ALERTA-Net: A Temporal Distance-Aware Recurrent Networks for Stock Movement and Volatility Prediction

Shengkun Wang[1], YangXiao Bai[2], Kaiqun Fu[2], Linhan Wang[1], Chang-Tien Lu[1], Taoran Ji[3]

[1]Department of Computer Science, Virginia Tech
[2]Department of Computer Science, South Dakota State University
[3]Department of Computer Science, Texas A&M University - Corpus Christi

shengkun@vt.edu, bai.yangxiao@sdstate.edu, kaiqun.fu@sdstate.edu,
linhan@vt.edu, ctlu@vt.edu, taoran.ji@tamucc.edu

Abstract—For both investors and policymakers, forecasting the stock market is essential as it serves as an indicator of economic well-being. To this end, we harness the power of social media data, a rich source of public sentiment, to enhance the accuracy of stock market predictions. Diverging from conventional methods, we pioneer an approach that integrates sentiment analysis, macroeconomic indicators, search engine data, and historical prices within a multi-attention deep learning model, masterfully decoding the complex patterns inherent in the data. We showcase the state-of-the-art performance of our proposed model using a dataset, specifically curated by us, for predicting stock market movements and volatility.

Index Terms—stock market prediction, twitter, google trends, sentiment analysis, macroeconomic data

I. INTRODUCTION

Significantly influencing other business sectors [1], the stock market serves as a vital mechanism and is crucial for companies to raise capital. With U.S. stock holdings expected to hit $40 trillion in 2023, equating to 1.5 times the nation's GDP, it stands as a major portion of the entire economy, highlighting the stock market's pivotal position as a benchmark for the U.S. economic landscape. Our research centers on blue-chip stocks[1], which mirror the broader dynamics of the stock market.

We've selected 41 blue-chip stocks from 10 Global Industry Classification Standard (GICS)[2] Sectors for our financial market study. Each of these stocks is considered investment-

---

[1]Blue chip stocks are shares issued by financially robust, well-established companies with stellar reputations.

[2]GICS classifies companies into specific economic sectors and industry groups that most accurately represent their business operations.

worthy[3] by both Moody's and S&P. Recognizing the intrinsic challenge in accurately predicting stock prices as highlighted by Nguyen et al. [2], we use blue-chip stocks in our research to anticipate upcoming stock price movements and volatility trends, as indicated by Feng et al. [3] and Xu et al. [4].

In the domain of stock market research, two primary methodologies prevail: technical analysis and fundamental analysis. Technical analysis utilizes past stock prices to predict future trends [5]. However, its heavy dependence on historical data can sometimes overlook sudden market changes due to unexpected events. Assuming a uniformly rational market behavior, this methodology can inadvertently create an echo chamber. This effect can cause trading signals to amplify themselves, eventually becoming disconnected from the actual economic context. Conversely, fundamental analysis integrates both price features and external information, including data from social media [6] and search engines [7]. Mao et al. [8] demonstrated an enhanced accuracy in forecasting the S&P 500 closing price when integrating Twitter[4] data into their model. While these data sources frequently reflect not only the financial market but also vital economic indicators, the prevailing research in fundamental analysis tends to emphasize the financial market, neglecting the symbiotic relationship between the broader economy and the stock market. Moreover, while existing models mainly center on forecasting trend shifts [9], they often neglect the importance of the scale of these changes. In the realm of stock behavior, the magnitude of these shifts holds significant weight.

In this paper, we propose ALERTA-Net: **A**ttentional Tempora**L** Distanc**E** Awa**R**e Recurren**T** Neur**A**l Networks. To our best knowledge, it is the first paper to use the combination of social media, macroeconomic data and search engine information to predict both stock price movement and volatility. Our contributions can be summarized as follows:

---

[3]Companies rated Baa or higher by Moody's and Standard & Poor's (S&P) are considered to be of high quality and deemed investment-worthy.

[4]Despite the recent rebranding of Twitter to "X", this article retains the use of its original name, "Twitter".

• **Proposing a framework enabling the fusion of social media, macroeconomic factors and search engine data for stock movements and volatility.** By integrating above information with our method, we can not only predict stock price movements but also efficiently extract information from stock market volatility. This allows us to provide advance warning of any unusual fluctuations in the stock market in the future.

• **Formulating a temporal distance-aware, multi-attention mechanisms on multi-view market data.** The proposed ALERTA-Net shines in recognizing the dynamic, temporal distance-based relationships inherent within difference hidden states. Capitalizing the same day stock price movements, the model greatly amplifies its precision in predicting stock market volatility.

• **Validating the effectiveness and efficiency of the proposed model via experiments and comparisons.** We conduct experiments on one real-world dataset[5]. Both conventional methods and deep learning based methods for stock market movements and volatility are selected for comparisons. Evaluations of various metrics are presented, illustrating the effectiveness of our proposed model.

II. RELATED WORK

Predictive methods for stock movement can be broadly categorized into two main types: technical analysis and fundamental analysis. While technical analysis relies exclusively on past price data to anticipate future trends, fundamental analysis adopts a more comprehensive approach, considering not only historical prices but also information from textual sources, economic indicators, financial metrics, and a myriad of both qualitative and quantitative aspects.

Social Media: Numerous studies have explored stock market predictions using social media data. Contemporary research combines sentiment analysis with historical price data, extracting insights from platforms such as Yahoo's message board [10], blogs [11], Twitter [4], and Reddit [12]. This integration has revealed correlations with stock market trends.

News and Search Engine: Exploring the influence of public news and user browsing habits, research has probed into how traders respond to news events. For instance, Xiong et al. [13] regard Google trends and market data as catalysts for daily S&P 500 variations. Other research, works like Bordino et al. [14] draw connections between search query volume and stock activity. More sophisticated techniques, like the hierarchical attention mechanisms introduced by Hu et al. [15], extract news sequences directly from textual content to predict stock trends.

Macroeconomic Indicators: Numerous studies pinpoint various economic elements influencing stock returns. Notably, Ferson et al. [16] highlight the central role of interest rates in dictating stock returns. In addition, indicators like the relative T-Bill rate and the consumption-wealth ratio are underscored as crucial predictors by Jank et al. [17]. Beyond these,
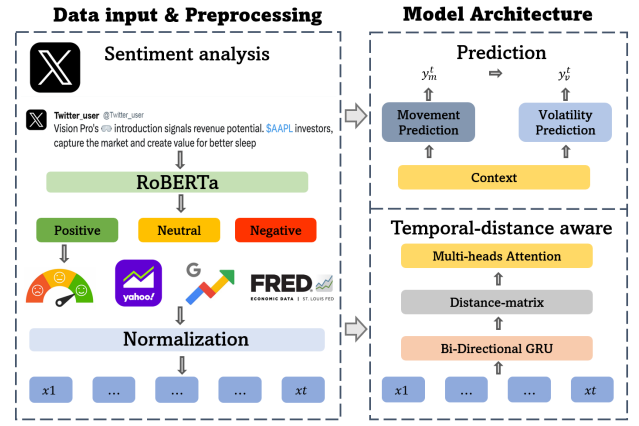

Figure 1: The architecture of ALERTA-Net is designed to predict the movement y_m^t, and volatility y_v^t on day t. In the data input and preprocessing phase, we extract textual information from Twitter and convert it into sentiment scores; Then, ALERTA-Net utilizes these scores, along with other features, to make predictions, taking temporal distance into account.

economic markers like unemployment rates, inflation, and commodity prices also exert significant influence on stock returns, as affirmed by sources like [18] and [19].

III. PROBLEM SETUP

Let $p = \{p^1, \ldots, p^t\}$ denote the stock's daily adjusted closing price. We formulate the actual labels movement y_m and volatility y_v as follows:

$$y_m^t = \mathbb{1}(p^t - p^{t-1}), \tag{1}$$

$$y_v^t = \mathbb{1}(p^t - p^{t-1})/p^{t-1}. \tag{2}$$

The actual labels can be represented as $y_m = \{y_m^1, \ldots, y_m^t\} \in \mathbb{R}^T$ and $y_v = \{y_v^1, \ldots, y_v^t\} \in \mathbb{R}^T$. Let $X^T = \{x^1, \ldots, x^t\} \in \mathbb{R}^{D \times T}$ represents the sequential input features (e.g., sentiment scores, stock adjusted prices) from the previous T time-steps, where D signifies the dimension of the features. Since our goal is to utilize the sequence of features X^T to predict the next time-step movements $\hat{y}_m$ and volatility $\hat{y}_v$ of blue-chip stocks. We can define our prediction functions $\hat{y}_m^t = f(X^T; \Theta_1), \hat{y}_v^t = g(X^T, \hat{y}_m^t; \Theta_2)$, where f the function with parameters Θ_1 aims to predict the movement of stock s at the next time-step from the sequential features X^T and g aims to predict the unusual fluctuation from both X^T and $\hat{y}_m^t$.

In practical settings, by varying the time lag in extensive historical stock data, we can often produce numerous training examples. However, for clarity in presenting our proposed technique, we zero in on a distinct time lag for predicting both movement and volatility. Additionally, our predictive models have assimilated the concept of adjusted stock prices [20]. This ensures that our models discern authentic stock value changes driven by market factors, rather than getting influenced by artificial shifts arising from corporate decisions.

---

[5]Our dataset is available at https://github.com/hao1zhao/ALERTA-Net

IV. Framework Components

In this paper, we introduce a new framework, ALERTA-Net, which incorporates temporal distance-aware, multi-attention mechanisms for processing multi-view stock data. The overall architecture is illustrated in Figure 1. The data input & preprocessing layer transforms both temporal and textual information into dense vectors. Then, our temporal distance-aware layer have the recurrent representation identifies hidden dependencies within the current stock data, based on past information. After that, distance-matrix context integrates these historical dependencies within the sequence of features X^T. Lastly, the predictions layer generates time-aware forecasts for the stock movements and volatility in the next time-step, thereby providing a complete and cohesive system for stock prediction.

Data input & preprocessing. We designate the textual data extracted from Twitter as α. To quantify the embedded sentiment within the Twitter text, we utilize the roBERTa-base sentiment model in combination with TweetNLP [21], facilitating the generation of tweet sentiment scores denoted as $\hat{\alpha}$. Then, we concatenate $\hat{\alpha}$ with other relevant historical data, yielding $E^T = \{e^1, \dots, e^t\} \in \mathbb{R}^{D \times T}$, in which the feature dimension is 17. Considering the fact that normalization provides a uniform scale to all features, thereby precluding any specific feature from dominating, we apply log normalization to E^T. For the purpose of avoiding numerical instability, we add a small constant $\epsilon = 1e - 8$, the formula is as follows:

$$X^T = \log E^T + \epsilon. \tag{3}$$

Temporal distance aware (TDA). In this layer, we propose a temporal-distance aware layer to enhance the modulation of conditional dependencies by allowing the model to access and directly attend to previous hidden states. First, given the proficiency in handling long-term dependencies, recurrent neural network is extensively employed for sequential data processing [22]. The general idea of recurrent unit is to recurrently project the input sequence into a sequence of hidden representations. At each time-step, the recurrent unit learns the hidden representations h_t by jointly considering the input x_t and previous hidden representation h_{t-1} to capture sequential dependency. To capture the sequential dependencies and temporal patterns in the historical stock features, an GRU [23] recurrent unit is applied to map $\{x^1, \dots, x^t\}$ into hidden representations $\{h^1, \dots, h^t\} \in \mathbb{R}^{U \times T}$, with the dimension of U.

Instead of just using the immediate previous hidden state h^t to update the current hidden state h^{t+1}, we are now considering a weighted sum of all previous hidden states. This not only enable the model to place greater emphasis on the impact of recent events on the stock market but also to consider a longer history. We utilized a temporal distance to represent the interval between two time steps.

$$w^i = \frac{1}{t - i + 1}, \tag{4}$$

where the weight w^i for a hidden state at time i based on its temporal distance from the current time t. By adding 1 to the denominator, we ensure that even the most recent hidden state gets a weight, preventing division by zero.

Then, we apply weights to the hidden states of the GRU:

$$c^t = \text{GRU}\left(x^t, \sum_{i=1}^{t} w^i \cdot h^i\right), \tag{5}$$

each hidden state h^i is multiplied by its respective weight w^i, and the results are summed up. The result is a context state c^t that captures the weighted influence of all previous states.

Prediction Layer. Instead of directly predicting stock movement, denoted as y_m^t, and volatility y_v^t, we first concatenate the context c^t with the previous hidden state h^t in the fusion layer. We then use the cross-entropy function as the optimizer for stock price movement prediction. Following this, we concatenate the last hidden state h^t, context c^t, and the output y_m^t to predict volatility y_v^t. For this prediction, we utilize the binary cross-entropy with logits as the loss function.

V. Experiment

A. Dataset Description

Our dataset coalescest three main components: Historical price data, Twitter data and macroeconomics data.

Yahoo Finance. We sourced historical data from Yahoo Finance, which monitored the trajectory of 41 blue-chip stocks between June 1, 2020, and June 1, 2023. To hone our prediction objectives, we designated a threshold range spanning from -0.5% to 0.5% to filter out negligible shifts. While Baker et al. [24] suggest daily stock price alterations exceeding 2.5% are deemed notable, our model's aim is to forecast atypical volatility. Aligning with Ding et al.'s insights [25] on distinct stock fluctuation parameters, we've established a loftier standard, recognizing a 5% swing as an outlier. As a result, we categorize samples with variances below 5% as 0 and those at or above 5% as 1.

Twitter. During the same date range, we included approximately 7.8 million tweets, gathered via Twitter's official API at a sampling rate of 10%. We were particular in our selection of tweets: they needed to contain at least one cashtag and had to be posted within standard U.S. trading hours, from 9 am to 4:30 pm. We recognized the significant influence that Twitter volume has on stock trading, a fact underscored by Cazzoli et al. [26]. Thus, we ensured that our model's input parameters included the daily count of processed Twitter posts.

Google Trends & Federal Reserve Economic Data. We engaged in targeted searches on Google and Federal Reserve Economic Data, using carefully curated keywords originating from the "Outline of Economics" Wikipedia page. In light of the disparate update intervals across various indicators, we broke down each data pull into smaller segments, ensuring normalization of the data across these windows.

B. Baseline Methods and Evaluation Metrics

We evaluate our model's effectiveness by comparing it with DP-LSTM [27], a renowned stock movement prediction network by using financial data.

Other benchmarks employed in our study include Extreme Gradient Boosting [28], Attention-based LSTM [29], and GRU [23]. Following similar procedures in (Xu et. al [22]; Zhang et. al [9]), we report our results in terms of Accuracy (Acc.) and Matthews Correlation Coefficient (MCC). Given that data points involving stock price changes greater than 5% only constitute a minor portion of our dataset, we've also chosen to utilize the Area Under the ROC Curve (AUC) as our performance metric in order to achieve a more robust and realistic evaluation.

TABLE I : Classification performance of ALERTA-NET and baseline methods, measured with the accuracy (Acc.), the Matthews Correlation Coefficient (MCC) and Receiver Operating Characteristic (ROC). The best is in bold, our ALERTA-NET shows the best performance in all evaluation metrics.

Models	Movement		Volatility		
	Acc.	MCC	Acc.	MCC	AUC
XGBoost	0.5030	0.01461	0.5240	-0.02275	0.5561
GRU	0.5121	0.04217	0.5633	0.06241	0.5993
DP-LSTM	0.5036	0.00741	0.5684	0.05297	0.5890
AT-LSTM	0.5165	0.03779	0.5886	0.07028	0.5850
ALERTA-Net	**0.5238**	**0.05179**	**0.6136**	**0.07465**	**0.6197**

TABLE II : An ablation study of ALERTA-Net, where ALERTA-Net(P) is baseline without any tweet and macroeconomic data. ALERTA-Net(S) is only relies on sentiment data. ALERTA-Net(W/O M) is without macroeconomic data. Both tweet sentiment and macroeconomic data improve the performance of the framework.

Models	Movement		Volatility		
	Acc.	MCC	Acc.	MCC	AUC
ALERTA-Net(P)	0.4988	0.0016	0.5238	0.07168	0.6019
ALERTA-Net(S)	0.4952	0.0092	0.5150	0.03213	0.5753
ALERTA-Net(W/O M)	0.5114	0.0378	0.5247	0.03992	0.5989
ALERTA-Net	**0.5238**	**0.05179**	**0.6136**	**0.07465**	**0.6197**

C. Results

The performances of our proposed models and the established baselines are detailed in TABLE I. AT-LSTM is observed to be the superior baseline model in terms of accuracy and MCC for movement prediction, while GRU shows an outstanding performance in the AUC (Area Under the Curve) and MCC for volatility prediction. ALERTA-Net surpasses both of these models by significant margins. In terms of accuracy, ALERTA-Net achieves a score of 0.5238 and 0.6136, outperforming GRU and AT-LSTM by 1.4% and 4.2% respectively. Additionally, for MCC, ALERTA-Net outperforms GRU and AT-LSTM by 22.8% and 6.2% respectively, and outshines both in AUC by a margin of 3.4%. Overall,

these results reinforce the effectiveness of our proposed model ALERTA-Net.

D. Ablation Study

To conduct an in-depth analysis of the core components of ALERTA-Net, we have constructed three variations alongside the fully-loaded model. Each variant is specifically tailored to handle certain types of input data: ALERTA-Net(P) solely relies on closed price data, ALERTA-Net(S) exclusively processes Twitter-derived sentiment data, while ALERTA-Net(W/O M) incorporates both price and sentiment data, but omits macroeconomic information. As shown in TABLE II, our ablation study revealed that incorporating macroeconomic data significantly enhances the predictive capabilities of the model for stock movement and volatility to varying degrees.

VI. CONCLUSION

We introduced ALERTA-Net, a deep generative neural network architecture, to showcase the efficacy of combining search engine data, macro-economy data, and social media data when trying to predict stock movements and volatility. We tested our model on a new comprehensive dataset and showed it performs better than strong baselines. In future studies, we plan to enhance accuracy by integrating a variety of text and audio sources, including earnings calls and financial reports.

REFERENCES

[1] M. Billah, S. Waheed, and A. Hanifa, "Stock market prediction using an improved training algorithm of neural network," in 2016 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE). IEEE, 2016, pp. 1–4.

[2] T. H. Nguyen, K. Shirai, and J. Velcin, "Sentiment analysis on social media for stock movement prediction," Expert Systems with Applications, vol. 42, no. 24, pp. 9603–9611, 2015.

[3] F. Feng, X. He, X. Wang, C. Luo, Y. Liu, and T.-S. Chua, "Temporal relational ranking for stock prediction," ACM Transactions on Information Systems (TOIS), vol. 37, no. 2, pp. 1–30, 2019.

[4] Y. Xu and S. B. Cohen, "Stock movement prediction from tweets and historical prices," in Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Melbourne, Australia: Association for Computational Linguistics, Jul. 2018, pp. 1970–1979. [Online]. Available: https://aclanthology.org/P18-1183

[5] W. Lu, J. Li, J. Wang, and L. Qin, "A cnn-bilstm-am method for stock price prediction," Neural Computing and Applications, vol. 33, pp. 4741–4753, 2021.

[6] J. A. Cookson and M. Niessner, "Why don't we agree? evidence from a social network of investors," The Journal of Finance, vol. 75, no. 1, pp. 173–228, 2020.

[7] Z. Wang, Y. Huang, B. Cai, R. Ma, and Z. Wang, "Stock turnover prediction using search engine data," Journal of Circuits, Systems and Computers, vol. 30, no. 07, p. 2150122, 2021.

[8] Y. Mao, W. Wei, B. Wang, and B. Liu, "Correlating s&p 500 stocks with twitter data," in Proceedings of the First ACM International Workshop on Hot Topics on Interdisciplinary Social Networks Research, ser. HotSocial '12. New York, NY, USA: Association for Computing Machinery, 2012, p. 69–72. [Online]. Available: https://doi.org/10.1145/2392622.2392634

[9] Q. Zhang, C. Qin, Y. Zhang, F. Bao, C. Zhang, and P. Liu, "Transformer-based attention network for stock movement prediction," Expert Systems with Applications, vol. 202, p. 117239, 2022.

[10] S. R. Das and M. Y. Chen, "Yahoo! for amazon: Sentiment parsing from small talk on the web," For Amazon: Sentiment Parsing from Small Talk on the Web (August 5, 2001). EFA, 2001.

[11] M. De Choudhury, H. Sundaram, A. John, and D. D. Seligmann, "Can blog communication dynamics be correlated with stock market activity?" in Proceedings of the nineteenth ACM conference on Hypertext and hypermedia, 2008, pp. 55–60.

[12] S. Long, B. Lucey, Y. Xie, and L. Yarovaya, ""i just like the stock": The role of reddit sentiment in the gamestop share rally," Financial Review, vol. 58, no. 1, pp. 19–37, 2023.

[13] R. Xiong, E. P. Nichols, and Y. Shen, "Deep learning stock volatility with google domestic trends," arXiv preprint arXiv:1512.04916, 2015.

[14] I. Bordino, S. Battiston, G. Caldarelli, M. Cristelli, A. Ukkonen, and I. Weber, "Web search queries can predict stock market volumes," PloS one, vol. 7, no. 7, p. e40014, 2012.

[15] 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 eleventh ACM international conference on web search and data mining, 2018, pp. 261–269.

[16] W. E. Ferson and C. R. Harvey, "The variation of economic risk premiums," Journal of political economy, vol. 99, no. 2, pp. 385–415, 1991.

[17] S. Jank, "Mutual fund flows, expected returns, and the real economy," Journal of Banking & Finance, vol. 36, no. 11, pp. 3060–3070, 2012.

[18] J. D. Hamilton, "Causes and consequences of the oil shock of 2007-08," National Bureau of Economic Research, Tech. Rep., 2009.

[19] P. J. Kehoe, P. Lopez, V. Midrigan, and E. Pastorino, "Asset prices and unemployment fluctuations," National Bureau of Economic Research, Tech. Rep., 2019.

[20] B. Xie, R. Passonneau, L. Wu, and G. G. Creamer, "Semantic frames to predict stock price movement," in Proceedings of the 51st annual meeting of the association for computational linguistics, 2013, pp. 873–883.

[21] J. Camacho-collados, K. Rezaee, T. Riahi, A. Ushio, D. Loureiro, D. Antypas, J. Boisson, L. Espinosa Anke, F. Liu, E. Martínez Cámara et al., "TweetNLP: Cutting-edge natural language processing for social media," in Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing: System Demonstrations. Abu Dhabi, UAE: Association for Computational Linguistics, Dec. 2022, pp. 38–49. [Online]. Available: https://aclanthology.org/2022.emnlp-demos.5

[22] Y. Xu and S. B. Cohen, "Stock movement prediction from tweets and historical prices," in Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2018, pp. 1970–1979.

[23] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," arXiv preprint arXiv:1406.1078, 2014.

[24] S. R. Baker, N. Bloom, S. J. Davis, and M. C. Sammon, "What triggers stock market jumps?" National Bureau of Economic Research, Tech. Rep., 2021.

[25] Y. Ding, H. Zhang, and J. Zhang, "Private vs state ownership and earnings management: Evidence from chinese listed companies," Corporate Governance: An International Review, vol. 15, no. 2, pp. 223–238, 2007.

[26] L. Cazzoli, R. Sharma, M. Treccani, and F. Lillo, "A large scale study to understand the relation between twitter and financial market," in 2016 third European network intelligence conference (ENIC). IEEE, 2016, pp. 98–105.

[27] X. Li, Y. Li, H. Yang, L. Yang, and X.-Y. Liu, "Dp-lstm: Differential privacy-inspired lstm for stock prediction using financial news," arXiv preprint arXiv:1912.10806, 2019.

[28] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, 2016, pp. 785–794.

[29] Y. Wang, M. Huang, X. Zhu, and L. Zhao, "Attention-based lstm for aspect-level sentiment classification," in Proceedings of the 2016 conference on empirical methods in natural language processing, 2016, pp. 606–615.