Journal of Computer Science and Technology Studies
ISSN: 2709-104X
DOI: 10.32996/jcsts
Journal Homepage: www.al-kindipublisher.com/index.php/jcsts

JCSTS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

The Influence of Social Media on Stock Market: A Transformer-Based Stock Price Forecasting with External Factors

Mst Zannatun Ferdus[1], Nishat Anjum[2]✉ Tuan Ngoc Nguyen[3], Afjal Hossain Jisan[4] and Md Azad Hossain Raju[5]
[1]Department of Science in Information Technology (MSIT), Washington University of Science and Technology Alexandria, USA
[2]Department of Computer Science, University of South Dakota, Vermillion, South Dakota, USA
[3]VNDirect Securities, 97 Lo Duc, Hai Ba Trung, Hanoi, Vietnam
[4]Department of Supply Chain & Information Systems, The Pennsylvania State University, University Park, Pennsylvania, USA
[5]Department of Computer Science, University of South Dakota, Vermillion, South Dakota, USA
Corresponding Author: Nishat Anjum, **E-mail**: nishat.anjum@coyotes.usd.edu

| ABSTRACT
This paper introduces an innovative approach to forecasting stock prices. Forecasting stock prices is crucial in assisting investors in making informed decisions. Our research presents a unique method that utilizes transformer-based machine learning approach for stock price forecasting. This method exploits self-attention mechanisms to grasp intricate patterns and dynamics within historical stock price data. To bolster our model performance, we integrate investors' sentiment collected from social media by using sentiment analysis with the help of natural language processing. Utilizing the variation caused by investors' sentiment over time, as well as external macroeconomic factors, our proposed model outperforms benchmark models. Through extensive comparisons with various benchmark machine-learning algorithms, results produced by our proposed method are favorably comparable to those produced by conventional approaches. Across multiple machine learning models, our preferred model demonstrates superior performance, achieving an RMSE value of 0.96 compared to the RMSE value of 1.58 obtained from LSTM model.

| KEYWORDS

Stock Price Forecasting, Transformer Models, Sentiment Analysis, Investment, Social Media

| ARTICLE INFORMATION

ACCEPTED: 02 February 2024 **PUBLISHED:** 13 February 2024 **DOI:** 10.32996/jcsts.2024.6.1.20

1. Introduction

Navigating stock prices is a challenging endeavor, yet it can be highly rewarding when forecasts are accurate. The scholarly community has focused its research on this domain, characterized by its complexity arising from the substantial volatility in the movement of stock prices. Stock prices are influenced by various technical indicators and information sources, making prediction challenging given the vast amount of available data. Nevertheless, technological advancements, especially in processing extensive temporal data, contribute to ongoing improvements in achieving higher prediction accuracy within this field. Yet, advanced approaches such as RNNs face challenges in capturing long-term dependencies in stock price data because of the gradient vanishing and exploding.

Various methods have been developed to forecast future stock prices, encompassing statistical techniques and machine learning methods. Statistical methods, such as auto-regression (AR), autoregressive moving average (ARMA), and autoregressive integrated moving average (ARIMA), leverage ground truth data to learn trends and patterns. Machine Learning and Deep learning approaches, including convolutional and recurrent neural networks, have also been applied to model stock price data. However, these models face challenges like "gradient vanishing and exploding" in RNNs and limitations of convolutional filters, affecting their ability to capture long-term and complex relations in sequence data. In this study, we introduce a novel time series forecasting approach based on the Transformer architecture (Vaswani et al., 2017). Unlike conventional sequencial models, Transformers process the entire sequence simultaneously, using self-attention mechanisms to learn dependencies. This makes Transformer-based models adept at capturing complex dynamics in time series data. In order to train the proposed Transformer model, we combine stock price data from Yahoo Finance with investors' sentiment obtained from social media and process it with Natural Language Processing (NLP). Within the realm of finance, the Efficient Market Hypothesis posits that asset prices aren't solely dependent on past information but rather react to new information, such as financial news articles and social media blogs (Titan, 2015). These sources possess the ability to impact the sentiments of investors and traders. We further enrich our dataset by amalgamating indicators for the macroeconomic environment as well as future anticipated macroeconomic policy variables. Recent literature on finance provides empirical evidence that investors' behavior and, therefore, stock prices can be affected by current macroeconomic conditions as well as anticipated future macroeconomic policies. Future macroeconomic policies can often be anticipated ahead of time as central banks engage in discussions about their intention, and such macroeconomic policies can affect the investment market. Investment market agents, therefore, try to take advantage of that by adjusting their current behavior based on anticipated future macroeconomic policies. In particular, we incorporate future interest rates as a representation of anticipated future interest rates alongside other macroeconomic indicators such as the Consumer Price Index (CPI), unemployment rates, and prevailing interest rates. All these factors have the potential to exert a substantial influence on the stock market. Our contribution to the literature includes the following:

1. We propose a transformer model that resolves gradient vanishing and exploding and utilizes self-attention mechanisms to learn long-term dependencies.
2. We enrich our data by amalgamating historical stock price data with investors' sentiment obtained from social media and processed by natural language processing approaches. We further enrich our data by combining macroeconomic indicators and anticipated future macroeconomic policy changes.

Relying on the advanced transformer model, we improve the accuracy of stock price predictions by amalgamating information from social media, which captures investors' sentiments, along with historical stock price related data and external current and expected future macroeconomic indicators. To our knowledge, our study first utilizes such a comprehensive approach to outperform conventional approaches. Our proposed model outperforms the benchmark model with an RMSE value of 0.96 compared to the LSTM model with an RMSE value of 1.58.

2. Literature Review

Whereas early literature mostly relied on statistical models, such as ARIMA, SARIMA, etc., recent literature has grown substantially, exploring the application of advanced algorithms, such as the RNN family, for stock price forecasting, with LSTM being the most widely used algorithm. With the help of a combination of short-term and long-term cells, LSTM can capture dependencies over longer periods. Given the intricacies and volatilities of stock data, researchers have also considered applying hybrid models incorporating LSTM layers for stock prediction. Another version of the algorithm under the umbrella of Deep Neural Networks, GRU, involves fewer parameters, making it computationally faster and more practical, especially for large stock datasets. Attention, a mechanism addressing the relationship between current and previous information, shows great potential in capturing sequential interactions. The remainder of this section will discuss various applications of such algorithms present in the current literature for forecasting stock prices.

LSTM has drawn much attention in the current literature for solving stock price predictions [Srijiranon et al. [2017] and Nguyen and Yoon [Srijiranon, 2022]. In [Salloum, 2017], news and historical data undergo preprocessing using FinBERT [Nguyen, 2019] and Principal Component Analysis (PCA) [Araci, 2019], respectively. Srijiranon (2022) incorporates prior knowledge through transfer learning [Abdi, 2010], resulting in significant improvements in prediction performance. Selvin et al. [2009] integrate LSTM with Convolutional Neural Network (CNN) to forecast stock prices. The data first enters a CNN layer to capture locational features, followed by an LSTM layer. During experimentation, the model, trained on the Infosys dataset, demonstrates good performance when applied to TCS and Cipla. Kim and Won [2017] combine LSTM with various GARCH-type models to predict stock price index volatility. The input includes GARCH parameters, EGARCH parameters, and explanatory variables. After concatenation, the data undergoes a single LSTM layer and two dense layers to obtain the final prediction. In experiments, the proposed model surpasses GARCH models and LSTM under the KOSPI 200 dataset. Sunny et al. [2024] apply Bi-Directional LSTM (BI-LSTM) to forecast stock price trends. The model comprises a RELU activation layer, two hidden BI-LSTM layers, one dropout layer, and one dense layer. Adjusting hyperparameters leads to a lower RMSE compared to LSTM models, enabling stock traders to achieve financial profits with a sustainable strategy. Zhang et al. [2018] utilize an LSTM-based model for forecasting stock price trends with limited order book data, capturing instant bid and ask information every 3 seconds. The model, named DeepLOB, consists of six CNN layers, three parallel inception layers, and one LSTM layer. Experimental results demonstrate that DeepLOB outperforms time-series related models when applied to the London Stock Exchange.

Hossain et al. [2020] employ GRU-based models for stock prediction. Utilizing S&P 500 historical time series data, the proposed model demonstrates superior performance compared to former state-of-the-art methods. Shahi et al. [2019] utilize LSTM and GRU for stock market forecasting. The model involves three data transformation processes: preprocessing stock and news data, followed by concatenation into a time series sequence generator. In the NEPSE dataset experiment, it is revealed that news data aids in prediction, with GRU outperforming LSTM. Gao et al. [2018] propose a method that combines GRU with LASSO or PCA for price trend prediction. The model initially uses LASSO or PCA for dimension reduction, followed by an LSTM or GRU layer for the final prediction. The experiment, using Shanghai Composite Index data, shows that GRU with LASSO dimension reduction performs better. Wu et al. [2023] use GRU with a Tree Regularization technique (GRU-Tree) to forecast the price trend. In this method, a GRU-based neural network is converted into a decision tree. Results using a dataset of the SSE Composite Index in China indicate that GRU-Tree outperforms both the decision tree and GRU-based methods.

A multi-input LSTM and attention-based neural networks were first integrated by Li et al. [2020]. The proposed method generates output with significantly improved performance compared to several existing benchmark methods. Empirical evidence from a transformer-based model was offered by Muhammad et al. [2021]. The model was trained on Dhaka Stock Exchange (DSE) stock price data and was used for forecasting daily, weekly, and monthly stock prices, yielding promising results. Other studies that considered transformer-based models include Liu et al. [2023], Zhang and Zohren [2019] and Sridhar and Sanagavarapu [2016]. We contribute to the literature by using the original transformer model comprised of encoder and decoders, proposed by Vaswani et al. (2017), trained on historical stock price data from Yahoo Finance combined with other external factors (sentiment index and macroeconomic factors). Our proposed model outperforms existing benchmark models.

3. Methods

3.1 Transformer Model

The original Transformer model proposed by Vaswani et al. in 2017 is a sequence-to-sequence model and consists of an encoder-decoder configuration. It takes a sequence from the input and generates the output sequence. Despite potential disparities in sequence length between the source and target sequences, the model encodes the source sequence into a fixed-length representation for subsequent decoding in an auto-regressive manner. This auto-regressive property imposes the need for information to propagate back to the sequence's beginning, a constraint applicable to time-series analysis as well. However, the auto-regressive nature of machine learning models during training can sometimes lead to the memorization of past observations rather than generalizing to new data.

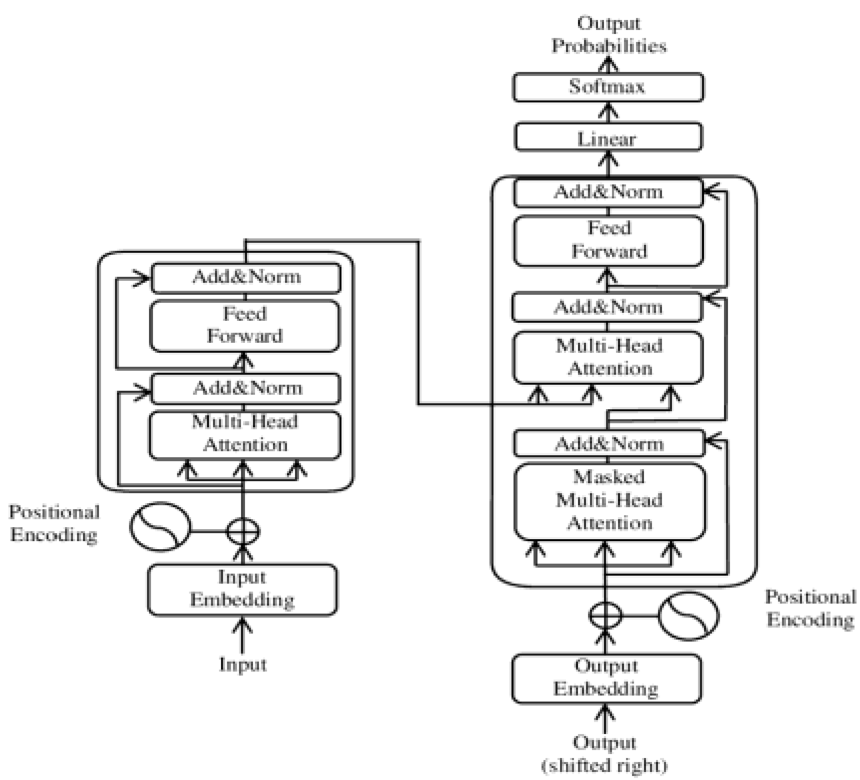

Figure 1: Basic Transformer Architecture

To address these challenges, Transformers employ self-attention and positional encoding techniques, enabling them to jointly attend to and encode ordered information as they analyze current data samples in a sequence. Unlike traditional recurrence-based models, Transformers preserve sequential information while eliminating the memorization problem. These techniques also leverage parallelism, thereby ensuring the maximum utilization of compute resources. Recent research endeavors have explored incorporating recurrent components into Transformers.

In the classical model, each embedded word is processed sequentially, relying on the output from the previous cell. In contrast, the Transformer processes the entire input sequence simultaneously, eliminating the need for sequential data processing. The sequence order is maintained through positional encoding. In the context of the self-attention operation, the Transformer architecture relies on finding associations between input segments by employing the dot product after incorporating positional information.

3.2 Dataset

The data utilized in this research has been gathered from diverse sources. We obtained historical daily stock price information for S&P500 ticker symbols from publicly accessible Yahoo Finance. Out of the 500 ticker symbols, we selected 10 symbols that have the strongest partial autocorrelation for any lags, ensuring that identified ticker symbols demonstrate robust serial autocorrelation, a crucial factor in constructing effective time-series models. Data on interest rates were sourced from Fred's data repository, while unemployment rates and Consumer Price Index (CPI) data were extracted from the World Bank's World Development Indicators (WDI) database. The research also incorporates the Tweets dataset from Kaggle. We applied various text cleaning methods that included converting all text to lowercase and removing numeric characters, Tweeter mentions, URLs, some special characters, and other texts that may not carry any relevant information. This tweet data is then classified using natural language processing to represent investors' sentiments reflected in the tweets. These distinct datasets were then consolidated into a unified dataset by appending them together. The analysis is limited to the duration of December 2022 to March 2023. The fusion of these datasets creates a comprehensive empirical corpus, allowing a correlation between daily Twitter sentiment and daily stock price movements.

4. Results and Discussions

Table 1: Performance matrices of various models

	RMSE		MAE	
	With Sentiment	Without Sentiment	With Sentiment	Without Sentiment
Transformer Model	0.96	0.97	0.99	1.27
LSTM	1.33	1.58	1.44	1.45
Random Forests	1.59	1.61	1.77	1.78
Light GBM	1.89	1.81	3.15	3.01

In Table 1, we present results obtained from the Transformer model and compare its performance with those of other benchmark models, such as Random Forests and Light GBM. Data is split into 80% training set and 20% test set, and then the Transformer model along with LSTM, Random Forests and Light GBM models were trained and tested on two datasets: one with sentiment and one without that. We perform various hyperparameter tunings and determine hyperparameter values for optimized performance. We assess various models' performances using Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). Results from the transformer model demonstrate superior performance across both measures and datasets. RMSE value obtained from the transformer model trained using investors' sentiment is 0.96, whereas that from the model trained without sentiment is 0.97. Similarly, the transformer model demonstrates superior performance in terms of MAE as well. The transformer model trained on data that includes sentiments demonstrates superior performance in terms of MAE as well, with an MAE of 0.99 compared to that obtained from the model trained on data without sentiments. Furthermore, results from all four models trained on the dataset containing investors' sentiment exhibit superior performance compared to those trained without that, as indicated by both RMSE and MAE. Performances across several models demonstrate the supremacy of the transformer model. Stock prices are often characterized by long term intricate signals that can affect future stock prices. Conventional models such as GRUs/LSTMs suffer weakness as these models fail to capture long term dependencies because of vanishing gradients and gradient explosion. Transformer model architecture overcomes this weakness by introducing a self-attention mechanism. Transformer models are thus the most appropriate for explaining future stock prices. Results from empirical experiments also support this assertion.

5. Conclusion

This study introduces an innovative approach to forecast stock prices utilizing the transformer model, integrating investors' sentiment reflected in social media along with macroeconomic indicators and expected shifts in macroeconomic policies into the models. We empirically establish the supremacy of the transformer model in predicting stock prices. Subsequently, we compare model performance on two datasets: one with investor sentiment and other external variables and one without. Model evaluation, using RMSE and MAE, consistently favors datasets incorporating investor sentiment across machine learning models. These results provide compelling evidence that our proposed method outperforms existing approaches, holding significant promise for academic research and practical industry applications. This approach could potentially disrupt the investment market, offering enhanced return on investment. Investors can leverage our method for more informed and advantageous investment decisions. Further investigations could explore the sensitivity of predictive accuracy in relation to more stock specific sentiment in social media.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Araci, D. (2019). Finbert: Financial sentiment analysis with pre-trained language models. arXiv 2019, arXiv:1908.10063.
[2] Ahmed, A. H., Ahmad, S., Sayed, M. A., Ayon, E. H., Mia, T., & Koli, T. (2023). Predicting the Possibility of Student Admission into Graduate Admission by Regression Model: A Statistical Analysis. *Journal of Mathematics and Statistics Studies*, *4*(4), 97-105.
[3] Abdi, H. (2010). Williams, L.J. Principal component analysis. Wiley Interdiscip. Rev. Comput. Stat. 2010, 2, 433–459. [CrossRef]
[4] Amin, M. S., Ayon, E. H., Ghosh, B. P., MD, M. S. C., Bhuiyan, M. S., Jewel, R. M., & Linkon, A. A. (2024). Harmonizing Macro-Financial Factors and Twitter Sentiment Analysis in Forecasting Stock Market Trends. *Journal of Computer Science and Technology Studies*, *6*(1), 58-67.
[5] Bhuiyan, M. S., Chowdhury, I. K., Haider, M., Jisan, A. H., Jewel, R. M., Shahid, R., & Ferdus, M. Z. (2024). Advancements in Early Detection of Lung Cancer in Public Health: A Comprehensive Study Utilizing Machine Learning Algorithms and Predictive Models. *Journal of Computer Science and Technology Studies*, *6*(1), 113-121.
[6] Bristi, U., Rajib, M. I., & Shoukat, S. (2020). Implementation of corporate social responsibility in the ready-made garment industry of Bangladesh in order to achieve sustainable development goals: a study on Bangladeshi garment manufacturers. *Journal of Textile Science and Technology*, *6*(4), 232-246.
[7] Gao, Y., Wang, R. and Zhou, E. (2021). Stock Prediction Based on Optimized LSTM and GRU Models. Sci. Program. 2021, 2021, 1–8. [CrossRef]
[8] Ghosh, B. P., Bhuiyan, M. S., Das, D., Nguyen, T. N., Jewel, R. M., Mia, M. T., & Cao, D. M. (2024). Deep Learning in Stock Market Forecasting: Comparative Analysis of Neural Network Architectures Across NSE and NYSE. *Journal of Computer Science and Technology Studies*, *6*(1), 68-75.
[9] Hossain, M.A., Karim, R., Thulasiram, R., Bruce, N.D and Wang, Y. (2018). Hybrid deep learning model for stock price prediction. In Proceedings of the 2018 IEEE Symposium Series on Computational Intelligence (ssci), Bengaluru, India, 18–21 November. 1837–1844.
[10] Hridoy, S M., Abu S., Rahman, S., Rishad, S M., Shadul I and Bhuiyan, Mohammad S. A Comprehensive Framework for Evaluating Software Engineering Technologies (December 1, 2023). Available at SSRN: https://ssrn.com/abstract=4650826 or http://dx.doi.org/10.2139/ssrn.4650826
[11] Islam, M. T., Ayon, E. H., Ghosh, B. P., MD, S. C., Shahid, R., Rahman, S., ... & Nguyen, T. N. (2024). Revolutionizing Retail: A Hybrid Machine Learning Approach for Precision Demand Forecasting and Strategic Decision-Making in Global Commerce. *Journal of Computer Science and Technology Studies*, *6*(1), 33-39.
[12] Kim, H.Y and Won, C.H. (2018). Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. Expert Syst. Appl. 2018, 103, 25–37. [CrossRef]
[13] Khan, R. H., Miah, J., Rahman, M. M., Hasan, M. M., & Mamun, M. (2023, June). A study of forecasting stocks price by using deep Reinforcement Learning. In *2023 IEEE World AI IoT Congress (AIIoT)* (pp. 0250-0255). IEEE.
[14] Li, H., Shen, Y., Zhu, Y. (2018). Stock price prediction using attention-based multi-input LSTM. In Proceedings of the Asian Conference on Machine Learning, PMLR, Beijing, China, 14–16 November. 454–469.
[15] Liu, J., Lin, H., Liu, X., Xu, B., Ren, Y., Diao, Y and Yang, L. (2019). Transformer-based capsule network for stock movement prediction. In Proceedings of the FirstWorkshop on Financial Technology and Natural Language Processing, Macao, China, 12 August 2019. 66–73.
[16] Muhammad, T., Aftab, A.B., Ahsan, M., Muhu, M.M., Ibrahim, M., Khan, S.I., Alam, M.S. (2022). Transformer-Based Deep Learning Model for Stock Price Prediction: A Case Study on Bangladesh Stock Market. arXiv 2022, arXiv:2208.08300.
[17] Nguyen, T.T and Yoon, S. (2019). A Novel Approach to Short-Term Stock Price Movement Prediction Using Transfer Learning. Appl. Sci, 9, 4745. [CrossRef]
[18] Pan, S.J and Yang, Q. (2009). A survey on transfer learning. IEEE Trans. Knowl. Data Eng. 2009, 22, 1345–1359. [CrossRef]
[19] Rahat, M. A. R., Islam, M. T., Cao, D. M., Tayaba, M., Ghosh, B. P., Ayon, E. H., ... & Bhuiyan, M. S. (2024). Comparing Machine Learning Techniques for Detecting Chronic Kidney Disease in Early Stage. *Journal of Computer Science and Technology Studies*, *6*(1), 20-32.
[20] Rana, M. S., Hossain, M. M., Jewel, R. M., & Islam, M. R. (2017). Evaluating Customers Satisfaction of Electronic Banking: An Empirical Study in Bangladesh. *The SIJ Transactions on Industrial, Financial & Business Management*, *5*(03), 07-12.
[21] Salloum, S.A., Al-Emran, M., Monem, A.A and Shaalan, K. (2017). A Survey of Text Mining in Social Media: Facebook and Twitter Perspectives. *Adv. Sci. Technol. Eng. Syst. J*, *2*, 127–133.
[22] Srijiranon, K., Lertratanakham, Y., Tanantong, T. (2022). A Hybrid Framework Using PCA, EMD and LSTM Methods for Stock Market Price Prediction with Sentiment Analysis. Appl. Sci. 2022, 12, 10823. [CrossRef]

[23] Selvin, S., Vinayakumar, R., Gopalakrishnan, E., Menon, V.K and Soman, K. (2023). Stock price prediction using LSTM, RNN and CNNsliding window model. In Proceedings of the 2017 International Conference on Advances in Computing, Communications and Informatics (icacci), Manipal, Karnataka, India, 13–16 September. 1643–1647. Appl. Sci. 2023, 13, 222 18 of 18

[24] Sunny, M.A.I., Maswood, M.M.S and Alharbi, A.G. (2020). Deep learning-based stock price prediction using LSTM and bi-directional LSTM model. In Proceedings of the 2020 2nd Novel Intelligent and Leading Emerging Sciences Conference (NILES), Giza, Egypt, 24–26 October. 87–92.

[25] Sarkar, M., Ayon, E. H., Mia, M. T., Ray, R. K., Chowdhury, M. S., Ghosh, B. P., ... & Puja, A. R. (2023). Optimizing E-Commerce Profits: A Comprehensive Machine Learning Framework for Dynamic Pricing and Predicting Online Purchases. *Journal of Computer Science and Technology Studies*, *5*(4), 186-193.

[26] Shahi, T.B., Shrestha, A., Neupane, A. and Guo, W. (2020). Stock price forecasting with deep learning: A comparative study. Mathematics 2020, 8, 1441. [CrossRef]

[27] Shoukat, S. (2016). Now and then: The neckline history of women. *American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS)*, *26*(2), 33-52.

[28] Sridhar, S and Sanagavarapu, S. (2021). Multi-head self-attention transformer for dogecoin price prediction. In Proceedings of the 2021 14th International Conference on Human System Interaction (HSI), Gdansk, Poland, 8–10. 1–6.

[29] Shoukat, S., & Kwon, W. S. (2024, January). Effects of AI Agent Anthropomorphism on Consumers' Affective, Cognitive, and Social Shopping Experiences. In *International Textile and Apparel Association Annual Conference Proceedings* (Vol. 80, No. 1). Iowa State University Digital Press.

[30] Tayaba, M., Ayon, E. H., Mia, M. T., Sarkar, M., Ray, R. K., Chowdhury, M. S., ... & Puja, A. R. (2023). Transforming Customer Experience in the Airline Industry: A Comprehensive Analysis of Twitter Sentiments Using Machine Learning and Association Rule Mining. *Journal of Computer Science and Technology Studies*, *5*(4), 194-202.

[31] Tayaba, M., Ayon, E. H., Mia, M. T., Sarkar, M., Ray, R. K., Chowdhury, M. S., ... & Puja, A. R. (2023). Transforming Customer Experience in the Airline Industry: A Comprehensive Analysis of Twitter Sentiments Using Machine Learning and Association Rule Mining. *Journal of Computer Science and Technology Studies*, *5*(4), 194-202.

[32] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł and Polosukhin, I. (2017). Attention is all you need. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 30.

[33] Wu,W., Wang, Y., Fu, J., Yan, J., Wang, Z., Liu, J., Wang,W., Wang, X. (2019). Preliminary study on interpreting stock price forecasting based on tree regularization of GRU. In Proceedings of the International Conference of Pioneering Computer Scientists, Engineers and Educators, Guilin, China, 20–23. 476–487.

[34] Zhang, Z., Zohren, S., Roberts, S. (2019). Deeplob: Deep convolutional neural networks for limit order books. IEEE Trans. Signal Process. 2019, 67, 3001–3012. [CrossRef]

[35] Zhang, Z and Zohren, S. (2021). Multi-horizon forecasting for limit order books: Novel deep learning approaches and hardware acceleration using intelligent processing units. arXiv 2021, arXiv:2105.10430.