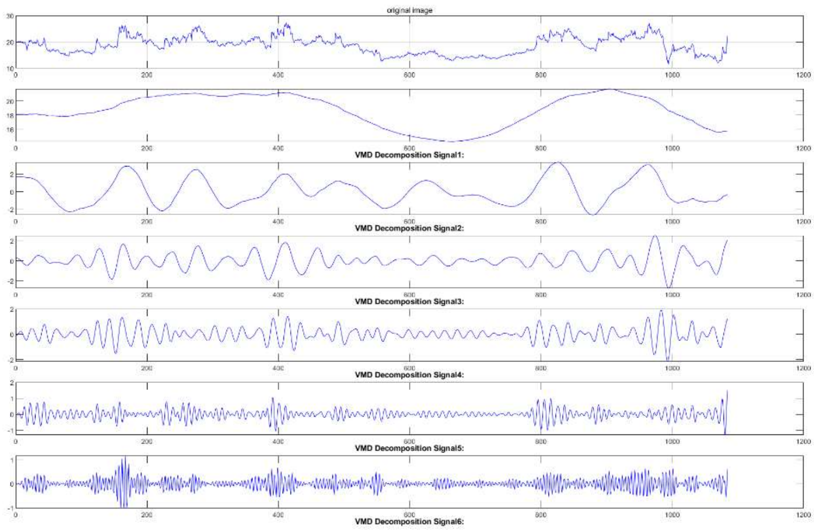

Figure 4. Variational Modal Decomposition (VMD) decomposition results

The top graph shows the original signal, which may contain a superposition of multiple frequency components exhibiting complex fluctuation characteristics. VMD Decomposition Signals: The six subplots below show each of the six IMFs obtained by VMD decomposition, with each IMF representing a specific frequency band or mode in the original signal. Signal 1: Usually contains the highest frequency component and may be a high-frequency noise or rapidly changing part of the signal. In the figure, this signal may appear as a rapid fluctuation. Signal 2: May contain the next highest frequency component, representing some medium frequency periodic variation in the signal. This signal may fluctuate more slowly than Signal 1, but still relatively quickly. Signal 3: This signal may represent the low-medium frequency component of the signal that fluctuates more slowly than Signal 2 and may reflect some moderate periodic variation in the signal. Signal 4: May represent a low-frequency component of the signal, such as some slower periodic changes or trends. Signal 5: This signal may contain a lower frequency component and may represent a long-term trend or very slow change in the signal. Signal 6: Usually contains the lowest frequency component and may represent a long-term trend or very slow change in the signal, perhaps even close to a constant. Amplitude and Frequency: The y-axis of each subplot represents the amplitude of the signal, where the variation in amplitude of the different IMFs can be observed. The x-axis represents the time, where the fluctuating characteristics of the different IMFs over time can be observed.

Figure 5-10 shows the results of 6 iterations. The R value is getting closer to 1 with the increase of iterations; the RMSE is 0.22012 from the iteration result of drop by drop to the result of the 6th iteration of 0.079422; there is a large reduction. The MSE with the increase of iterations also has a large reduction; the model convergence results are better.

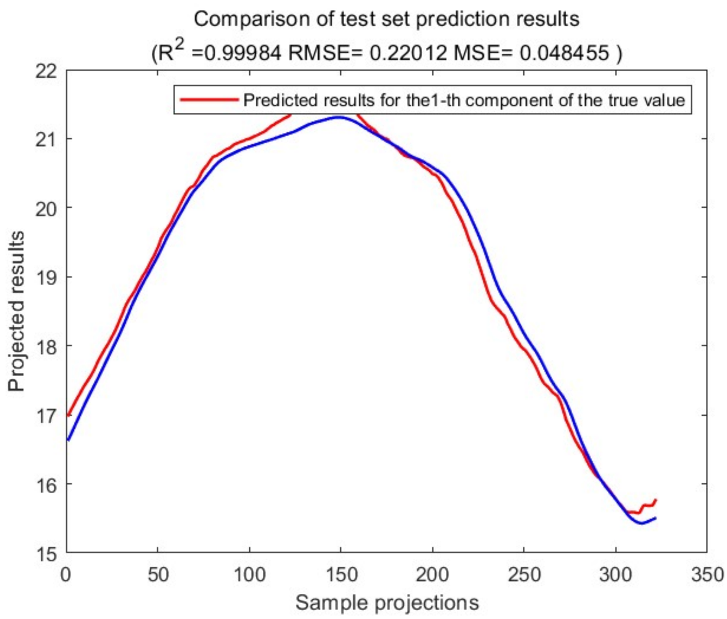

Figure 5. VMD-BILSTM model 1th decomposition prediction effect plot

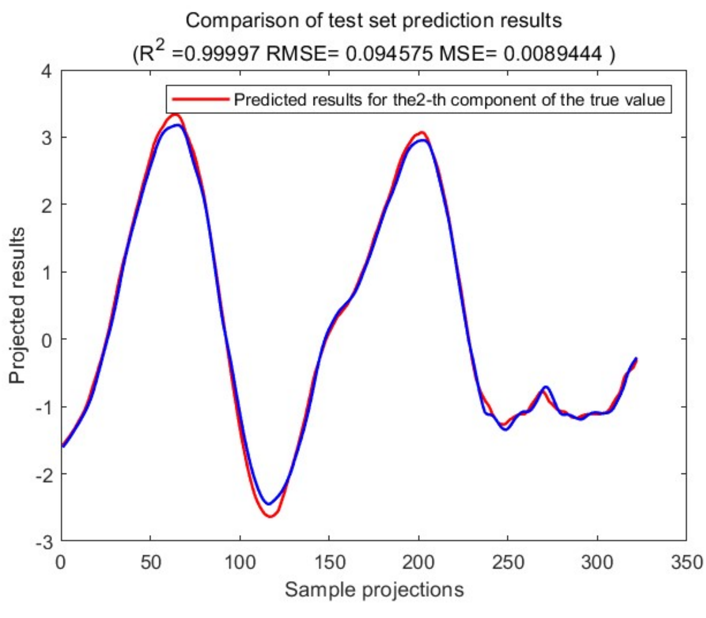

Figure 6. VMD-BILSTM model 2th decomposition prediction effect plot

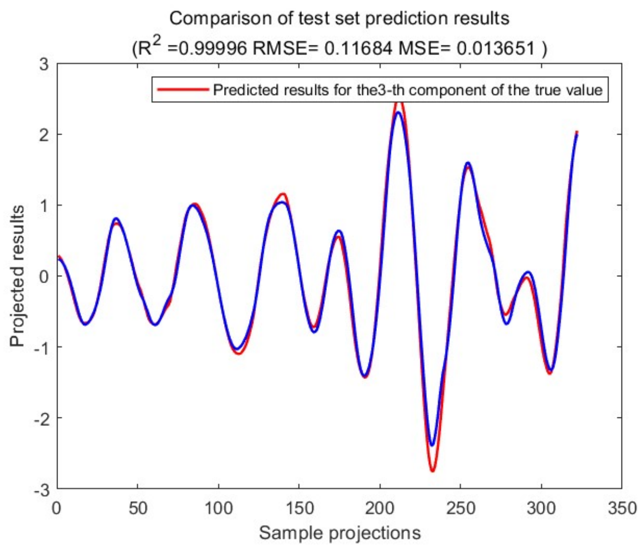

Figure 7. VMD-BILSTM model 3th decomposition prediction effect plot

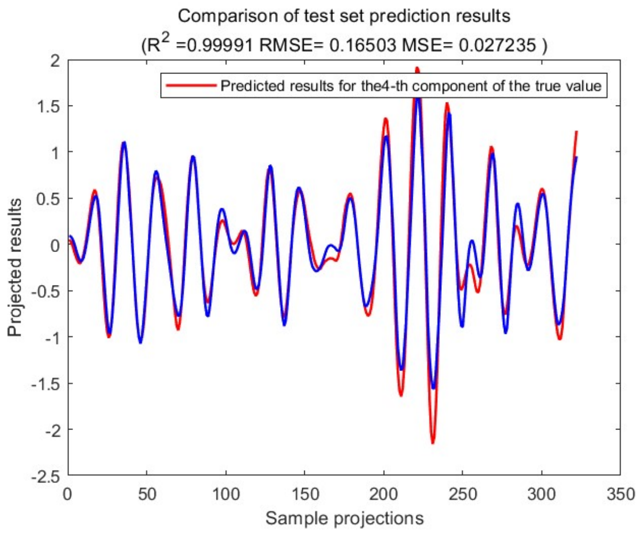

Figure 8. VMD-BILSTM model 4th decomposition prediction effect plot

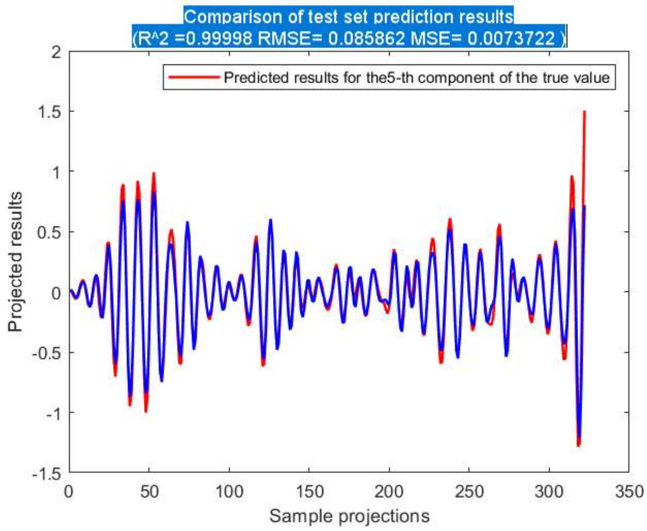

Figure 9. VMD-BILSTM model 5th decomposition prediction effect plot

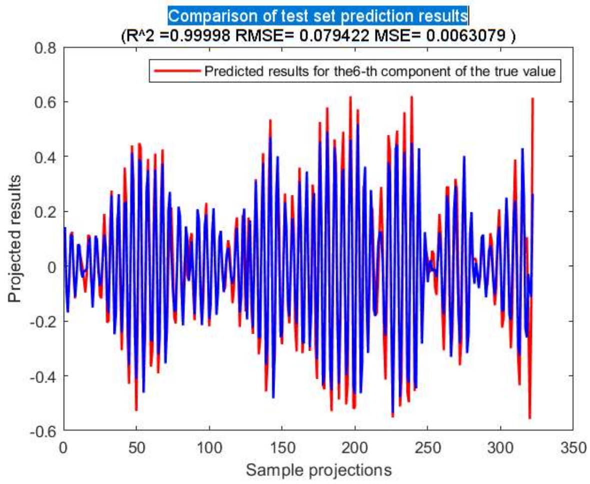

Figure 10. VMD-BILSTM model 6th decomposition prediction effect plot

After the stock price dataset has been processed by variational modal decomposition (VMD), this paper then employs the constructed bidirectional long-short time memory (BiLSTM) model for sequence prediction. By superimposing the prediction results of each subsequence, the final stock price closing price prediction is obtained. From the evaluation metrics, the VMD-BiLSTM model shows better prediction performance and higher accuracy.

Table 2. Comparison of model evaluation indicators

Precision indicators	RMSE	MAE	R^2
LSTM	0.8778	0.6144	0.9355
BiLSTM	0.8857	0.6147	0.9343
SSA-LSTM	0.9031	0.6458	0.9316
VMD-LSTM	0.8205	0.6280	0.9437
VMD-BiLSTM	0.58532	0.42133	0.97133

To further analyze the difference between the predicted values and the raw stock price data, the specific fit is shown in figure 11.

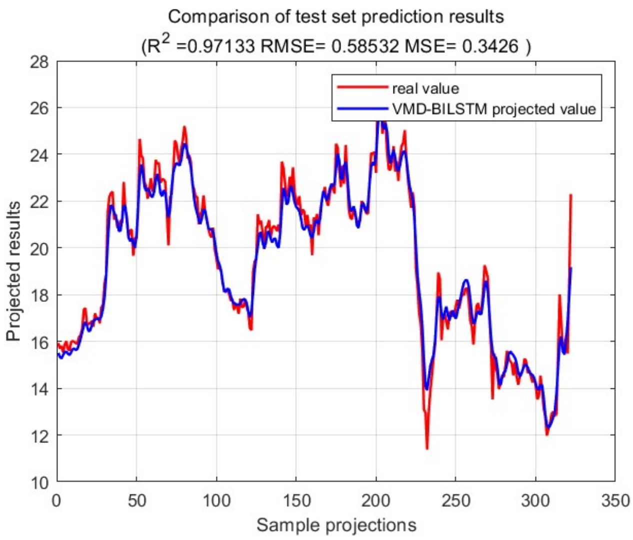

Figure 11. VMD-BILSTM Model Stock Price Forecast

From the evaluation results, the model performs well in predicting stock price movements and is able to capture the main movements of the stock price by and large. However, the model's prediction accuracy decreases for regions where the stock price fluctuates dramatically and reaches extreme values. This may be due to the fact that stock prices change too rapidly during certain periods, making it difficult for the model to accurately identify and reflect these instantaneous and complex fluctuations.

4. Conclusion

The article first analyzes the research background of stock price forecasting in depth and clarifies its research value and the necessary conditions for making forecasts. Subsequently, a large amount of literature is synthesized, on the basis of which a model integrating multiple approaches is proposed. Afterwards, this integrated model is empirically analyzed using a stock price dataset. In order to comprehensively assess the performance of the proposed algorithm, this paper designs multiple sets of experiments comparing a single model (e.g., LSTM and BiLSTM) with a composite model (e.g., SSA-BiLSTM, VMD-LSTM, and VMD-BiLSTM) and validates the model's effectiveness through multiple evaluation metrics. Ultimately, the main findings are derived by visualizing the prediction results of each model:

1) Deep learning is more robust in stock price prediction than traditional machine learning algorithms. While traditional machine learning algorithms vary greatly in their fitness for different datasets, deep learning is able to perform well across different datasets.

2) The VMD-BiLSTM model is used to predict the stock dataset better than the VMD-LSTM model. After further decomposition process, the root mean square error (RMSE) and mean absolute error (MAE) of the model are reduced, while the coefficient of determination (R^2) is improved. This indicates that through the secondary decomposition, the model can extract the features of the time series data more efficiently, which enhances the accuracy of the prediction.

Acknowledgments

China College Students' Entrepreneurship Training Program Project (Letter of Department of Higher Education [2024] No. 13)), Hunan Province College Students' Innovation and Entrepreneurship Training Program Project (Xiangtong [2024] No. 118), Hunan Institute of Technology College Students' Innovation and Entrepreneurship Training Program Project (S202411528027); Hunan Province College Students' Innovation and Entrepreneurship Training Program Project (Xiangtong [2023] No. 132), Hunan Engineering College Students' Innovation and Entrepreneurship Training Program Project (S202311528013);Research Project of Hunan University of Technology: Financial Risk Measurement Based on Time-Varying Copula Models.

References

[1] Box G E P .Time series analysis, forecasting and control rev. ed[M].Holden-Day,1976.

[2] Li Y , Zhu Z , Kong D ,et al.EA-LSTM: Evolutionary attention-based LSTM for time series prediction[J]. Knowledge-Based Systems, 2019, 181(Oct.1):104785.1-104785.8.DOI:10.1016/ j.knosys.2019. 05.028.

[3] Wang W , Yang X , Ooi B C ,et al.Effective deep learning-based multi-modal retrieval[J].Vldb Journal - the International Journal on Very Large Data Bases, 2016, 25(1):79-101.

[4] Lu W , Li J , Wang J ,et al.A CNN-BiLSTM-AM method for stock price prediction[J].Neural Computing and Applications, 2020:1-13.DOI:10.1007/s00521-020-05532-z.

[5] Manuel Méndez, Merayo M ,M. Núez.Long-term traffic flow forecasting using a hybrid CNN-BiLSTM model[J].Eng. Appl. Artif. Intell. 2023, 121:106041.DOI:10.1016/j.engappai.2023.106041.

[6] Kulshrestha A , Krishnaswamy V , Sharma M ,et al.Bayesian BILSTM approach for tourism demand forecasting[J].Annals of Tourism Research, 2020, 83.DOI:10.1016/j.annals.2020.102925.

[7] Wu M T , Li Z , Herencsar N ,et al.A graph-based CNN-LSTM stock price prediction algorithm with leading indicators[J].Multimedia systems, 2023.

[8] Lu W , Li J , Wang J ,et al.A CNN-BiLSTM-AM method for stock price prediction[J].Neural Computing and Applications, 2020:1-13.DOI:10.1007/s00521-020-05532-z. [9]Banerjee Soham, Mukherjee Diganta. Short Term Stock Price Prediction in Indian Market: A Neural Network Perspective[J]. Studies in Microeconomics, 2022, 10(1):23-49.

[9] Li Z , Yang S , Chen Y ,et al.A Road Roughness Estimation Method based on PSO-LSTM Neural Network[J].SAE Technical Paper Series, 2023.

[10] Mustaqeem, Sajjad M , Kwon S .Clustering-Based Speech Emotion Recognition by Incorporating Learned Features and Deep BiLSTM[J].IEEE Access, 2020, 8:79861-79875.DOI:10.1109/ ACCESS. 2020.2990405.