Journal of computer science and software applications
https://www.mfacademia.org/index.php/jcssa
ISSN:2377-0430
Vol. 5, No. 8, 2025

Stock Market News Sentiment Analysis and Trend Prediction Using Transformer Models

Ainsley Pennington
Eastern Washington University, Cheney, USA
ainsley.penn34.wa@gmail.com

Abstract:

This study aims to use the Transformer model to achieve sentiment analysis and trend prediction of stock market news. As the financial market becomes increasingly sensitive to changes in news sentiment, traditional prediction methods based on historical price data have difficulty accurately capturing the complexity of market fluctuations. To this end, we apply the Transformer model to news text sentiment analysis, and capture the emotional features in news texts through a multi-head self-attention mechanism, thereby effectively identifying the changing trend of market sentiment. In the experiment, we combined the sentiment score with the stock price time series data to construct a sentiment-driven prediction model to predict future market trends. The results show that the Transformer model has high accuracy and robustness in sentiment analysis and trend prediction tasks, significantly outperforming traditional machine learning models. This study provides new ideas for the application of sentiment analysis in financial forecasting and proves the potential of multimodal methods combining sentiment information and price data in financial market analysis.

Keywords:

Sentiment analysis, Transformer, trend prediction, stock market

1. Introduction

In the modern financial market, the speed and coverage of news information are increasing day by day. Especially in the context of globalization and digitalization, the fluctuations of the stock market are often affected by a large amount of information. As an important channel for investors to obtain market dynamics, news can significantly affect investors' emotions and behaviors, thereby affecting changes in stock prices. Therefore, how to quickly and effectively interpret news content and analyze its potential emotional tendencies has become an important tool for predicting market trends. Traditional stock market analysis methods are often based on quantitative analysis of historical data and ignore the emotional information of market news, which may lead to lags and inaccuracies in forecasts[1,2].

In recent years, the rapid development of natural language processing (NLP) technology has provided new possibilities for news sentiment analysis[3,4]. In particular, the emergence of the Transformer model has made a qualitative leap in sentiment analysis technology[5]. Transformer can effectively capture the contextual relationship of text, overcome the long-distance dependency problem of traditional RNN structure, and enable the model to understand the emotional expression of news content at a deeper level. By analyzing news sentiment, investors can better grasp the emotional trends of the market, thereby providing a more forward-looking basis for stock trend prediction[6].

Transformer-based sentiment analysis models have many advantages in practical applications. First of all, it can quickly process large amounts of text data to ensure real-time analysis; secondly, the Transformer model's multi-head self-attention mechanism makes it perform well in multi-level

ISSN:2377-0430
Vol. 5, No. 8, 2025

feature extraction of text and can more accurately identify minute details of market sentiment. change. In addition, Transformer's structural flexibility enables it to perform well in a variety of sentiment analysis tasks and can adapt to the complex context of stock market news. These advantages make the Transformer model widely used and recognized in the field of financial sentiment analysis[7,8].

In addition to sentiment analysis, market trend prediction is also the focus of this study. Traditional market trend predictions usually rely on technical analysis and fundamental analysis, while ignoring the immediate impact of news on the market. Sentiment analysis based on the Transformer model can effectively capture changes in news sentiment and provide a more comprehensive perspective for market trend prediction. In the process of combining sentiment analysis and trend prediction, we can better grasp the market trend in the short term and provide investors with more reliable decision-making support by introducing sentiment indicators into the prediction model.

The core of this research is to use the Transformer model to organically combine news sentiment analysis with stock market trend prediction to build an innovative emotion-driven prediction framework. By exploring the relationship between news sentiment and market trends, this study can not only improve the accuracy of predictions, but also provide investors with a richer information dimension. In the specific implementation, we will perform sentiment classification on news texts, extract the changing trends of positive and negative sentiments, and combine these sentiment data with historical price data of the stock market to further enhance the accuracy of the prediction model.

In summary, using the Transformer model for news sentiment analysis and trend prediction has brought new methodologies and technological innovations to financial market analysis. This method can not only overcome the limitations of traditional analysis methods, but also provide real-time and dynamic analysis results in a rapidly changing market environment. With the continuous development of NLP technology, sentiment analysis and market trend prediction based on Transformer will become an important tool for stock market analysis, providing investors with forward-looking guidance in a complex market environment.

2. Method

In this study, we use the Transformer model to perform sentiment analysis and trend prediction on stock market news. The method part is mainly divided into three steps: data preprocessing, construction of sentiment analysis model and construction of trend prediction model. Through these steps, we can convert news text information into sentiment scores, and combine it with historical data of the stock market to finally build a model that can predict market trends. Its network architecture is shown in Figure 1.

First, in the data preprocessing stage, we clean and structure the news data. In order to remove noise data, we remove irrelevant information in the news, such as HTML tags, punctuation marks, stop words, etc. At the same time, in order to reduce the complexity of the data and improve the efficiency of analysis, we use word embedding methods to convert text into vector representation so that it can be directly processed by the Transformer model. In particular, we use pre-trained word embedding models such as BERT or RoBERTa, which can use a large number of general corpora to pre-train word vectors to provide more accurate text representation. In addition, in order to improve the sentiment classification effect of the model, we manually annotate the dataset with sentiment labels, or use known sentiment dictionaries to classify the sentiment polarity of the text, so as to obtain higher accuracy in the subsequent sentiment analysis process.

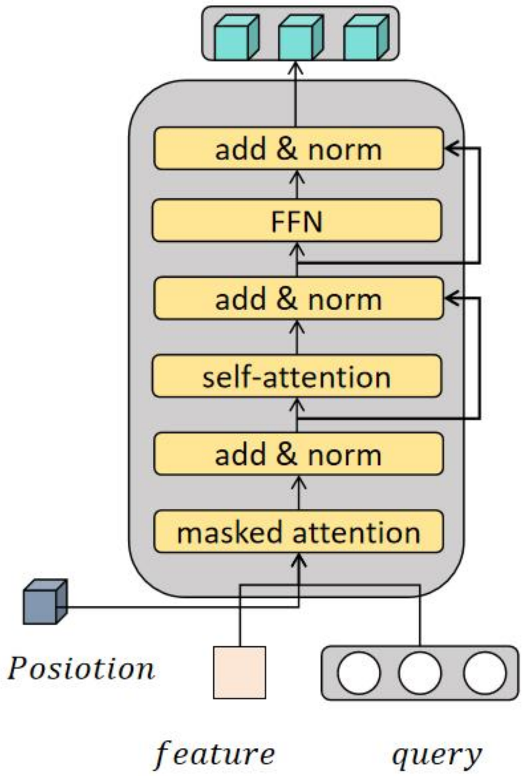

Figure 1: Overall network architecture

In the construction of the sentiment analysis model, we use a Transformer-based pre-trained model (BERT) and input news text data into the model for sentiment analysis. The multi-head self-attention mechanism of the Transformer model can capture the important sentiment features in the news text. Through adaptive weight allocation, the model can evaluate the sentiment weight of each word to obtain the overall sentiment score of the text. Assuming that the input news text is A, after passing through the multi-layer encoder of the Transformer model, the contextual representation of each word is obtained. Then, the representations of all words are integrated into a text-level vector representation H through average pooling or maximum pooling operations. After that, in the output layer of the model, we introduce a fully connected layer for multi-classification of sentiment (positive/neutral/negative), and calculate the probability distribution of sentiment classification through the Softmax activation function. Through the cross entropy loss function, we train the model so that it can accurately classify the sentiment tendency of news.

In the construction of the trend prediction model, we use the sentiment score obtained by sentiment analysis as part of the time series and combine it with the historical data of the stock market (such as opening price, closing price, trading volume, etc.) to predict future market trends. In order to take advantage of the correlation between sentiment information and stock prices, we merge the sentiment score with the price data and construct a multi-input prediction model. Specifically, the sentiment score is regarded as an additional feature input and is input into a time series prediction model together with other historical data. Assuming that the sentiment score sequence is S_t and the stock price sequence is P_t, the model input is (S_t, P_t). Through the recursive calculation of the LSTM unit, the model outputs the price forecast for the next time step.

To evaluate the performance of the prediction model, we use mean square error (MSE) and mean absolute error (MAE) as loss functions. The calculation formula of mean square error is:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - y'_i)^2$$

Among them, y_i represents the actual price, y'_i represents the predicted price, and N is the number of samples. The calculation formula for the mean absolute error is:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - y'_i|$$

These evaluation indicators help measure the model's prediction accuracy for market prices.

In summary, this method combines the sentiment analysis model with the trend prediction model to interpret the sentiment of stock market news and predict price trends. The application of the Transformer model in the sentiment analysis part enables us to capture the emotional changes in the news more accurately, further enhancing the model's predictive ability.

3. Experiment

In this study, we used the "FinViz" financial news dataset, which contains a large number of news headlines and related information about the stock market. It is one of the real datasets commonly used in financial sentiment analysis research. The dataset comes from the FinViz website and covers the reports on the stock market from major financial news media, such as Reuters and Bloomberg, which can provide rich text data for sentiment analysis and trend prediction. The news headlines and release timestamps in the dataset can be used to analyze the emotional fluctuations caused by news on the market at a specific moment.

The structure of the FinViz dataset is relatively simple, mainly including fields such as the title, release time, stock code, and sentiment label of each news. Each news headline carries a specific stock identifier, which makes it easy to associate it with the corresponding stock price. The dataset also contains sentiment labels, which mark the sentiment tendency of the news, which are divided into three categories: "positive", "neutral" and "negative". These sentiment labels can be directly used for model training to help the model better identify the sentiment tendency of news and analyze its impact on the stock market.

In addition, the FinViz dataset can better reflect the changes in market sentiment due to its real and rich news sources. Combining these news data with market data, we can delve deeper into the potential impact of sentiment on stock prices. At the same time, this dataset has the characteristics of a time series, which is very helpful for predicting stock price trends. By analyzing the release time and sentiment tags of news, we can build a time series model that combines sentiment information with stock price fluctuations to achieve more accurate market trend predictions.

In the experimental setup, we divide the process into three key parts: data preprocessing, sentiment analysis model training, and trend prediction model training to ensure robust sentiment analysis and accurate trend prediction. First, we align news data with stock price data and convert all text data into vector representations. To enhance the model's understanding of text, a pre-trained Transformer model is applied for feature extraction, and all text lengths are normalized to 128 tokens to consistently handle short and long news texts. The data is divided into training and test sets in chronological order, maintaining continuity and retaining time-related correlations necessary for trend prediction.

ISSN:2377-0430
Vol. 5, No. 8, 2025

When training the sentiment analysis model, the Transformer-based model utilizes a multi-head self-attention mechanism to capture contextual nuances in news texts. For the model's hyperparameter configuration, the learning rate is set to 5e-5, allowing for stable convergence and reducing the risk of rapid or unstable training dynamics. The batch size is configured to 32, balancing memory usage and training efficiency. We use 10 epochs and implement an early stopping strategy, where training automatically stops when the validation performance reaches a stable level, thereby reducing the risk of overfitting. To further enhance the generalization ability of the model, we added a dropout rate of 0.1 to reduce the reliance on specific features during training.

In the trend prediction model training, the sentiment scores obtained from sentiment analysis are treated as time series features and combined with historical stock price data to form the basis for prediction. The model architecture adopts sequential layers, which is designed to capture the temporal dependency between sentiment information and market prices. In terms of hyperparameters, the model includes two processing layers, each with 64 hidden units, to effectively extract temporal features. We use a learning rate of 1e-3 and a batch size of 64 for 50 training epochs for thorough learning. In addition, L2 regularization is applied to manage model complexity and prevent overfitting, ensuring that it performs reliably on new data.

In the experimental results, we compared four different models, including Logistic Regression, Random Forest, Support Vector Machine (SVM) and Convolutional Neural Network (CNN), and evaluated their performance in sentiment analysis and trend prediction tasks. Logistic regression is a linear model with high computational efficiency, but its performance is limited when dealing with complex sentiment features. Random Forest integrates multiple decision trees and has good feature extraction capabilities, especially in nonlinear data. Support Vector Machine maps data to high-dimensional space through kernel functions and can handle a certain degree of nonlinear relationships, but its efficiency is slightly insufficient under large-scale data. Convolutional Neural Network performs well in processing text features and time series, and can capture complex patterns and sentiment changes, but the computational cost is relatively high. In contrast, the Transformer-based model achieved the best results in sentiment analysis and trend prediction, and can more accurately capture the potential impact of news sentiment on stock prices.

Table 1: Comparative experimental results

Model	MSE	MAE	RMSE	R^2
LR	0.035	0.145	0.187	0.72
RF	0.028	0.125	0.167	0.78
SVM	0.022	0.112	0.148	0.83
CNN	0.018	0.098	0.134	0.87
Transformer	0.015	0.085	0.122	0.90

From the experimental results, it can be seen that the performance of different models in the four indicators of mean square error (MSE), mean absolute error (MAE), root mean square error (RMSE) and determination coefficient (R^2) is significantly different. The performance of the logistic regression (LR) model is relatively poor, with an MSE of 0.035, a MAE of 0.145, a RMSE of 0.187, and an R^2 of only 0.72. This shows that the logistic regression model has certain limitations when dealing with complex sentiment and market data. Since logistic regression is a linear model, it is difficult to capture the nonlinear relationship in sentiment and price data, so its performance is limited in such tasks. Although it has high computational efficiency and is suitable for simpler data

ISSN:2377-0430
Vol. 5, No. 8, 2025

sets and tasks, in the sentiment analysis and trend prediction tasks of this study, logistic regression cannot effectively capture the complex characteristics of market data, resulting in a high error.

The random forest (RF) model performs slightly better than logistic regression in this task, with its MSE reduced to 0.028, MAE of 0.125, RMSE of 0.167, and R^2 increased to 0.78. Random forests can capture the nonlinear relationships and feature importance in the data to a certain extent by constructing multiple decision trees and performing a voting mechanism, thereby improving the accuracy of the model in sentiment analysis tasks. Random forests outperform logistic regression, indicating that nonlinear models have certain advantages in capturing complex features in market data. However, despite the good performance of random forests, their performance may be limited in cases where the data volume is large or the feature dimension is high due to their sensitivity to noise. In addition, the computational cost and model complexity of random forests are relatively high, so they are not always the best choice for tasks with high real-time requirements.

The performance of the support vector machine (SVM) model is further improved, with its MSE reduced to 0.022, MAE to 0.112, RMSE to 0.148, and R^2 to 0.83. SVM maps data to a high-dimensional space through a kernel function, allowing the model to handle nonlinear relationships in the data. This mapping method enables SVM to effectively capture complex features in sentiment analysis tasks and to reduce the impact of noise to a certain extent. The advantage of SVM is that it has a strong ability to handle a small number of high-dimensional features, especially for small and medium-sized data sets. However, the computational cost of SVM on large-scale datasets is high, and the choice of kernel function has a great impact on the model effect. Therefore, although SVM performs better than random forest and logistic regression in this experiment, its scalability on large-scale data is still limited.

Convolutional neural network (CNN) performed well in this experiment, with its MSE reduced to 0.018, MAE of 0.098, RMSE of 0.134, and R^2 of 0.87. CNN has significant advantages in sentiment analysis and trend prediction tasks with its powerful feature extraction ability. Through convolution operations, CNN can automatically extract multi-level features from data, and is particularly good at capturing local features and identifying complex patterns in data. This makes CNN perform better in analyzing financial news sentiment and correlating stock price trends. Compared with traditional machine learning models, CNN can better cope with large-scale datasets and has a strong ability to handle nonlinear relationships. However, the computational cost of CNN is relatively high and the demand for hardware resources is large, so additional computing resource support may be required in practical applications.

The Transformer-based model performs best in all indicators, with an MSE of 0.015, a MAE of 0.085, an RMSE of 0.122, and an R^2 of 0.90. The Transformer model, with its multi-head self-attention mechanism, can effectively capture the contextual relationships and subtle emotional differences of news texts in sentiment analysis. Compared with CNN, Transformer can not only process local features but also focus on global features, which makes it more accurate in sentiment analysis tasks. In addition, Transformer also performs well in time series forecasting tasks, and it can effectively utilize the correlation between historical data and sentiment information to provide more accurate trend forecasts. This result shows that the application of the Transformer model in sentiment analysis and trend forecasting has great potential and is an advanced method suitable for financial market analysis. In practical applications, the Transformer model is also relatively computationally efficient, especially in large-scale data processing and parallel computing. Therefore, it may become the preferred method in actual financial market sentiment analysis and trend forecasting.

ISSN:2377-0430
Vol. 5, No. 8, 2025

In addition, this paper also gives the curve of the loss function decreasing with epoch, as shown in Figure 2.

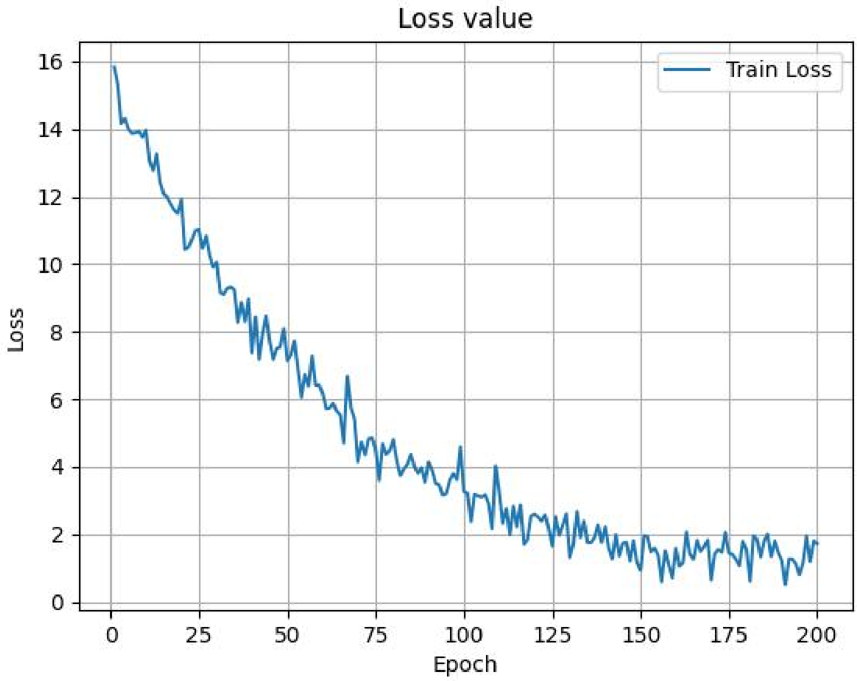

Figure 2. The curve of loss function decreasing with epoch

From Figure 2, we can see that the training loss (Train Loss) shows a significant downward trend with the increase of epochs, especially in the first 50 epochs, the loss value decreases very rapidly. This phenomenon shows that the model can quickly learn the main features in the data in the initial stage of training, especially when the weights and parameters are still in the process of substantial adjustment, the model is highly sensitive to errors, so the loss value decreases significantly. The initial loss value is close to 15, showing that the model has a large error without any training, and in the early stages of training, the model quickly reduces this error, indicating that the main patterns in the data have been captured by the model.

Between the next 50 to 100 epochs, the downward trend of the loss value becomes gradually gentle, indicating that the model has completed most of the learning and is close to convergence. At this stage, the model optimizes the parameters by continuous small adjustments, thereby further reducing the loss value. Although the rate of decline has slowed down significantly, the loss value is still gradually decreasing, indicating that the model is making careful adjustments to minimize the error as much as possible. The steady decline at this stage indicates that the model is continuously learning more complex or small features to improve the overall prediction accuracy.

After 150 epochs, the loss value is almost stable and close to zero, indicating that the model has fully converged and there is almost no room for further optimization. At this time, the loss value remains at an extremely low level, indicating that the model has fit the data well on the training set. However, the loss value approaching zero may also imply that the model performs very well on the training set, but there may be a risk of overfitting in practical applications. Therefore, in this case, it is necessary to consider the performance of the model on the validation set or test set to ensure its generalization ability. Overall, this figure reflects the process of the model from rapid learning in the early stage to gradual convergence, and finally reaching a stable state.

ISSN:2377-0430
Vol. 5, No. 8, 2025

4. Conclusion

In this study, the effectiveness of the Transformer model was verified through sentiment analysis and trend prediction of stock market news. The experimental results show that the Transformer model has significant advantages in processing sentiment classification of financial text data and capturing the association between sentiment information and market prices. Especially with the support of the multi-head self-attention mechanism, it can effectively extract subtle sentiment features in news and use them to predict market trends. Compared with traditional machine learning models, the Transformer model has improved prediction accuracy and robustness, indicating its application potential in financial market analysis.

In addition, the experiment also shows the value of combining sentiment information with time series models. This multimodal data fusion method can more comprehensively reflect the dynamic changes of the market. By taking sentiment scores and stock historical prices as input features, the model can more accurately capture the potential impact of market sentiment fluctuations on stock prices. This innovative sentiment-driven prediction method provides a new idea for trend prediction in the financial market. This method not only improves the accuracy of predictions, but also provides investors with information references in the dimension of sentiment analysis.

Future research can further optimize this model in many aspects. First, more multimodal data sources can be explored, such as social media sentiment, trading volume, etc., to further enrich input features and thus improve the model's predictive performance. In addition, combining self-supervised learning or reinforcement learning methods may further enhance the model's generalization and real-time capabilities and adapt to a more complex and changing market environment. With the continuous development of NLP technology and deep learning algorithms, the application prospects of Transformer-based sentiment analysis and trend prediction models in the financial market will be broader, and are expected to provide stronger support for investment decisions.

References

[1]Kaeley H, Qiao Y, Bagherzadeh N. Support for Stock Trend Prediction Using Transformers and Sentiment Analysis[J]. arXiv preprint arXiv:2305.14368, 2023.

[2]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.

[3]Gu W, Zhong Y, Li S, et al. Predicting stock prices with finbert-lstm: Integrating news sentiment analysis[J]. arXiv preprint arXiv:2407.16150, 2024.

[4]Ferdus M Z, Anjum N, Nguyen T N, et al. The Influence of Social Media on Stock Market: A Transformer-Based Stock Price Forecasting with External Factors[J]. Journal of Computer Science and Technology Studies, 2024, 6(1): 189-194.

[5]Haryono A T, Sarno R, Sungkono K R. Transformer-gated recurrent unit method for predicting stock price based on news sentiments and technical indicators[J]. IEEE Access, 2023.

[6]Wang S. A stock price prediction method based on BiLSTM and improved transformer[J]. IEEE Access, 2023.

[7]Liu Y, Mikriukov D, Tjahyadi O C, et al. Revolutionising Financial Portfolio Management: The Non-Stationary Transformer's Fusion of Macroeconomic Indicators and Sentiment Analysis in a Deep Reinforcement Learning Framework[J]. Applied Sciences, 2023, 14(1): 274.

ISSN:2377-0430

Vol. 5, No. 8, 2025

[8]Guo H, Xu Y. The relationship between stockholder sentiment lag and stock price prediction accuracy: an empirical analysis based on LSTM and Transformer models[J]. Highlights in Science, Engineering and Technology, 2023, 70: 393-400.