ISSN: 2395-7852

International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

Volume 11, Issue 2, March 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

IMPACT FACTOR: 7.583

International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 7.583 | Bimonthly, Peer Reviewed & Referred Journal|

| Volume 11, Issue 2, March 2024 |

A Stock Price Prediction Method Based on BILSTM and Improved Transformer

R.VIJAY, ATLA PENCHALA KRISHNA, SUNKARA SHESHU, VENKATA HAREESH JAGARLAMUDI

Assistant Professor, Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

ABSTRACT: In this study, five main sets of experiments were performed to verify the validity of the proposed model. First, we compared the prediction performance with that of various baseline models to investigate model feasibility. Second, four different variants of the TEA model were constructed and tested to analyze the impact of different components. Third, the effects of the attention mechanisms were intuitively investigated to confirm the contributions of the multihead attention and temporal attention mechanisms to the overall performance. Then, trading simulations were performed as a profitability test. Finally, we performed an error analysis to analyze TEA from various perspectives. The results indicate that the proposed model successfully achieves the stock movement prediction task with satisfactory experimental performance. Through comparisons with state-of-the-art methods, we find that the proposed method is more suitable for practical application and can help traders avoid financial risks and make more favorable decisions.

KEYWORDS: Structured multi-head attention, Stock prediction, Heterogeneous financial data

I. INTRODUCTION

In recent years, the use of social media information to predict the financial market has attracted the attention of more and more researchers, and some satisfactory experimental results have been achieved. This is due to the fact that social media information contains investor-related attitudes and subjective sentiments towards the financial market, resulting in many investment banks and hedge funds trying to dig out valuable information from media information to help better predict financial markets, which plays a key role in predicting the market. At the same time, the efficient market hypothesis (EMH) introduced by [1] points out that the current price of the asset reflects all the prior information that is immediately available. Therefore, using social media information and the actual price of the current financial market seems to be able to complete the market forecasting more accurately.

The data collected from social media is mainly text, which can be regarded as a natural language processing problem (NLP). Currently, commonly used texts are mainly public news and Twitter, and Twitter is more time-sensitive than public news. In addition, there is a large amount of data reflecting market sentiment in social media, which, of course, or indirectly, affects the trading psychology of investors and thus has an impact on stock prices. Therefore, many recent studies only use tweets to predict stock trends and at the same time explore the characteristic information of these texts at a deeper level [2, 3]. However, due to the time series characteristics of the financial market, historical data at different times has different impacts on the target trading day.

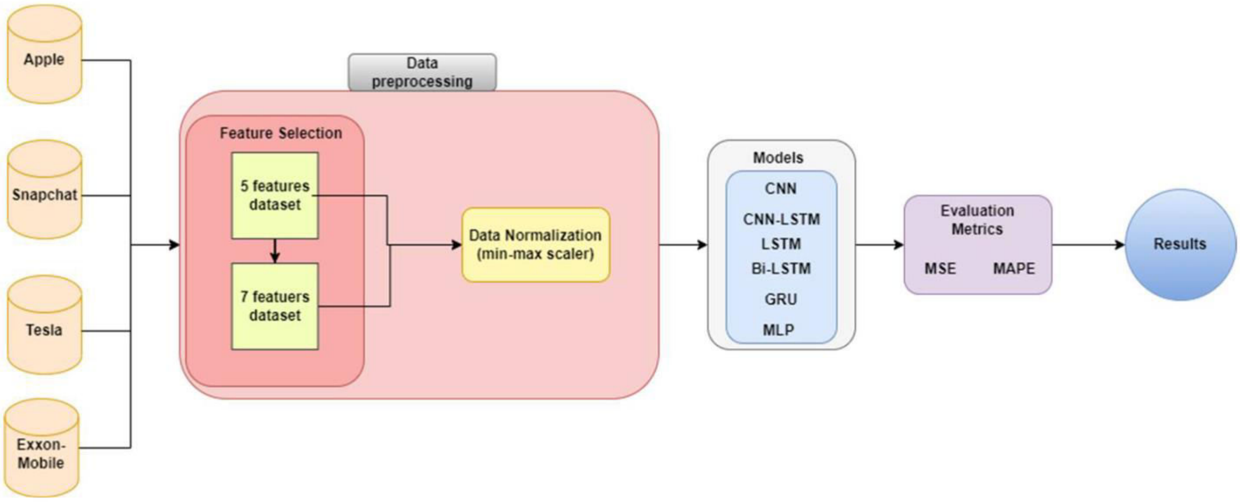

Fig 1: A New Stock Price model

Based on the abovementioned research, we can find that few research studies integrate text, and stock prices to realize stock market forecasts. Experiments show that the method of fusing tweets and actual stock prices with a small time window is superior to the prediction results generated by using text or stock prices alone [4, 5]. This reflects that the wisdom of the crowd comes from the interaction between individuals and the gathering of group opinions. The collective judgment of the group may be much better than the judgment made by a single person. However, currently, there are few studies using this method to predict the stock market; therefore, it is valuable to propose a research method that integrates social media information and stock prices.

Considering the current results, we still face some challenges: first, the stock market is regarded as time-series data, so the data is temporal dependent, and the value of historical information may decrease over time. In addition, the use of tweets and historical stock prices and other financial data to more accurately analyze and predict the rise and fall of stocks is still a problem worthy of in-depth discussion. In order to effectively solve the above problems, we propose a novel transformer encoder attention (TEA) network architecture, which is a network framework for deep extraction of financial data features and further integration, including a *feature extractor* and a *concatenation processor*, where the historical data uses the information of five calendar days from the target trading day as the training data. In addition, transformer is used to extract text features in-depth, and the attention mechanisms are used to obtain key information from financial data. Through the above framework, one can effectively predict the rise and fall of the target trading day.

For comprehensive experiments, we chose several evaluation metrics to verify the effectiveness of the proposed method in various situations. First, we investigated the feasibility of TEA by comparing its predictive performance with that of other baseline models. Second, we conducted an ablation study of the integrated TEA framework, including ablation of the input data and the model components. Then, visual experiments are presented to demonstrate the effects of various attention mechanisms. Through the performance comparison with related methods, the effectiveness of the method proposed in this paper is fully demonstrated. Furthermore, stock trading simulations are used to illustrate the potential application of the proposed model in actual market trading. Finally, we analyze the performance of the TEA model under special circumstances to detect weaknesses in this method.

II. RELATED WORK

In recent years, deep learning (DL)-based stock prediction methods have been extensively studied and satisfactory results have been achieved [5–7]. Vargas et al. [8] use financial news headlines and prices to estimate the movement of a certain day. The main framework uses a traditional convolutional neural network (CNN) and recurrent neural network (RNN) architectures. Sohangir and Wang [9] propose financial forecasting through DL methods, combined with stock Twitter data, to help investors make decisions. The approach proposes a more innovative analytical approach that provides insights for future solutions. Li et al. [10] proposed a unique approach to text-data fusion that was based on

LSTM's model, using four different sentiment dictionaries for a five-year stock market prediction experiment on stock technical indicators and text news data from Hong Kong. However, Lee and Kim [11] proposed a capsule network model based on the transformer encoder (CapTE), which uses the transformer encoder to extract the deep semantic features of social media texts, and then capture the structural relationships of these texts through the capsule network. In addition, attention mechanisms have been applied in many fields since their emergence, and they have now been applied to financial market research problems, especially the temporal attention mechanism, to adapt to the time series of financial data and obtain critical information. For example, [12] developed a hybrid attention network (HAN) to predict stock trends based on the sequence of recent related news items, which can significantly increase annual returns. With the help of tweets and stock prices that is based on incorporative attention mechanisms and multilevel local features, Hu et al. [13] proposed a stock price prediction network (SMPN). Wang et al. [14] have conducted an in-depth analysis of the timeliness and target sensitivity of stock investment reviews, the reliability of investors, public stock reviews, and their application in stock trend forecasting. The paper proposes a new framework based on dynamic experts' attitudes for aggregate stock trend forecasting, in which scores mainly use temporal attention to achieve decision-making. From the above research, it can be found that deep learning has shown great potential in the field of financial market forecasting, and it is expected that better results will be achieved in the future.

In summary, the use of the advanced DL framework to analyze financial data is an effective way to improve stock prediction results. The combined framework that we designed requires in-depth extraction of text and stock price features, and on this basis, can effectively integrate the extracted features of different types of information so as to completely realize the Accuracy of prediction and enhance its application in the actual financial market value. However, the current learning framework still has a lot of room for development in integrating text and numerical data. Therefore, we have designed a complete network framework that can process financial time series data and obtain key information from corresponding tweets. Among them, we adopted a transformer architecture that is completely different from the traditional architecture, mainly considering that it is not troubled by long-term dependence problems, avoiding information dependence problems, and thus greatly reducing training time. In addition, a small-time window, that is, the historical data of 5 days from the target trading day, is used for forecasting so as to obtain more valuable information and avoid focusing on too many resources. Correspondingly, we use an improved temporal attention mechanism to analyze financial data from the perspective of dependence and information content to extract features that are more useful for predicting the stock market. In addition, we adopted an auxiliary forecasting strategy, that is, using previous forecasts to assist in determining the trend of the target trading day. In general, the abovementioned deep learning framework is used to realize the fusion and prediction of text and stock prices so as to be applied to actual stock prediction scenarios.

III. METHODS

Next, we compare the baseline models and TEA on the composition of the framework. First of all, StockNet and Adv-LSTM adopt the bidirectional gated recurrent unit (BGRU) and LSTM structure, respectively, in the text feature extraction stage, while TEA adopts the transformer encoder; the subsequent architectures are basically the same. The results show that the transformer-based method has significant performance advantages over the traditional RNN method. At the same time, CapTE also uses the transformer encoder to extract deep text features. Overall, the experimental results of this method are indeed significantly better than other baseline methods. However, TEA's results are still better than other transformer-based methods, mainly because CapTE's source data only uses social media text information, without considering real stock market prices. In addition, by comparing the overall prediction results of the baseline model, it can be seen that there is still a lot of room for improvement in the performance of the HAN model and the Adv-LSTM model. The fundamental reason is the lack of practical applicability of the submodel and the onesidedness of reflecting emotions. It can be seen that choosing a suitable submodel has a crucial influence on the experimental results. Finally, the experimental results of the CH-RNN model based on RNN are still quite satisfactory. This model uses a variety of different types of attention mechanisms as the mainframe to conduct bidirectional analysis of the target text, thus demonstrating the advantages of the attention mechanism in obtaining key information. SMPN also adopts the idea of fusing text and prices with the attention mechanism as the main body and adopts the method of concatenation. Similarly, model1 considers news text and integrates emotional information, but the model it uses is a more traditional method and does not deeply integrate the two types of information.

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 7.583 | Bimonthly, Peer Reviewed & Referred Journal|

| Volume 11, Issue 2, March 2024 |

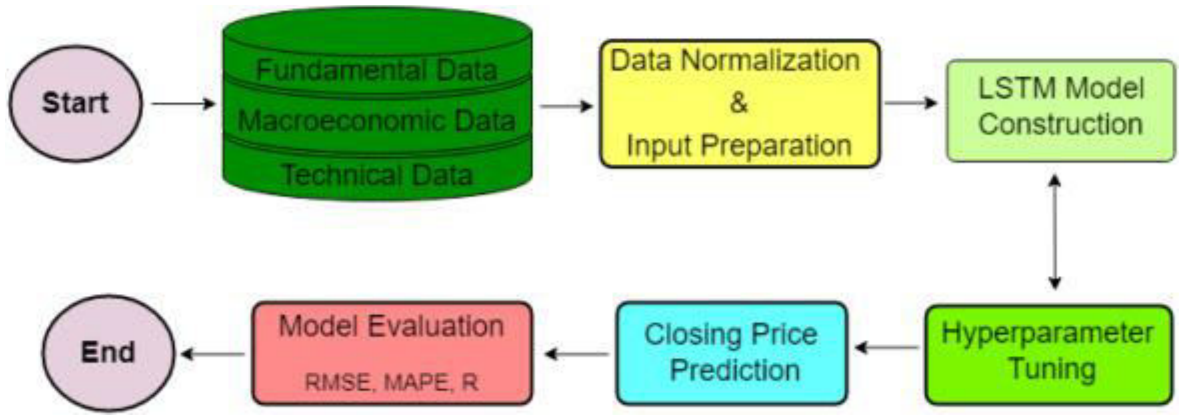

Fig 2: Predicting stock market index using LSTM

In recent years, there have been some new methods to solve the problem of stock prediction, especially when combined with knowledge graphs, we can see that LSTM-RGCN obtained great results using the graph. And the model of CNN-BiLSTM-AM combined the traditional method and deep learning model to predict the stock. Maybe how to combine them better is a problem to explore. And also, we find that FOCUS used the fuzzy system to achieve trading, it seems that this kind of method is a novel direction. However, the results of the above methods behave poorer than our method, there are some reasons: first, there are differences in extracting features, which may play an important influence on prediction; second, historical data and time window are also key factors. It is observed that our method obtained satisfactory results on three datasets, which demonstrate the great effect.

IV. RESULT ANALYSIS

In order to further illustrate the impact of the information used in the improved temporal attention mechanism and the dependent scores on the actual stock market, we observe the relationship between stock price changes and tweet information. In Figure 4, we have plotted the rise and fall of three stocks over time. It can be concluded that the trend of stocks is not only affected by historical stock prices but also by tweets, which directly proves that the improved timing attention mechanism we proposed is comprehensive. In fact, investors can carefully synthesize relevant investor comments and prices to better assess their impact on the underlying stocks. Therefore, the ideal framework for simulating this analysis process should integrate and interpret such information over a continuous period of time, rather than analyzing these factors separately.

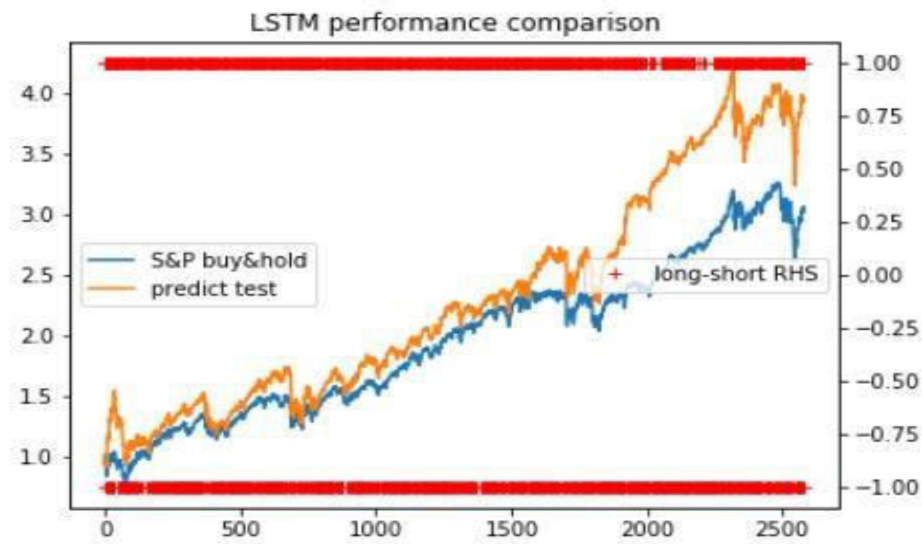

Fig 3: predict price Result analysis

In this article, we compare the performance of TEA and CapTE and analyze the situations where TEA produces wrong results but CapTE is correct. There are two scenarios: firstly, if a trader writes a tweet with negative sentiment, for example, the tweet "Mobile phones are depreciating faster than Vanke A is falling," the comment was actually talking about the devaluation of mobile phones, but the trader's sentiment extended to the stock, leading to a malicious comment about the corresponding stock. Secondly, comments about market issues from long ago may appear on Weibo. For example, the financial tsunami discussed in "The 2008 US-induced financial tsunami caused Chinese stock prices to fall by almost 80%" occurred in 2008 and is not relevant to the current stock market, yet in this case, it would be difficult for our model to obtain correct predictions without introducing relevant information.

V. CONCLUSION

In this paper, we have proposed a novel DL model called TEA, which can use the historical stock prices from five calendar days in combination with social media tweet representations to predict stock movements by means of a transformer encoder and attention mechanisms. To solve the problems of temporal dependence in financial data and insufficient effectiveness in fusing information from tweets and stock prices, an architecture consisting of a *feature extractor* and a *concatenation processor* is adopted. For the *feature extractor*, the structure consists of a transformer encoder, attention mechanisms, and a normalization strategy to effectively extract features and learn key information through relevant reviews of tweets and preprocessing of stock prices. The *concatenation processor* then processes the fused features to capture the temporal dependence. The overall framework realizes effective processing and analysis of tweets and stock prices to improve the *Accuracy* of stock movement prediction.

REFERENCES

1. E. F. Fama, L. Fisher, M. C. Jensen, and R. Roll, "The adjustment of stock prices to new information," *International Economic Review*, vol. 10, pp. 34–105, 1969.
   View at: Publisher Site | Google Scholar
2. J. Liu, X. Liu, H. Lin et al., "Transformer-based capsule network for stock movements prediction," in *Proceedings of the First Workshop on Financial Technology and Natural Language Processing*, pp. 66–73, Macao, China, August 2019.
   View at: Google Scholar
3. D. Araci, "FinBERT," *Financial Sentiment Analysis with Pre-trained Language Models*, 2019, https://arxiv.org/abs/1908.10063v1.
   View at: Google Scholar
4. Y. Xu and S. B. Cohen, "Stock movement prediction from tweets and historical prices," in *Proceedings of the ACL 2018 - 56th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference (Long Papers)*, pp. 1970–1979, Melbourne, Australia, July 2018.
   View at: Publisher Site | Google Scholar
5. H. Wu, W. Zhang, W. Shen, and J. Wang, "Hybrid deep sequential modeling for social text-driven stock prediction," in *Proceedings of the 2018 Association for Computing Machinery*, pp. 1627–1630, Torino, Italy, October 2018.
   View at: Publisher Site | Google Scholar
6. Q. Ding, S. Wu, H. Sun, J. Guo, and J. Guo, "Hierarchical multi-scale Gaussian transformer for stock movement prediction," in *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI-20)*, pp. 4640–4646, Christian, Bessiere, July 2020.
   View at: Publisher Site | Google Scholar
7. F. Feng, X. He, X. Wang, C. Luo, Y. Liu, and T. S. Chua, "Temporal relational ranking for stock prediction," *ACM T INFORM SYST*, vol. 37, 2019.
   View at: Publisher Site | Google Scholar
8. M. R. Vargas, C. E. M. Dos Anjos, G. L. G. Bichara, and A. G. Evsukoff, "Deep leaming for stock market prediction using technical indicators and financial news articles," in *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 971–978, New Orleans, Louisiana, USA, July 2018.
   View at: Publisher Site | Google Scholar
9. S. Sohangir and D. Wang, "Finding expert authors in financial forum using deep learning methods," in *Proceedings of the 2018 Second IEEE International Conference on Robotic Computing (IRC)*, pp. 399–402, Laguna Hills, CA, USA, February 2018.
   View at: Publisher Site | Google Scholar

10. X. Li, P. Wu, and W. Wang, "Incorporating stock prices and news sentiments for stock market prediction: a case of Hong Kong," *Information Processing & Management*, vol. 57, no. 5, Article ID 102212, 2020.
    View at: Publisher Site | Google Scholar
11. S. W. Lee and H. Y. Kim, "Stock market forecasting with super-high dimensional time-series data using ConvLSTM, trend sampling, and specialized data augmentation," *Expert Systems with Applications*, vol. 161, Article ID 113704, 2019.
    View at: Publisher Site | Google Scholar
12. Z. Hu, W. Liu, J. Bian, X. Liu, and T. Y. Liu, "Listening to chaotic whispers: a deep learning framework for news-oriented Stock trend prediction," in *Proceedings of the WSDM 2018 - Proceedings of the 11th ACM International Conference on Web Search and Data Mining*, pp. 261–269, Los Angeles, CA, USA, February 2018.

IJARASEM

INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

निस्केयर
NISCAIR

International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

| Mobile No: +91-9940572462 | Whatsapp: +91-9940572462 | ijarasem@gmail.com |