Academic Journal of Science and Technology
ISSN: 2771-3032 | Vol. 13, No. 1, 2024

Stock Market Prediction Model Based on Deep Learning and Enhancement of Interpretabilit

Jiazhen Liu

Hongwen school Qingdao campus, Qingdao, China

Abstract: This paper's objective is to delve into the utilization of deep learning technology within the realm of stock market forecasting, specifically emphasizing the enhancement of model interpretability. To accomplish this, we employ a deep learning model rooted in the long-term and short-term memory network (LSTM). We proceed to construct three distinct models: the foundational LSTM model, an LSTM model augmented with an attention mechanism, and an LSTM model incorporating an integrated learning strategy. By conducting comparative experiments, we assess the effectiveness of these models in both regression prediction, particularly forecasting the closing price of the following day, and fluctuation prediction. The outcomes of these experiments reveal that integrating an attention mechanism and an integrated learning strategy notably boosts the prediction accuracy of the LSTM model. The attention mechanism, by dynamically assigning weights to various time steps and features, amplifies the model's focus on crucial information, thereby enhancing prediction accuracy. At the same time, the attention mechanism also provides an intuitive explanatory perspective. The integrated learning strategy improves the overall stability and generalization ability by combining the predictions of multiple models. These findings help investors and financial institutions make more informed decisions.

Keywords: Stock market forecast; Deep learning; Explanatory; Investor.

1. Introduction

Under the double wave of globalization and informatization, the volatility and complexity of the stock market are increasing. Investors, financial institutions and even policy makers have high expectations for accurately predicting the dynamics of the stock market in order to gain competitive advantages or avoid risks in the rapidly changing market [1]. However, the stock market is influenced by many factors such as macroeconomic environment, policy changes, market sentiment and international situation, and its trend prediction has always been regarded as a major challenge in the financial field [2]. The deep learning model, especially the long-term and short-term memory network (LSTM), can effectively capture the long-term dependence in time series data through its unique memory cell design [3]. LSTM can not only deal with the time series characteristics of data, but also alleviate the problem of gradient disappearance or explosion that is easy to occur when traditional neural networks train long series to some extent [4].

Although the deep learning model has made remarkable progress in forecasting accuracy, its "black box" feature, that is, the internal operation mechanism of the model, is difficult to understand intuitively, which has become a major obstacle to its wide application in the financial field [5]. In financial decision-making, the interpretability of the model is very important. It is not only related to the transparency and credibility of decision-making, but also an important basis for regulators to evaluate model risks and ensure market fairness [6]. Therefore, how to improve the interpretability of the deep learning model while maintaining its predictive performance has become a hot topic in the current financial artificial intelligence research.

Based on this background, this study aims to explore the stock market forecasting model based on deep learning, especially the LSTM model, and how to enhance its interpretability through innovative methods while improving the forecasting accuracy. It is expected that this study can provide investors with more accurate market forecasting tools and promote the trust and acceptance of deep learning models by financial institutions and regulators.

2. Theoretical Basis

2.1. Basic principles of deep learning

Deep learning, a subset of machine learning, mimics the human brain's learning mechanism through the establishment of deep neural networks. It achieves automatic feature extraction and pattern recognition for intricate datasets. In contrast to conventional machine learning methods, deep learning excels in managing high-dimensional data and addressing nonlinear relationships. The essence of deep learning involves transforming raw data into a high-dimensional feature space via multiple layers of nonlinear mappings, enabling the capture of fundamental data characteristics.

2.2. Structural characteristics of LSTM model

To address the gradient vanishing or exploding issues encountered by traditional Recurrent Neural Networks (RNNs) when handling long sequences, the LSTM network was developed as a specialized recurrent architecture. LSTM achieves long-term information retention and selective amnesia by incorporating a memory unit alongside three regulatory gates: the input gate, the forget gate, and the output gate [7]. The memory unit is tasked with preserving extensive sequence information and ensuring informational continuity via linear self-recurrence. The input gate governs the influx of new information into the memory unit, determining updates based on the present input and prior hidden state [8]. Meanwhile, the forget gate selects which pieces of information to eliminate from the memory unit, preventing the accrual of irrelevant data. Lastly, the output gate modulates the memory unit's impact on the immediate output, generating the hidden state in conjunction with the current input and memory contents.

2.3. The importance of interpretability and existing methods

In the financial field, the interpretability of the model means that the model can provide a reasonable explanation of its decision-making process and prediction results, so that users can understand why the model makes such a prediction [9]. Interpretability is very important for enhancing users' trust, assisting decision-making and meeting regulatory requirements. At present, the methods to improve the interpretability of machine learning models can be mainly divided into the following categories: model simplification: improving the interpretability by reducing the complexity of the model or using simpler models (such as decision trees and linear regression) [10]. Feature importance analysis: evaluate the influence of each feature on the prediction results of the model, such as using SHAP value, weight analysis and other methods. Visualization technology: the interpretability of the model is enhanced by visualizing the decision-making process of the model, the relationship between features or the trend of model output changing with input. Agent model: build a simple and explainable agent model to approximate the behavior of complex model, so as to realize the explanation of complex model.

3. Methodology

The LSTM network represents a specialized form of RNN, tailored to mitigate the gradient vanishing issue inherent in traditional RNNs when addressing long-term dependencies. Our study outlines an LSTM model comprising an input layer, LSTM layer, fully connected layer, and output layer. The input layer processes pre-transformed stock market data encompassing price, volume, and technical indicators. At the model's heart, the LSTM layer hosts multiple LSTM units, each equipped with a forget gate, input gate, and output gate to regulate information flow and retention. Succeeding the LSTM layer, the fully connected layer translates its output into predictive targets like future prices or trends. According to different forecasting tasks, the output layer may use sigmoid function for binary classification (such as price forecasting) or linear activation function for regression forecasting (such as price forecasting). The LSTM model structure of this paper is shown in Figure 1.

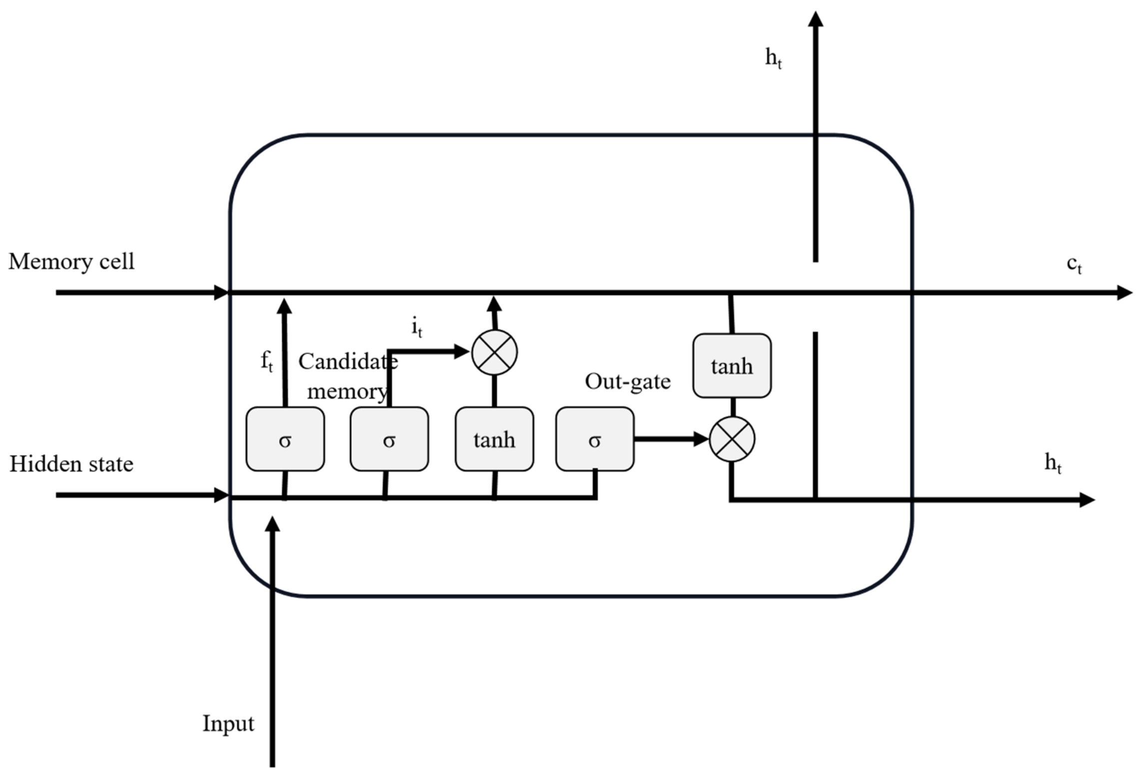

Figure 1. LSTM model

In order to improve the stability and generalization ability of the model, the ensemble learning method is adopted, and the final prediction is obtained by voting or weighted average combining the prediction results of several LSTM models. Transfer learning allows knowledge learned in one task to be transferred to another related task. In the stock market forecast, the general features are learned from a large number of historical data through pre-training models, and then fine-tuned for a specific market or time period.

Assuming the innovation sequence $\{et\}$ of k assets' liquidity follows a multivariate normal distribution with mean 0 and covariance matrix H_t, while k asset represents an independently and identically distributed white noise process (where $e_t|\Omega_{t-1} \sim N(0, N_t)$), let Ω_{t-1} denote the information set of r_t at time t. The dynamic correlation structure is defined as follows:

$$r_t = u_t + e_t \tag{1}$$

$$e_t|\Omega_{t-1} \sim N(0, H_t) \tag{2}$$

$$H_t = D_t R_t D_t \tag{3}$$

$$Q_t = \left(1 - \sum_{m=1}^{M} a_m - \sum_{n=1}^{N} \beta_n\right)\overline{Q} + \sum_{m=1}^{M} a_m \left(\varepsilon_{t-m}\varepsilon_{t-m}^{'}\right) + \sum_{n=1}^{N} \beta_n Q_{t-n} \tag{4}$$

Here, $\overline{Q} = T^{-1}\sum_{t=1}^{T}\varepsilon_t\varepsilon_t^{'}$ signifies the unconditional variance matrix for standardized residuals, $R_t = \left(Q_t^*\right)^{-1} Q_t \left(Q_t^*\right)^{-1}$, and Q_t^* represents the diagonal value in Q_t.

Attention mechanism can provide dynamic weights for the model and reveal which input features or time steps contribute the most to the prediction results. By adding attention layer after LSTM layer, we can visualize these weights and understand the basis of model decision. By visualizing the internal state of LSTM layer, the distribution of attention weight and the change of model output with time, we can intuitively show how the model processes the input data and make predictions.

4. Experimental Results and Analysis

4.1. Experimental design

Our experimental dataset comprises daily trading records of a stock market index spanning from 2010 to 2022, encompassing opening, closing, highest, and lowest prices, as well as trading volumes. This data is partitioned into a training set (2010-2019), a validation set (2020), and a test set (2021-2022). The objective of our forecasting is to predict the closing price of the subsequent day.

4.2. Model description

In this study, three stock market forecasting models based on LSTM are adopted: LSTM-Base, whose architecture includes two stacked LSTM layers, each with 128 units; Add the LSTM model of attention mechanism (LSTM-Attn), and introduce the attention layer to the basic model to dynamically allocate weights; And the integrated learning LSTM model (LSTM-Ensemble), using Bagging strategy to integrate the prediction results of five basic LSTM models. All models use mean square error as loss function, and Adam optimizer optimizes them.

4.3. Evaluating indicator

To conduct a thorough assessment of the model's performance, this study employed four evaluation metrics: Mean Square Error (MSE) and Mean Absolute Error (MAE) for quantifying the discrepancy between predicted and actual values; accuracy rate to gauge the correctness of fluctuation predictions; and Sharp ratio to holistically evaluate the model's risk-adjusted returns, reflecting its practical application performance.

4.4. Experimental results

Table 1 shows the regression prediction performance of the three models on the test set (the next day's closing price prediction).

Table 1. Comparison of Regression Prediction Performance

Model	MSE	MAE
LSTM-Base	0.0052	0.054
LSTM-Attn	0.0048	0.052
LSTM-Ensemble	0.0045	0.050

Compared with the basic LSTM-Base, LSTM-Attn with attention mechanism has lower MSE and MAE, which shows that attention mechanism is helpful to improve the prediction accuracy. LSTM-Ensemble, which adopts integrated learning strategy, further reduces the error and achieves the best performance.

Table 2 shows the performance of the three models in the fluctuation prediction (binary classification problem).

Table 2. Comparison of Up/Down Prediction Performance

Model	Accuracy	Sharpe Ratio
LSTM-Base	56.8%	0.45
LSTM-Attn	58.2%	0.48
LSTM-Ensemble	59.5%	0.51

In the ups and downs forecast, LSTM-Ensemble also showed the best performance, the accuracy rate increased to 59.5%, and the Sharp ratio also increased significantly, indicating that ensemble learning not only improved the prediction accuracy, but also optimized the balance between risks and benefits.

In addition to predicting performance, this study also pays attention to the interpretability of the model. Through the attention mechanism, the influence of different time steps and features in the LSTM-Attn model on the prediction results is analyzed. The results show that historical closing price, trading volume and specific technical indicators play a key role in the forecast.

4.5. Discussion

The results show that the stock market forecasting model based on LSTM can improve the forecasting accuracy and interpretability by introducing attention mechanism and integrated learning strategy. Attention mechanism enhances the attention of the model to important information and provides an intuitive explanatory perspective; However, ensemble learning improves the overall stability and generalization ability by combining the predictions of multiple models.

Although some achievements have been made in this study, there is still room for improvement. For example, we can further explore more complex model architecture (such as Transformer), incorporate more kinds of market data (such as news sentiment analysis) and develop more elaborate interpretable tools to comprehensively improve the forecasting ability of the model.

5. Conclusions

This study tackles stock market forecasting by investigating LSTM-based models and their enhancement techniques. We constructed three models: a basic LSTM, an LSTM with an attention mechanism, and an LSTM employing an ensemble learning strategy. These were compared and analyzed for their performance in regression and fluctuation predictions. The results indicate that incorporating an attention mechanism and ensemble learning significantly boosts the LSTM model's prediction accuracy. The attention mechanism improves prediction precision by dynamically weighting different time steps and features, offering an intuitive and interpretable view of the model's prediction process. Meanwhile, the ensemble strategy mitigates single-model bias and variance, enhancing overall stability and generalization.

In summary, the LSTM-based stock market forecasting model, augmented with these improvements, excels in both accuracy and interpretability. Future research could delve into more sophisticated model architectures and integrate a wider array of market data.

References

[1] Kumar G, Jain S, Singh U P. Stock Market Forecasting Using Computational Intelligence: A Survey[J]. Archives of Computational Methods in Engineering, 2021, 28(3):1069-1101.

[2] Zhang Yongan, Yan Binbin. A deep learning compound forecasting model of the stock market [J]. Computer Science, 2020, 47(11):255-267.

[3] Zhang Qianyu, Yan Dongmei, Han Jiatong. Research on Stock Price Forecasting Based on Deep Learning and Decomposition Algorithm [J]. Computer Engineering and Application, 2021, 57(05): 56-64.

[4] Lin Peiguang, Zhou Jiaqian, Wen Yulian. SCONV: A method of forecasting financial market trends based on sentiment analysis [J]. Computer Research and Development, 2020, 57(08): 1769-1778.

[5] Zhang Xiaoxu, Gao Zhentao, Wu Lei, et al. Stock price prediction based on hybrid quantum-classical neural network model [J]. Journal of University of Electronic Science and Technology of China, 2022,51(01):16-23.

[6] Zhang Yong, Zheng Fengqi, Yang Xingyu, et al. Prediction of financial stress index based on long-term and short-term memory neural networks [J]. Systems Engineering, 2023, 41(5):115-123.

[7] Yang Xu, Liu Jiapeng, Yue Han, et al. Research on Portfolio Strategy and Automated Trading Based on Deep Reinforcement Learning Algorithm [J]. Modern Electronic Technology, 2024,47(06):154-160.

[8] Zhao Cheng, Qin Chu, Liu Wei, et al. Composite data stock trend prediction based on two-stage dimension reduction [J]. Shanghai Management Science, 2023, 45(3):117-124.

[9] Jiang Huimin, Chen Feng. Financial Time Series Modeling and Forecasting Based on Deep Learning and Knowledge Map [J]. Microcomputer System, 2022,43(09):1793-1800.

[10] Cong Jingqi, Cheng Pengfei, Zhao Zhenjun. Integrated forecasting model of stock index based on CEEMD-CNN-LSTM [J]. Systems Engineering, 2023, 41(4):104-116.