Vol. 19 (2026): Proceedings of 2026 International Conference on Financial Technology, Mathematical Statistics, and Business Analytics (ICFMB 2026)
E-ISSN: 3105-0522 P-ISSN: 3105-0514
DOI: https://doi.org/10.70267/icfmb.202619331336

A Review of Stock Forecasting Methods Based on Improved LSTM Series Models

Mengqi Li*

Hainan Normal University, Cao County, Heze City, Shandong Province, China

**Corresponding author: Mengqi Li.*

Abstract

The trend of stock prices not only immediately affects policymakers' decisions and income, but also reflects macroeconomic conditions and market status. In the Chinese capital market, scientific and reliable methods for stock price prediction are highly relevant to practice. Financial time series exhibit nonlinearity, high noise, and strong randomness. And it is affected by sentiment and emergency events—the traditional and fundamental models. LSTMs have not met the demand for predictive performance. This paper systematically reviews the research process for stock prediction based on improved LSTMs, elaborates on the different improved LSTMs and their drawbacks, and summarizes the principles, advantages, and disadvantages of the four models: PCA-LSTM, CNN-LSTM-Attention, BERT-Bi-LSTM, and Event LSTM-Attention. This paper compares the distinctions among this model's Data Adaptation, Feature Mining, and Market Response, and outlines future development directions for Multimodal Fusion, Interpretability Enhancement, and Extreme Market Condition Modeling. This paper provides theoretical foundations and methodological support for model selection and innovation research in stock price prediction.

Keywords

LSTM, stock price prediction, neural network, time series

1. Introduction

The stock market is a core component of the financial system. Stock price fluctuations exhibit nonlinearity, high noise, nonstationarity, and sudden shocks. They are affected by the interplay of multiple factors, including the macroeconomy, policy events, and market sentiment. So, the prediction is highly complex. Scientific and effective methods not only serve as decision-making references for investors and reduce investment risks, but also provide statistical support for market supervision and policy formulation. Thus, this is of great importance to sustain a stable financial market. As our country's capital market continues to develop, the prominence of a high proportion of individual investors and drastic fluctuations in market sentiment has become increasingly evident. The traditional prediction method is difficult to adapt to a complex and volatile market environment. More efficient and reliable intelligent forecasting models are urgently required to provide a more scientific basis for market participants' decision-making.

According to current findings, machine learning methods used for stock price prediction in early literature were mostly supplanted by traditional methods. For example, Xiang Chaoju [1] applied the Grey Wolf Algorithm to optimize the BP neural network. Zou Jie and Li Lu [2] constructed an SA-BiGRU model based on random forests to predict the next day's closing price of stocks. Lin Mingsong, Yang Xiaomei, and Yang Zhixia [3] proposed a structured maximum-margin twin support vector machine, which mitigated the impact of high noise in stock data on classification decisions to some extent. In recent years, with the rapid development of deep learning, research methods for stock prediction have shifted from traditional machine learning to deep learning approaches such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs). Researchers found that RNNs have clear limitations when handling long time series, including weak long-term memory and eventual forgetting of early information. Meanwhile, as the time series length increases and the model depth deepens, gradient vanishing and gradient explosion become unavoidable, limiting its use in financial time series forecasting.

To address this problem, the LSTM model emerged as an improved model of RNN. By introducing a gating mechanism, it effectively alleviates the gradient vanishing problem and can better capture the long-term trend regularity of financial time series. Consequently, it has rapidly become a mainstream method for predicting stock prices and return rates. At present, numerous scholars worldwide have applied LSTM and its improved models to the field of stock price prediction. Zhao Hongyan [4] proposed an LSTM-CNN-CBAM hybrid model that integrates time-series features, deep features, and attention mechanisms, and verified its predictive effectiveness using the SSE Composite Index. Jiang Shuyu [5] directly adopted the LSTM model to predict the stock price of Pudong Development Bank, and the experimental results demonstrate that the model achieves low error and high accuracy. Wu Yin [6] et al. constructed an event-aware EvenLSTMATT model, which integrates non-periodic public announcement events with time-series stock price data, significantly improving the prediction accuracy of stock price trends. Hu Shuaiwen [7] developed PCA-LSTM and LASSO-LSTM models. By performing dimensionality reduction and screening on stock technical indicators, he found that PCA-LSTM achieves the optimal accuracy and stability. Xu Xuechen and Tian Kan [8] design a BERT-Bi-LSTM model to predict stock prices through sentiment analysis of financial texts.

This paper focuses on improved LSTM models for stock price prediction, with the main research content as follows: (1) Analyze the characteristics of financial time series and the adaptability of PCA, CNN, BERT, attention mechanisms and event modules; (2) Systematically review the principles, advantages, disadvantages and desirable scenarios of four types of models: PCA-LSTM, CNN-LSTM-Attention, BERT-Bi-LSTM and EventLSTM-Attention; (3) Horizontally compare the performance differences among the models; (4) Summarize the existing limitations and prospect future research trends.

2. Discussion on the Adaptability of Basic Theories Related to Stock Prediction in Financial Time Series

Financial time-series data exhibit complex characteristics, such as nonlinearity, high redundancy, strong noise, sentiment sensitivity, and policy and market emergencies. A single algorithm can’t adapt to every market rule. Simultaneously, each basic theory exhibits differentiated adaptability when addressing different flaws in financial data.

Principal Component Analysis (PCA) is well-suited to the traits of financial indicators, including complex dimensionality, information redundancy, and so on. It screens core features via linear dimensionality reduction, eliminates ineffective indicators due to multicollinearity, compresses high-dimensional trading data, reduces the computational burden on models, and addresses the disorder and redundancy of original high-dimensional financial data.

Convolutional Neural Networks (CNNs) excel at mining local spatial features of financial data. By leveraging convolutional kernel sliding, they extract short-term fluctuation patterns of candlestick patterns and technical indicators, filter random noise, accurately capture short-term price oscillation modes, and enhance the capability of micro-pattern recognition.

The BERT sentiment embedding layer adapts to the sentiment-sensitive nature of the financial market. It performs semantic encoding on news and announcement texts, quantifies market sentiment and sentiment tendency, converts unstructured texts into computable features, and aligns with the characteristic that stock prices fluctuate under the influence of investor sentiment.

The Even Perception Unit is specifically designed to handle abrupt financial fluctuations. By adding an independent event control gate, it embeds non-periodic announcements and policy shocks into the model, compensates for the defect of traditional algorithms in failing to identify sudden external events, and adapts to abrupt changes in market trends.

The attention mechanism optimizes the weight distribution of financial time-series data. It assigns high weights to critical time nodes and authoritative news sources, weakens the influence of useless noise, amplifies the influence of key information, and aligns with the long-term memory and uncertainty characteristics of financial data.

In summary, dimensionality reduction algorithms, convolutional networks, sentiment-encoding event modules, and attention mechanisms respectively target five major pain points of financial data: redundancy, local fluctuations, public opinion interference, abrupt shocks, and information screening. They provide a comprehensive theoretical foundation for the improved prediction models discussed in subsequent chapters.

3. Research Review of Mainstream Improved LSTM Prediction Models

3.1 Research Review of the Basic LSTM Stock Prediction Model

The LSTM neural network is a variant of the RNN, specifically designed for processing time-series data such as stock prices. Its core principle is to predict future stock prices using historical window sequences. The LSTM consists of an input gate, an output gate, and a memory cell. The mathematical expression of the forget gate is$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$ The threshold value obtained from this calculation ranges between 0 and 1. A higher value indicates greater information retention. The input gate determines which data to store in the memory cell, the cell state integrates information, and the output gate controls the output of cell-state data for stock price prediction. The LSTM inherently has a noise-reduction capability and can capture both long-term trends and short-term fluctuations, while resolving the gradient vanishing problem of RNNs. However, it lacks a local feature extraction module and the capability to purify features. Stock data contains numerous features. So directly inputting original high-dimensional redundant features will introduce correlations and information redundancy that interfere with model training, thereby generating substantial invalid noise.

3.2 Literature Review of PCA-LSTM Hybrid Prediction Model

To address the problem of excessive technical indicators in stock data, the Principal Component Analysis (PCA) algorithm is applied before the original Long Short-Term Memory (LSTM) network. First, linear dimensionality reduction and redundancy elimination are applied to the original high-dimensional features of the stock data. Highly correlated and information-redundant features are eliminated, while the principal components that contribute the most information to the data are retained. The resulting low-dimensional purified features are then fed into the LSTM network to complete the time-series prediction. This method effectively addresses the information redundancy introduced by multi-dimensional features in stock prediction. It extracts the principal components that represent the impact factors, using the minimum number of principal components as input variables for the LSTM neural network model, thereby improving calculation speed and prediction accuracy. However, PCA is a linear dimensionality reduction method. It performs poorly at dimensionality reduction for non-linear data with complex correlations, leading to the loss of key non-linear information. In addition, it lacks a local feature extraction module and cannot identify local fluctuations and abrupt short-term market changes in stock data.

3.3 Literature Review of CNN-LSTM-Attention Model

The CNN-LSTM-Attention model integrates a local feature extraction module and an Attention mechanism based on the original LSTM network. A Convolutional Neural Network (CNN) can extract local, short-term features, such as K-line patterns, short-term trends, and fluctuations. The CNN feature vectors are converted into time series for input to the LSTM, enabling the CNN-LSTM model to capture both local and global information more effectively. On this basis, the Attention mechanism is introduced to dynamically weight each LSTM time step, automatically focusing on key dates such as financial statement release days, policy announcement days, and abrupt market decline days, thereby enhancing the importance of key features and adapting to complex K-line trends. The CNN-LSTM-Attention model can capture long-term dependencies and trend patterns in time series, accounting for both short-term abrupt changes and long-term trends. Nevertheless, it generalizes poorly under abnormal market conditions, such as policy shocks and sharp rises and crashes, which can easily lead to prediction bias. It relies solely on data-driven modeling, without integrating the objective laws of the financial market, and lacks an understanding of sentiment. In the stock market, stock prices are affected not only by historical price data but also by financial news, which is a critical driver of price fluctuations. The CNN-LSTM-Attention model is incapable of capturing the sentiment impact derived from text information such as financial news and market public opinion.

3.4 Research Review of the BERT-Bi-LSTM Sentiment and Time-Series Fusion Model

The BERT-Bi-LSTM hybrid model effectively integrates the advantages of semantic mining and time-series modeling, and has been widely applied in research on stock price fluctuation prediction. Xu Xuechen and Tian Kan [8] proposed BERT-Bi-LSTM, an index prediction model based on financial text sentiment analysis. It uses BERT, a state-of-the-art natural language processing model, in combination with a bidirectional LSTM model for sentiment analysis. As a bidirectional Transformer-based encoding model, BERT possesses powerful contextual semantic comprehension. It can bidirectionally analyze the contextual information of financial texts and accurately extract market sentiment features, including optimism, pessimism, and neutrality, addressing the limitation that traditional word vectors cannot deeply understand financial semantics. Bi-LSTM consists of forward and reverse LSTMs, which can bidirectionally capture long-term dependencies in historical stock price sequences, alleviate the gradient vanishing problem, and comprehensively mine the time-series patterns of stock price fluctuations. This model first employs BERT to extract affective semantic features from financial news, then uses Bi-LSTM to learn time-series patterns in stock prices. It ultimately fuses text sentiment features and price time-series features as network inputs to predict stock price trends. Compared with the traditional LSTM, BERT-Bi-LSTM simultaneously incorporates quantitative transaction data and market public sentiment, offering a broader information dimension, superior time-series fitting performance, and higher accuracy in rise-fall classification. It demonstrates excellent performance in financial time-series prediction and has become a mainstream improved model for sentiment-integrated stock price prediction.

Nevertheless, the model still has limitations. The bidirectional structure of Bi-LSTM averages market data, thereby smoothing out sudden stock price changes. This results in overly smooth prediction outcomes, making it difficult to identify short-term sharp declines. Additionally, it is biased toward learning stable daily data and lacks sufficient sensitivity to sharp crashes and extreme fluctuations. Furthermore, BERT can only process publicly available text information and fails to capture early weak abnormal signals. When black-swan emergencies occur in the market, the original rise-fall patterns are disrupted, and the data distribution shifts.

3.5 Research Review of the State-of-the-Art EventLSTM-Attention Prediction Model

In view of the above defects existing in the BERT-Bi-LSTM model, Wu Tao et al. constructed the event-aware EventLSTM-Attention model, which is specially optimized for the extreme abrupt change scenario in the financial market. Aiming at the problem that the bidirectional structure of Bi-LSTM tends to smooth out stock price mutations and lead to over-smoothed prediction results, the model strengthens the feature extraction ability for abrupt market conditions, effectively reduces the prediction errors caused by over-smoothing, and greatly improves the recognition accuracy of extreme fluctuations such as stock crashes and short-term rapid declines. Meanwhile, the attention mechanism is introduced to adaptively amplify weak key signals, such as trading volume and abnormal fluctuations in public sentiment, which are often overlooked by conventional models, thereby addressing their limited sensitivity to extreme events. To address the issues that BERT cannot capture hidden early abnormal signals and struggles to adapt to sudden changes in market patterns under black-swan events, the EventLSTM-Attention model deeply integrates historical transaction time-series data with public event information. The specific process is as follows: First, the embedding layer maps historical transaction data and public event information into a low-dimensional semantic space. Then, the bottom EventLSTM unit performs fusion modeling of the two data types via an event information control gate to capture the non-periodic impact of events on stock prices. The upper traditional LSTM unit further learns complex temporal features and long-term dependencies. The attention layer weights the hidden states at each time step to focus on key event information that significantly impacts stock price prediction. Finally, the feed-forward neural network outputs a prediction of the stock return rate. This model effectively breaks through the limitation of traditional models that overly rely on stable historical data and fail to adapt to sudden changes in data distribution under black-swan events, achieving superior financial prediction performance in the stock market with high noise and strong suddenness.

3.6 Horizontal Comparison of Models

In summary, this paper systematically sorts through and compares four mainstream improved LSTM-based stock prediction models. Obvious differences exist among these models in terms of data adaptation scenarios, feature mining capability, and market adaptability. The PCA-LSTM model excels at processing high-dimensional, redundant technical indicators; it simplifies financial input data through feature dimensionality reduction and achieves high computational efficiency, yet it exhibits limited capability for mining nonlinear information, such as K-line patterns and textual events. The CNN-LSTM-Attention model can effectively extract local temporal features of K-lines and technical indicators and optimizes long-term information dependence via attention weights, delivering excellent prediction performance under conventional trend-based market conditions. Still, it struggles to respond to sudden policy shocks and bullish and bearish events. The BERT-Bi-LSTM model focuses on mining the emotional semantics of financial news, adapts to stock price fluctuations driven by market sentiment, and adeptly captures changes in public opinion. However, it suffers from insufficient stability in fitting long-term price trends. The EventLSTM-Attention model integrates sudden announcement information via an independent event gating structure, making it well-suited to the sudden characteristics of the financial market and enabling it to achieve significantly higher prediction accuracy than other models under event-impacted market conditions.

4. Conclusion

Currently, most models only adopt single price–time series data for prediction, leading to insufficient mining of market information. In the future, LSTM models will further integrate multi-dimensional information, including stock market quotation data, financial news texts, social media public opinions, macroeconomic indicators, and policy emergencies. Using BERT and large language models, in-depth encoding of text semantics and market sentiment will be achieved. Public opinion sentiment features, event signals, and time-series data will be uniformly embedded into model inputs, thereby opening up the transmission chain of “data-sentiment-event-stock price”. This can address the shortcomings of pure time-series models that ignore market sentiment and emergency information, and ultimately enable multimodal integrated modeling.

References

[1] Xiang, C. J. Stock Price Prediction of the BP Neural Network Optimized by the Grey Wolf Optimizer[J]. Science & Technology Information, 2024, 10: 253-256.

[2] Zou, J., Li, L. Research on Stock Price Prediction Based on Random Forest and SA-BiGRU Model[J]. China Price, 2023, 11: 52-56.

[3] Lin, M. S., Yang, X. M., Yang, Z. X. Structured Maximum Margin Twin Support Vector Machine and Its Application in Stock Trend Prediction[J]. Computer Engineering and Applications, 2024, 60(11): 346-353.

[4] Zhao, H. R., Xue, L. Research on Stock Forecasting Based on LSTM-CNN-CBAM Model[J]. Computer Engineering and Applications, 2021, 57(3): 203-207.

[5] Hu, Y. W. Stock Forecast Based on Optimized LSTM Model[J]. Computer Science, 2021, 48(6A): 251-257.

[6] Wu, X., Zhao, G. Y., Liu, H. Z. Stock Price Trend Prediction Based on Event-Aware LSTM Model[J]. Computer Engineering and Applications, 2025, 61(22): 295-303.

[7] Jiang, S. Y. Stock Price Prediction based on LSTM Model[J]. Jiangsu Commercial Forum, 2025, 1: 83-86.

[8] Xu, X. C., Tian, K. A Novel Financial Text Sentiment Analysis-Based Approach for Stock Index Prediction[J]. The Journal of Quantitative & Technical Economics, 2021, 38(12): 121-141.

[9] Xu, Y. E., Luo, Y. Literature Review of Stock Price Forecasting Based on Combination Models[J]. Science Technology and Industry, 2022, 22(11): 214-220.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Open Access