Application of an LSTM-Transformer Hybrid Model Incorporating Public Sentiment in Predicting CSI 300 Stock Prices

Zhao Hanwen
School of Science, Beijing Information Science and Technology University, Beijing, 102206, China
hwzhao_432@163.com

ABSTRACT

The short-term price changes in the stock market are influenced by many factors, so it is difficult to predict. In this paper, the daily trading data of CSI 300 index from 2016 to 2025 and the emotional text of stock forum are used as inputs to construct the LSTM-Transformer hybrid model. It uses CFSD China Financial Emotion Dictionary to quantify emotions, and evaluates the predictive advantages of emotional information through ablation experiments. The results show that the average MAE of the mixed model is 77.59, RMSE is 107.81, and r is 0.9218, which is better than the LSTM and Transformer models alone. After adding emotional characteristics, the average MAE decreased by 21.6%, RMSE decreased by 15.4%, and R decreased from 0.8715 to 0.9146, and remained stable in the 1-5-day forecast. This study provides a practical solution to predict, mix sensory data with mixed models.

Keywords: stock price forecasting; LSTM-Transformer hybrid architecture; public sentiment data; multi-source data fusion; multi-step forecasting; CFSD dictionary

1. INTRODUCTION

1.1 Research Background and Significance

The short-term price fluctuation of the stock market is influenced by a complex mixture of many factors. Although accurate prediction is very important for investors' decision-making and financial market stability, it is still the main challenge of financial research. At present, the research on stock price forecasting mainly focuses on model optimization and data expansion, and gradually shifts from traditional linear model to intelligent model, mixed architecture and multi-source data fusion. But there are still some big problems: it is difficult for a simple model to consider both time dependence and overall stock price correlation. The coordination mechanism of the mixed model has not been fully optimized; Moreover, merging data from multiple sources is usually a simple addition, which can not correctly match the data characteristics with the model architecture. At the same time, with the rise of social media, investor sentiment has become an important variable affecting stock prices, and the limitations of traditional forecasting-mainly focusing on historical transaction data, but ignoring emotional factors-have become increasingly clear.

Theoretically, this study attempts to build a hybrid LSTM-Transformer model to understand how to capture time dependence and extract global features, thus providing a new method to view stock price forecasts. In addition, by adding public opinion data and verifying its nonlinear advantages, this study enriches the theory of merging data from multiple sources, and shows how public opinion and stock price fluctuations are interrelated in complex ways. In terms of practical importance, this study uses CSI 300 index to forecast short-term stock prices in multiple stages, providing investors with more useful decision-making tools. Finally, the integration of public opinion and forecasting model developed by this research with various sources provides practical technical solutions for financial institutions in monitoring market risks and formulating quantitative strategies.

1.2 Literature Review

Short-term prediction of stock market trends using mathematical models has long been an important topic in financial research. In the past, researchers often used classical mathematical models, including autoregressive model (AR) and moving average model (MA), and their mixture: integral autoregressive moving average model (ARIMA). However, because there are many factors that affect the changes of the stock market, and most of these relationships are nonlinear, the accuracy of the traditional time series model is limited, and the results are usually average [1].

With the increasing use of machine learning and deep learning models in the financial field, many researchers in China and abroad have tried to use these models to predict stock prices. For example, Vizat Khan and his colleagues used several machine learning algorithms, among which SMO algorithm was the most effective, with a prediction accuracy of 68.56% [2]. Pooja Mehta and his colleagues used machine learning and deep learning models to make predictions, and found that the LSTM model performed better than some traditional machine learning models [3].

However, the model itself has some limitations. Therefore, using many mixed models to take advantage of different models may help to improve the accuracy of the final prediction. Pratyush Ranjan Mohapatra and his colleagues used CNN method to select features, and then applied the combined model of XGB and LSTM to predict these features, thus improving the accuracy [4]. Anindya Sarker and her colleagues combined VAE, Transformer and LSTM models to find the linear and nonlinear relationships in stock price changes [5]. MDR Kabir and his colleagues used LSTM-mTrans (mTransformer) and MLP model [6].

In addition to the continuous improvement and updating of the model algorithm, the data used to predict the stock market is no longer just the historical data of the stock. Researchers such as Ning Zhang extend this method by adding stock market sentiment data, and use CNN-LSTM mixed model to analyze these data [7]. Jihwan Kim and his colleagues use FinBert to analyze the emotions of financial texts [8]. Sri Varsha Mulakala and his colleagues use the pre-trained Transformer model to process comments and emotional data, while FinBert is used to analyze abstracts and news headlines [9]. Guo Kaifeng and his colleagues focused on the relationship between social media sentiment and the stock market, trying to use deep learning to discover the impact of Twitter sentiment on stock prices. They adjusted RoBerta on a set of stock market sentiment data to determine the mood of tweets (positive, neutral, negative). Finally, they use RNN combined with historical stock prices and sentiment analysis to predict price trends [10].

In short, the research of stock market forecasting has turned to intelligent model, hybrid architecture and multi-source data fusion. However, there are still some problems: for model integration, most hybrid models have no cooperation mechanism between LSTM and Transformer. For the quantification of emotions, there is a lack of standardized quantification methods suitable for financial texts. For the ablation test, most studies using public emotional data failed to isolate the marginal contribution of emotions through rigorous ablation experiments, and ignored the time alignment. In order to solve these problems, this paper constructs the LSTM-Transformer collaborative hybrid architecture, designs the emotional quantification process of the stock forum based on the CFSD dictionary, and conducts ablation experiments under strict time alignment to quantify the predictive gain of emotional information.

1.3 Scope of the Study

This study discusses the main problems of dynamic stock price prediction. We advance step by step in logical order: "model improvement-data addition-merger verification".

Firstly, in order to solve the problem that a single model can't capture the time dependence and the overall stock price correlation at the same time, we design the LSTM-Transformer hybrid architecture. By comparing the pure LSTM model with the pure Transformer model, we show all the advantages of the hybrid architecture for feature extraction.

Secondly, because only using historical trading data can not fully show the influence of market sentiment on stock prices, we have added sentiment data as an additional source. However, because emotional information is unstructured text, it can't be directly used in forecasting model, so we created a special quantitative processing process for the stock market environment. This process converts the text into standardized emotional features, and analyzes their relationship with stock price changes, providing a database for the follow-up.

Finally, in order to verify whether the emotional data is really useful, we conducted an ablation experiment based on the hybrid architecture. We only use historical data (control group) and emotional data (experimental group) to build a prediction model. By comparing the differences between them, we measure the marginal gain of emotional information to predict the stock price. Therefore, we have conducted a complete research cycle, including model improvement, data processing and fusion evaluation.

2. DATA SOURCES AND PREPROCESSING

2.1 Data Sources

2.1.1 Historical Transaction Data

In this study, the daily transaction data of CSI 300 index (code: 000300) from 2016 to 2025 was retrieved through AkShare financial interface. These data cover five important indicators-opening price, highest price, lowest price, closing price and trading volume-including a total of 2,430 original records.

2.1.2 Public Opinion Data

The public opinion data used in this study comes from two sources: posts from January 2016 to June 2024 come from Kaggle's public data set "Chinese Stock CSI 300 Guba Forum Comments 2015-2024" (data set approved by MIT), which contains complete titles and post texts. Publications from 2024 to July 2025 were recovered from East Money Stock Forum using Octopus crawler. Because of the anti-crawling mechanism and the dynamic loading technology of website pages, we can only get the title of the publication; We can't collect a lot of texts.

For the data as of July 2024, this study uses a strategy: we use the title instead of the text. The main reasons are as follows: firstly, in recent years, publishing on mobile devices has become the norm, and users of stock forums are often accustomed to "the title is positive" when publishing on mobile devices; Many posts only fill in the title and leave a blank body, or the body just repeats the title. Then, even if the body has extra information, they usually feel the same as the title. Because the title shows the main idea of the user, it already has important information of emotional analysis. Thirdly, this study mainly regards emotions (positive/negative/neutral) as the main result, rather than a very fine understanding, and the title information is enough to reach this level.

According to us, this study still uses the title as the analysis text of posts published after July 2024, and retains the same preparation and functional work as before. When there is little data, this method helps to keep important information to public opinion as much as possible, which is reasonable and possible.

2.2 Data Preprocessing

2.2.1 Preprocessing of Historical Transaction Data

In order to make the data analysis results reliable, we prepared the original data. After verification, the data has no missing or strange values. In order to eliminate the influence of feature size difference on model learning, we use MinMaxScaler to place five features (opening price, high price, low price, closing price and trading volume) in the [0,1] interval. The standardized formula is as follows:

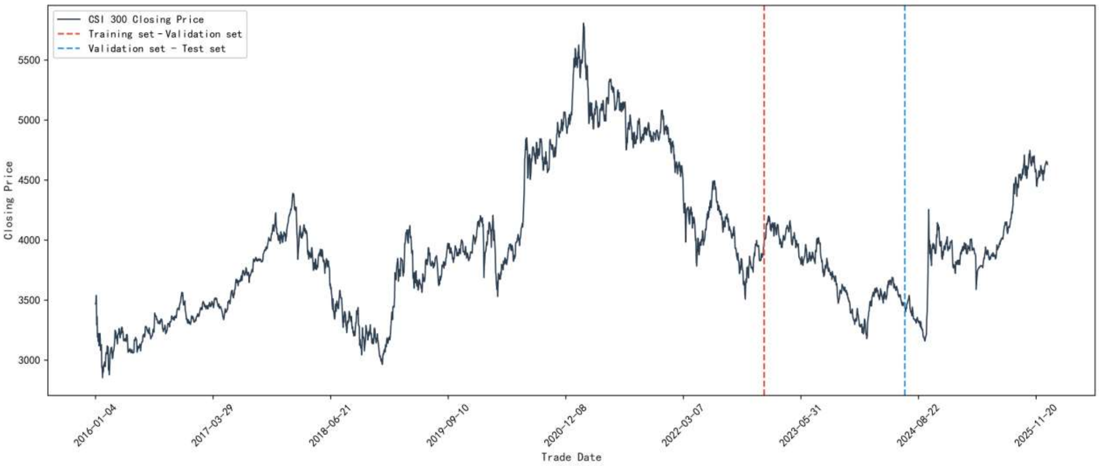
$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{1}$$

Here, x represents the original value of the feature, and xmin and xmax are the minimum and maximum values of the feature.

After standardization, we use sliding window method to construct time series samples. In this study, the length of the time window is set to 30 days, and the forecast horizon is set to 5 days. In other words, we use the characteristics of the past 30 trading days to predict the closing price of the next 5 trading days. A total of 2 395 valid time series samples are generated by sliding window, and the sample shapes are (2395, 30, 5). Then, we divide the samples into a training set (1,676), a verification set (359) and a test set (360) with the ratio of 7: 1 and 5: 1.5, which are arranged in strict chronological order to ensure that future data will not be used for training the model and avoid data leakage.

In order to show how the data is distributed and why we divide the data into several parts, we draw the closing price over time in the chart, and the result is shown in Figure 1. The graph shows the price changes of CSI 300 index from 2016 to 2025. The red dashed vertical line shows the separation of training set and verification set (January 3, 2023), and the blue dashed vertical line shows the separation of verification set and test set (July 1, 2024). This verifies that the time is continuous and the data is cut correctly.

Figure 1. Closing Price Trends of the CSI 300 Index from 2016 to 2025 and Data Set Segmentation



2.2.2 Preprocessing of Public Opinion Data

We cleaned up the raw data of public sentiment in several steps: when the main text is blank or blank, we will fill it with headlines. Delete short texts with less than 2 characters (548 entries) after filling in. We use the release time and the main text filled in as common keywords to eliminate duplicate data; We normalized the time format and extracted the date field. After pretreatment, we obtained 57,430 valid data items, and the cleaning rate was 0.95%.

Special characters and unnecessary spaces have been deleted by regular expressions. Use a combined dictionary of general keywords and stock forum specific keywords to filter and delete words that have nothing to do with emotion or investment; The stock market terminology dictionary is expanded by the segmentation tool Jieba, and it is segmented in an accurate mode to filter words less than 2 characters. After feature engineering, the final effective data set count is 56,289 records, and the well-structured segmentation results match the characteristics of stock market texts.

3. EVALUATION CRITERIA AND RESEARCH METHODS

3.1 Evaluation Criteria

This study uses three indicators-mean absolute error (MAE), mean square root error (RMSE) and determination coefficient (R^2)-to clearly understand the prediction error of the model and how it adapts to the data. MAE measures the average difference between the predicted value and the actual value, showing the total accuracy. RMSE is more sensitive to big mistakes and shows the stability of the model. R^2 tells us the degree to which the model can explain data changes, and the closer it is to 1, the better. The calculation formula is:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|\hat{y_i} - y_i| \tag{2}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y_i} - y_i)^2} \tag{3}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y_i})^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{4}$$

Here, y_i represents the actual closing price of the i-th sample, $\hat{y_i}$ represents the model's predicted closing price for the i-th sample, $\bar{y}$ represents the mean of all actual closing prices, and N represents the total number of samples in the test set.

3.2 Research Methods

3.2.1 LSTM Model

Lstm (long short-$ TERM memory) is helpful to solve the problem of gradient generalization in normal RNN. It uses gate mechanism (input gate, forget gate and output gate), which makes it beneficial to learn the long-term dependence of stock price. The LSTM model we used in this study has two stacked layers, the hidden layer size is 64 and the dropout rate is 0.2.

3.2.2 Transformer Models

Transformer is based on self-attention mechanism, which allows parallel extraction of global correlation features in sequences. In this study, a double-layer converter encoder is used, which has four points for attention, and the feedforward network dimension is 64, and a position coding is added to keep time information.

3.2.3 LSTM-Transformer Hybrid Architecture

The hybrid model combines the advantages of LSTM and Transformer: First, we use single-layer LSTM (input dimension = 5 or 6, hidden layer dimension = 64) to extract the short-term time features of the last 10 time steps in the first 30 time steps. Then, these short-term features are propagated throughout the sequence and connected with the original features. After being converted into a 64-dimensional space through the linear layer, the data is sent to a single-layer conversion encoder (note that the number of heads = 8 and the size of feedforward network = 128) to extract the global correlation features, and finally the closing price prediction of the next five time steps is obtained through the output layer. This design allows the model to capture the details of short-term fluctuations and learn long-term dependencies.

3.2.4 Methods for Quantifying Sentiment

Emotional score is calculated by using China Financial Emotional Dictionary (CFSD). This dictionary is constructed from the corpus of the financial industry, and contains 1,109 words of positive emotions and 1,488 words of negative emotions. For each text, we first use Jieba word segmentation for accurate pattern segmentation. Then calculate the frequency of positive and negative emotional words respectively, and use the following formula to calculate the emotional score:

$$Sentiment\ Score = \frac{Frequency\ of\ positive\ words - Frequency\ of\ negative\ words}{Frequency\ of\ positive\ words + Frequency\ of\ negative\ words} \tag{5}$$

The range score is [-1,1], in which positive values represent positive emotions and negative values represent negative emotions; Absolute value indicates the intensity of feeling. If no sensory word is found in the text, the sensory score will be marked as 0 (neutral). The daily public perception score is calculated by taking the arithmetic average of the perception scores of posts aggregated daily.

4. EMPIRICAL ANALYSIS

All models were implemented using the PyTorch framework, employing the Adam optimizer (learning rate = 0.001, weight decay $\lambda = 10^{-5}$), with the mean squared error (MSE) as the loss function, a batch size of 32, and 50 training epochs. To ensure the reproducibility of the experiments, this study fixed the random seed to 44 prior to model training. Control experiments maintained all configurations consistent except for the input feature dimensions, ensuring the comparability of results.

4.1 Multi-step Stock Price Forecasting Based on an LSTM-Transformer Hybrid Architecture

4.1.1 Overall Performance Comparison

In this study, three models are used for experiments: pure LSTM, pure Transformer and mixed LSTM-Transformer to see if the mixing is better. The average performance of the test set is shown in Table 1.

Table 1. Comparison of Average Performance Across the Three Model Test Sets

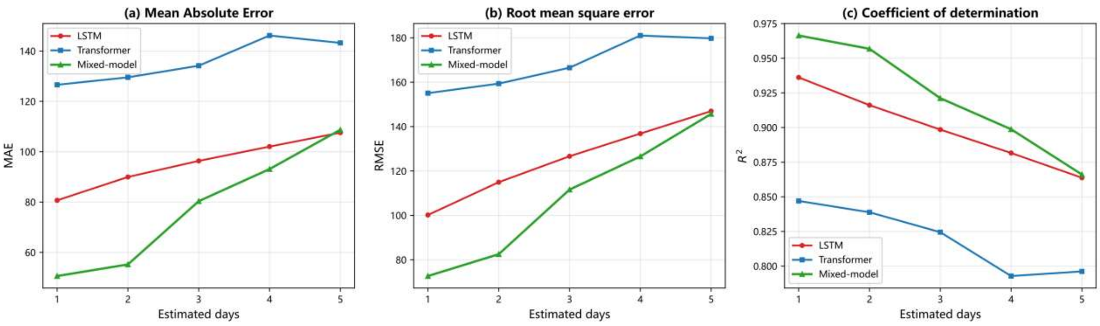
Model Type	Average MAE	Average RMSE	Average R^2
LSTM	95.31	125.11	0.8992
Transformer	135.94	168.34	0.8198
LSTM-Transformer Hybrid Model	77.59	107.81	0.9218

As shown in Table 1, the LSTM-Transformer hybrid model performs best in all tests: the average value of MAE is 77.59, which is 18.6% lower than the pure LSTM model (95.31) and 42.9% lower than the pure Transformer model (135.94). The average RMSE is 107.81, which is 13.8% lower than the pure LSTM model (125.11) and 36.0% lower than the pure Transformer model (168.34). The average value of R is 0.9218, which is higher than 0.8992 for pure LSTM model and 0.8198 for pure Transformer model. These results show that the design of the hybrid model is good. The LSTM part can learn short-term dependencies well, while the Transformer part can find big trends well. Together, they help us better understand how the stock price changes.

4.1.2 Multi-step Prediction Stability Analysis

The core challenge of multi-step stock price forecasting lies in the fact that "the longer the forecast horizon, the more significant the cumulative error becomes." This section compares the performance of three types of models over 1–5-day forecast horizons using three metrics: MAE, RMSE, and R^2.

Figure 2. Comparison of Multi-Step Prediction Performance Among Three Model Types



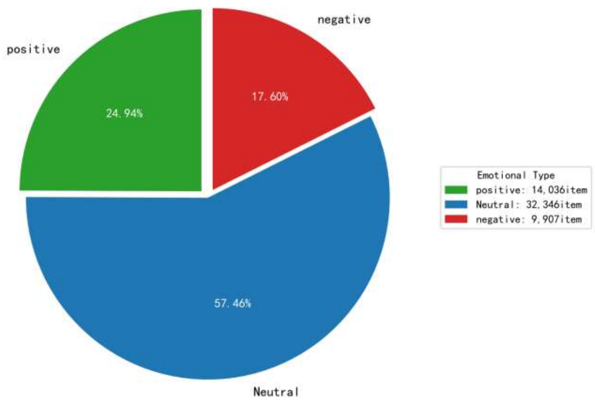
As shown in Figure 2, all three model types exhibit the common pattern of "increasing error and decreasing goodness of fit." The hybrid model significantly outperformed both the pure LSTM and pure Transformer models across all prediction steps: on Day 1, the hybrid model's MAE was 60.0% lower than that of the pure Transformer and 37.3% lower than that of the pure LSTM, with an R^2 of 0.9664; From days 2 to 4, the hybrid model continued to perform best; on day 5, its performance was roughly on par with the pure LSTM. In terms of the R^2 metric, the hybrid model maintained a lead from days 1 to 4. These results indicate that the hybrid model demonstrates a significant advantage in short-term (1–3 days) forecasting, effectively integrating temporal dependencies with global features to improve prediction accuracy and goodness of fit.

4.2 Quantitative Analysis of Public Sentiment Data and Analysis of Its Correlation with Stock Prices

4.2.1 Results of Sentiment Analysis

After completing the sentiment annotation, we performed statistical visualization of the sentiment distribution across 56,289 valid public opinion data points.

Figure 3. Pie Chart Showing the Distribution of Sentiment Types in Public Opinion Regarding the CSI 300 Index

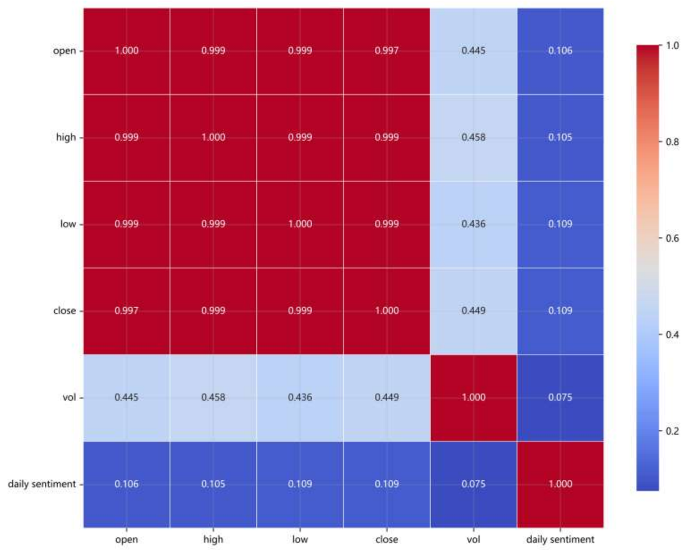


As shown in Figure 3, positive emotions accounted for 24.94% (14,036 posts), neutral emotions accounted for 57.46% (32,346 posts) and negative emotions accounted for 17.60% (9,097 posts). Generally speaking, the emotions in the stock forum show a distribution characterized by “neutral emotions as the leading category, followed by positive emotions and negative emotions as the lowest”. This result shows that the posts of stock forum users are mostly facts and neutral views, and rarely express positive or negative emotions. The number of positive messages is slightly higher than the number of negative messages.

4.2.2 Analysis of the Correlation Between Public Sentiment and Stock Prices

The daily average emotional score of public opinion data is matched with the daily trading data of CSI 300 index in time dimension, with a total of 2,422 valid data points. We analyzed the degree of linear correlation between variables using Pearson’s correlation coefficient.

Figure 4. Heatmap of the Correlation Between Public Sentiment Indicators and Stock Price Indicators



As shown in Figure 4, the Pearson correlation coefficients among the four price indicators (opening price, high price, low price and closing price) are all greater than 0.997, showing a very strong positive correlation. The display of trading volume is positively correlated with the moderate price index. In contrast, the absolute value of the correlation coefficient between the daily average emotional score and each type of stock price index is less than 0.11, which indicates that the correlation level is very low. This shows that there is no significant linear relationship between public sentiment and stock price fluctuation; On the contrary, there may be complex nonlinear or staggered correlation between them. This provides a reasonable basis for introducing mixed LSTM-Transformer model to discover potential nonlinear features.

4.3 Empirical Study of a Stock Price Forecasting Model Incorporating Public Sentiment Data

It is worth noting that the data used in the ablation experiment in this section are a small number of historical transaction data and date-related public sentiment data (2,422 trading days in total, 2,387 time series samples were generated after using the 30-day sliding window), while the complete experiment in section 4.1 is based on all historical trading days (2,430 trading days, 2,395 samples were generated after using the sliding window). Therefore, the performance of the control group in Table 2 cannot be directly compared with the hybrid model in Table 1-they serve different basic goals. In Table 2, the control group and the experimental group were compared under the same data conditions, so as to ensure that the contribution to sensory characteristics was properly quantified.

4.3.1 Overall Performance Analysis

In this section, using the LSTM-Transformer mixed model that we tested before, we compare the control group (only using 5 historical trading features) with the experimental group (5 trading features plus 1 emotional feature) to see the forecast gain brought by emotional features. The two groups used the same training settings.

Table 2. Comparison of the Impact of Public Sentiment Characteristics on the Performance of Stock Price Prediction Models

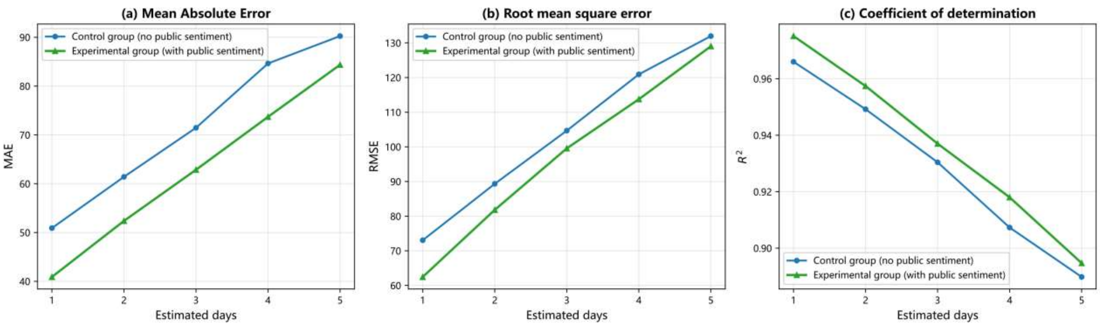
Model	Average MAE	Average RMSE	Average R^2
Control group (no public sentiment)	107.21	135.86	0.8715
Experimental group (with public sentiment)	84.01	115.00	0.9146

As shown in Table 2, the prediction performance of the experimental group (using public sentiment characteristics) is obviously better than that of the control group: the average MAE decreased from 107.21 to 84.01, a decrease of about 21.6%; The average RMSE decreased from 135.86 to 115.00, a decrease of about 15.4%. The average value of R^2 increased from 0.8715 to 0.9146, which indicated that the model had better adaptability. This tells us that the characteristics of public sentiment can help reduce stock price forecasting errors and improve the performance of the model. Therefore, it can be clearly seen that the use of public sentiment data is helpful to predict stock prices.

4.3.2 Multi-step Prediction Stability Analysis

Further analyze the gain in performance achieved by incorporating public opinion features at different prediction stages.

Figure 5. Comparison of multi-step forecasting indicators with and without public sentiment features



As shown in Figure 5, the experimental group (with public sentiment) has better prediction performance than the control group (only with past data). It has an average of 21.64% less than MAE, 15.36% less than RMSE, and 4.95% more than R^2. When the control group made a major mistake (such as the fourth day), the experimental group remained stable. It shows that public sentiment can correct the error of the model, and it proves that the data of public opinion can help to predict the stock price.

5. CONCLUSIONS AND OUTLOOK

5.1 Key Findings

This study involves CSI 300 index, and studies the optimization of mixed architecture and data merging from multiple sources. It solves three important problems in forecasting stock prices: how to consider time and global characteristics at the same time; How to measure public sentiment data; And how to use all this data together.

The hybrid architecture shows great advantages: the LSTM-Transformer hybrid model is superior to the pure LSTM model, the pure Transformer model with MAE, RMSE and R^2 indicators, and the average MAE of the hybrid model is 77.59, which is 18.6% lower than that of the pure LSTM model and 42.9% lower than that of the pure transformer model; The average R^2 is 0.9218. For multi-day forecasting, the mixed model keeps the best state from the first day to the fourth day, and the R^2 of the first day is 0.9664. This shows that LSTM can capture short-term time dependence well, while Transformer can extract global features, and they are very powerful together.

The method of emotion quantification is effective: we have created a process of emotion quantification, which is suitable for the stock market environment, and transformed unstructured emotional texts into standardized emotional features based on CFSD China Financial Emotion Dictionary. The analysis of emotional distribution shows that the emotions in the stock market forum follow a pattern: neutrality is the dominant emotion, active tracking, and negative emotions are the lowest, with positive emotions accounting for 24.94%, neutrality accounting for 57.46% and negative emotions accounting for 17.60%. Correlation analysis shows that there is no significant linear relationship between emotion and stock price (all absolute correlation coefficients are less than 0.11), which provides a reasonable basis for further exploring the nonlinear relationship.

The emotional data increased the value: compared with the control group, the average MAE decreased by 21.6%, the average RMSE decreased by 15.4% and the average R^2 increased by 0.8715 to 0.9146 in the experimental group. In the multi-stage prediction, especially when the control group shows significant prediction error, the experimental group remains stable, which shows that the public sentiment characteristics can effectively correct the model error and verify the value-added effect of integrating public sentiment data into stock price prediction.

In a word, the LSTM-Transformer hybrid architecture and public sentiment integration strategy proposed in this study have achieved good results in forecasting the short-term stock price of CSI 300 index. This provides useful hints for using data from multiple sources and mixed models in financial time series forecasting.

5.2 Future Research Directions

Although this study has achieved satisfactory prediction results, there are still many things to be improved:

Optimization of the method of quantifying feelings: This study uses CFSD dictionary to quantify feelings. This can well identify the feelings in financial texts, but it is not good at understanding complex sentences (such as negative phrases and satire). Then, we can use the trained financial model, such as FinBERT, to further improve the accuracy and robustness of quantifying emotions.

Expand data merging from multiple sources: This study only uses historical trading data and stock forum sentiment data. After that, you can add other data, such as news bulletins, macroeconomic indicators and industry indexes, to build a more comprehensive feature system and explore how to merge different data.

Verify the generalization ability of the model: This study mainly focuses on CSI 300 index. Then, we can apply this model to individual stock forecasting under different market conditions to verify its generalization ability and robustness, and explore how to use graphic neural network and other technologies in stock price forecasting.

REFERENCES

[1] Ma Yuanyuan, Liu Yanze, Liu Chenglong, et al. A Multi-Perspective Analysis of Chinese Investor Sentiment and Its Role in Stock Market Forecasting [J]. Journal of Northeastern University (Natural Science), 2022, 43(8).

[2] Khan W, Malik U, Ghazanfar M A, et al. Predicting stock market trends using machine learning algorithms via public sentiment and political situation analysis[J]. Soft Computing, 2020, 24(15): 11019-11043.

[3] Mehta P, Pandya S, Kotecha K. Harvesting social media sentiment analysis to enhance stock market prediction using deep learning[J]. PeerJ Computer Science, 2021, 7: e476.
[4] Mohapatra P R, Parida A K, Swain S K, et al. Gradient boosting and lstm based hybrid ensemble learning for two step prediction of stock market[J]. Journal of Advances in Information Technology, 2023, 14(6): 1254-1260.
[5] Sarkar A, Vadivu G. An Advanced Ensemble Deep Learning Framework for Stock Price Prediction Using VAE, Transformer, and LSTM Model[J]. arXiv preprint arXiv:2503.22192, 2025.
[6] Kabir M R, Bhadra D, Ridoy M, et al. LSTM–Transformer-Based Robust Hybrid Deep Learning Model for Financial Time Series Forecasting[J]. Sci, 2025, 7(1): 7.
[7] Zhang N, Xiong J, Zhao Z, et al. Dose my opinion count? A CNN-LSTM approach for sentiment analysis of Indian general elections[J]. Journal of Theory and Practice of Engineering Science, 2024, 4(05): 40-50.
[8] Kim J, Kim H S, Choi S Y. Forecasting the S&P 500 index using mathematical-based sentiment analysis and deep learning models: a FinBERT transformer model and LSTM[J]. Axioms, 2023, 12(9): 835.
[9] Sarkar A, Vadivu G. An Advanced Ensemble Deep Learning Framework for Stock Price Prediction Using VAE, Transformer, and LSTM Model[J]. arXiv preprint arXiv:2503.22192, 2025.
[10] Guo K, Xie H. Deep learning in finance assessing twitter sentiment impact and prediction on stocks[J]. PeerJ Computer Science, 2024, 10: e2018.