# Combining Graph Attention Networks and Modern Portfolio Theory for Enhanced Stock Portfolio Optimization

Junkyu Jang[a] , Daesan Oh[a] and Sung-Hyuk Park[a]*

[a] *Management Engineering Department, College of Business, Korea Advanced Institute of Science and Technology, Seoul, Korea*

E-mail: sunghyuk.park@kaist.ac.kr *corresponding author

**[Present/permanent address]**

[a]*Sung-Hyuk Park **(corresponding author)**

Management Engineering, KAIST College of Business

85, Heogi-ro, Cheongnyangri-dong, Dongdaemun-gu, Seoul, Korea

E-mail: sunghyuk.park@kaist.ac.kr

**Combining Graph Attention Networks and Modern Portfolio Theory for Enhanced Stock Portfolio Optimization**

**Abstract**

This paper compares the investment model that applies correlations learned from the Graph Attention Networks (GAT), an artificial intelligence that considers relationships between assets, with the correlations from traditional complex system-based optimization models, against the conventional Markowitz model and portfolios utilizing artificial neural networks. We backtest indicators such as Compound Annual Growth Rate (CAGR), Sharpe ratio, and beta, which can compare the returns and risks of portfolios, using daily data of assets belonging to the Nasdaq 100. We found that selecting asset compositions with GAT and applying the learned correlations as weights in the calculation yields better portfolio investment outcomes.

Keywords : Portfolio optimization , Portfolio Rebalancing, Investment , Graph Attention Network(GAT), Artificial Intelligence (AI) ,

## 1. Introduction

As the volume of data available in financial markets continues to expand, the complexity of analyzing and predicting stock market trends has significantly increased. This complexity necessitates delving into exploring advanced forecasting and portfolio optimization techniques (Nam and Seong, 2019). Artificial Intelligence (AI) has emerged as a pivotal tool in addressing these challenges, offering new avenues for informed decision-making and enhancing the efficiency of investment management and algorithmic trading (Song, et al., 2017).

Contemporary exploration in portfolio optimization has significantly evolved from the foundational principles of Modern Portfolio Theory (MPT) (Harlow, 1991, Jang and Park, 2016, Markowitz, 1952). Notably, AI methodologies, including Support Vector Machines (SVM), Long Short-Term Memory (LSTM) networks, Convolutional Neural Networks (CNN), and reinforcement learning, have been employed to refine asset selection processes and integrate these selections with MPT to yield optimized portfolio outcomes (Jang and Seong, 2023, Paiva, et al., 2019, Silva, et al., 2024, Zhou, et al., 2023). However, these methods of asset selection do not consider the relationships between assets. Despite these advancements, the consideration of inter-asset relationships, a cornerstone of MPT, often remains unaddressed (Fernández and Gómez, 2007). Acknowledging the significance of these correlations, this study proposes the utilization of the Graph Attention Network (GAT) (Veličković, et al., 2017) model to incorporate asset correlations into the portfolio selection process. GAT has been proven effective in learning the graph between nodes for various tasks (Brody, et al., 2021, Wang, et al., 2019). The primary objective of our paper is to effectively train the GAT model by appropriately utilizing each asset's technical indicators, thereby learning the relationships between assets. This will enable us to construct features for each asset and select stocks that are likely to perform well based on the given date. Following the selection, we aim to optimize these stocks based on methods derived from Modern Portfolio Theory and compare the results with those of traditional methods. Our research primarily focuses on combining the artificial intelligence-based GAT model, which considers the relationships between assets, with traditional optimization models.

We contribute to the recent literature in several ways: (i) The GAT model effectively learns the relationships between assets and uses this information to generate investment signals, resulting in superior backtesting performance compared to other AI-based methods. (ii) In conventional portfolio weight optimization methods, improvements in portfolio diversification were observed through the use of correlations by complex system-based methodologies (Seong and Nam, 2022). Building upon this approach, the current study extends and enhances traditional optimization methods by incorporating MPT based on complex systems with correlations learned through GAT.

(iii) We have confirmed that the portfolios constructed in this method outperform in terms of returns and metrics such as the Sharpe ratio when compared with the NASDAQ index and Markowitz portfolios.

## 2. Methodology

Data collection : This study gathered data on each asset included in the NASDAQ 100 index, spanning from September 17, 2012, to January 2, 2024, encompassing every trading day within this period. The dataset was procured utilizing Bloomberg software, and was subsequently employed for both training and evaluation purposes via Python. The dataset comprises daily metrics such as open, close, high, low prices, and trading volume for each asset, adjusted for stock splits and dividend payments. To ensure data consistency, we used 78 stocks that continuously remained in the NASDAQ 100 index throughout the collected data period. We divided the collected data into two parts: from September 17, 2012, to September 27, 2021, was designated for model training, while the second segment, from September 27, 2021, to January 2, 2024, was allocated for model testing.

Data preprocessing : In the realm of stock market analysis, technical indicators play an essential role in uncovering the dynamics, trends, and volatility of stock prices to provide information for future market predictions. This study utilizes open, high, low, close, and volume (OHLCV) data from 78 stocks to create technical indicators and use them as inputs to a graph attention network (GAT) model. The selection of technical indicators was based on a comprehensive review of prior research in the field (Barua and Sharma, 2022, Chen, et al., 2016, Henrique, et al., 2018, Ma and Yan, 2022, Peng, et al., 2021, Silva, et al., 2024). The technical indicators chosen as inputs for our GAT model, detailed in Table 1, were collected.

Table 1. Technical indicators used in GAT model.

| Input Type | Technical Indicators | Reference |
|---|---|---|
| Close | Moving Average, Exponential Moving Average, Rate of Change, Standard Deviation, Bollinger Bands, Triple Exponential Average | (Chen, et al., 2016, Gorenc Novak and Velušček, 2016, Ji, et al., 2022) |
| Open, High, Low, Close | Average True Range, Average Directional Movement Index, Commodity Channel Index, Relative Strength Index, WilliamsR, Slow Stochastic Oscillator, Vortex Indicator, Coppock Curve, Keltner Channel | (Barua and Sharma, 2022, Henrique, et al., 2018, Ma and Yan, 2022, Shynkevich, et al., 2017, Weng, et al., 2017) |
| Open, High, Low, Close, Volume | Money Flow Index, Force Index | (Peng, et al., 2021, Ramezanian, et al., 2019, Thawornwong, et al., 2003) |

Asset selection method through GAT : Once the data based on Technical Indicators is prepared, this study employs Graph Attention Networks (GAT) to encode it for predicting fluctuations in stock prices. GAT aims to understand the complex relationships between each stock by learning the interactions among nodes in a graph, thereby aiming to classify the stock price movements after a portfolio rebalancing period. GAT computes a new feature vector $h'_i$ for stock $i$ using its feature vector $h_i$ along with the set of feature vectors $\{h_j \mid j \in N\}$ of the entire set of assets $N$. During this process, the Attention mechanism is utilized to take a weighted average of the information from other stocks related or not to asset $i$. The Attention weight, denoted as $\alpha_{ij}$, is determined by a neural network that takes as input the feature vectors of asset $i$ and asset $j$ :

$$\alpha_{ij} = \frac{exp(ReLU(a^T[Wh_i \,||\, Wh_j]))}{\sum_{k\in N} exp(ReLU(a^T[Wh_i \,||\, Wh_j]))}$$

Here, $W$ represents the Linear weight to find latent features in the data vector created from Technical Indicators, and $a$ is the weight scaler value to calculate the attention weight. || denotes the concatenation of vectors, and *ReLU* is the activation function. The calculated attention weight represents the importance and correlation of asset *j* to asset *i*, and the GAT model uses this to compute the new feature for asset $h'_i$:

$$h'_i = \sigma\left(\sum_{j\in N} \alpha_{ij} W h_j\right)$$

In this context, $\sigma$ represents the sigmoid activation function. This model is not designed to forecast the stock market's short-term volatility; instead, it focuses on predicting fluctuations that hold significant meaning over a two-week rebalancing period. The goal is to classify the stock price movements that occur after this rebalancing interval, specifically identifying scenarios where prices ascend by more than 2% as class 1 (indicating an increase) and all other scenarios as class 0. Given the abundance of data points that reflect negligible price movement around 0% in the stock market, adopting a simplistic criterion for classifying increases could potentially lead the model to inaccurately magnify minor fluctuations—a 0.1% rise versus a 0.1% fall—as binary outcomes (1 and 0, respectively). Such a binary approach could inadvertently introduce confusion in the model's ability to forecast genuine value changes, hence the decision to implement a classification criterion that seeks to transcend the stock market's inherent volatility by identifying truly significant price movements. This nuanced approach underscores the model's sophistication in distinguishing between meaningful market trends and routine stock price movements, thereby enhancing the predictive accuracy and utility for investors aiming to navigate the complexities of financial markets.

Portfolio rebalancing method and backtesting process : After training the GAT model, the process of backtesting the rebalancing strategy for the stock portfolio was conducted. The backtest was based on the softmax output values of the GAT model, which predict the probability of stocks increasing by more than 2% during the rebalancing period. Through analyzing the predicted probabilities for each ticker, the top $k$ stocks with the highest probabilities were selected to set the asset universe for the given rebalancing date.

The weighting of the selected $k$ stocks was determined by analyzing each asset's expected return, variance, and the correlation between assets, reconstructed based on the theory of complex systems. Notably, the correlation between assets utilized a blend of the correlation learned by the GAT model and the actual correlation between assets at a ratio of 1:9. The rationale behind using this mixed correlation approach is supported by the findings of Seong and Nam (2022) which demonstrated that employing a complex systems-based approach to remove noise from the adj close values of each asset and calculate portfolio weights can enhance the efficiency of the rebalancing process. Thus, this research extends the methodologies of previous studies by using the correlation information learned through the GAT model in conjunction with noise-reduced correlation data to develop a more sophisticated portfolio rebalancing strategy.

Based on the determined returns, variances, and correlations of each asset, the optimal weights for each asset were calculated applying Modern Portfolio Theory. Through this optimization process, the study conducted the rebalancing of the portfolio composed of the selected stocks, thereby deriving a strategy that maximizes the portfolio's performance during the rebalancing period.

The backtesting process and portfolio rebalancing strategy of this research demonstrate the implementation of efficient asset management and rebalancing strategies amidst the volatility of the stock market, through the fusion of advanced analytical techniques utilizing the GAT model and noise reduction methodologies based on complex systems theory. This approach showcases the potential for achieving enhanced portfolio management and strategic rebalancing in the fluctuating stock market landscape.

## 3. Results

In Fig.1, we present the backtesting simulation's accumulated return results for four distinct portfolios: the market portfolio (equal proportions across all stocks), the Markowitz portfolio (optimized mean-variance for all stocks), a simple DNN model portfolio, and our GAT model portfolio. For the DNN and GAT model portfolios, we first selected the top five stocks most likely to achieve a 2% or more after the two-week rebalancing period (10 business days), subsequently determining the optimal asset weights. Specifically, in the case of the GAT model, we combined the correlation between assets derived from the training process of the GAT model and noise reduction methodology through complex system theory to find the optimal weights of assets, while in the case of the DNN model, we only applied noise reduction methodology. The results show that the accumulated return of the GAT model outperforms the other comparison portfolios. In addition to the accumulated return, we also compared the results against various metrics that are used to compare traditional backtesting results. (Table 2)

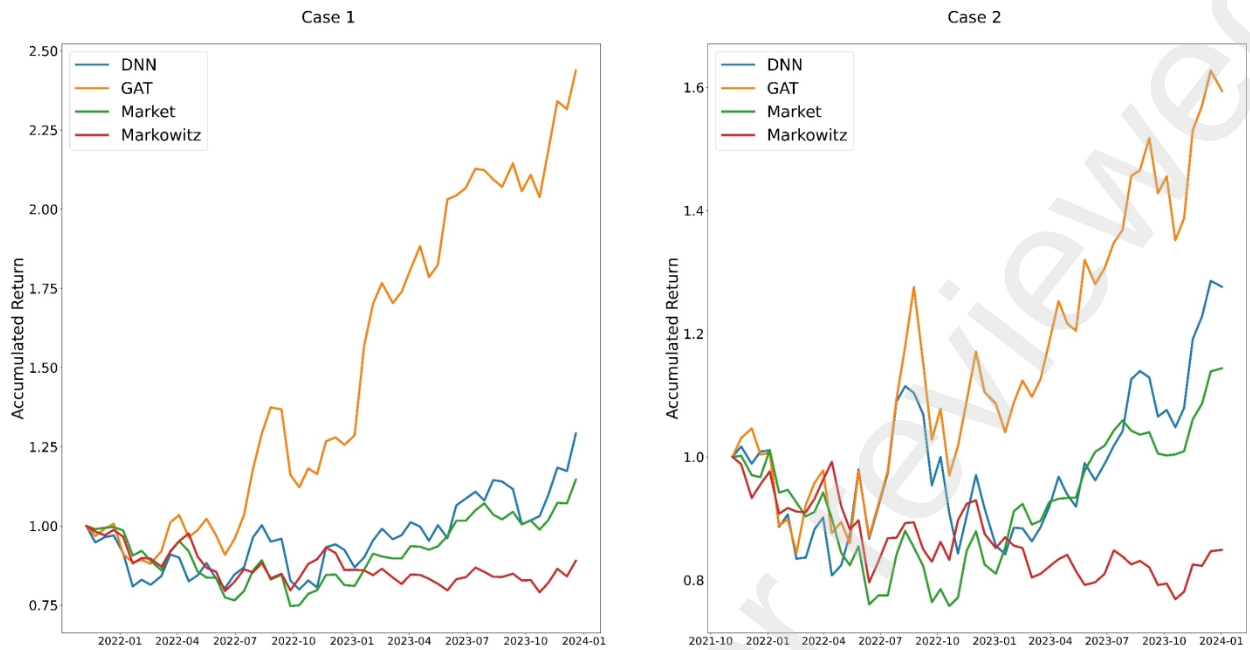


Fig. 1. Comparison of cumulative returns of portfolios

Table 2. Portfolio performance metrics for backtesting results.

| Metrics | Description | Equation |
|---|---|---|
| Accumulated Return | The value of a portfolio based on expected returns over a given period. | $Accumulated\ Return = \frac{Ending\ Value}{Beginning\ Value}$ |
| CAGR | Mean annual growth rate of an investment over a specified period. | $CAGR = \left(\frac{Ending\ Value}{Beginning\ Value}\right)^{\frac{Rebalance\ days}{Number\ of\ years}}$ |
| Sharpe Ratio | Performance of investment compared to a risk-free asset. | $Sharpe\ Ratio = \frac{R_p - R_f}{\sigma_p}$ |
| Beta | The volatility, or systematic risk, of a portfolio in comparison to the market. | $Beta\ (\beta) = \frac{Cov(R_p, R_m)}{Var(R_m)}$ |
| Sortino Ratio | Variation of the Sharpe Ratio that differentiates harmful volatility from total overall volatility by using the standard deviation of negative asset returns. | $Sortino\ Ratio = \frac{(R_p - R_f)}{\sigma_d}$ |
| Jensen's Alpha | The excess return of a portfolio over the expected market returns, based on the capital asset pricing model (CAPM). | $Jenesn's\ Alpha\ (\alpha) = R_p - (R_f + \beta(R_m - R_f))$ |
| Treynor Ratio | The returns earned in excess of that which could have been earned on a risk-free investment per unit of | $Treynor\ Ratio = \frac{(R_p - R_f)}{\beta}$ |

market risk.

$R_p$ is the return of the portfolio, $R_f$ is the risk-free rate, $R_m$ is the expected market return, $\sigma_p$ is the standard deviation of the portfolio's excess return, $\sigma_d$ is the standard deviation of the negative returns of the portfolio

Beyond accumulated return, we can see that the GAT model outperforms the Market, Markowitz, and DNN model portfolios on portfolio performance metrics such as CAGR and Jensen's Alpha with values of 1.2697 and 0.0075. However, Beta show slightly higher results compared to the other portfolios, which can be interpreted as evidence that the portfolio with the GAT model has aggressive investment characteristics. Though one might think that the aggressive characteristics of its investments mean it has some risk it also performs well compared to other portfolios on metrics like the Shape Ratio, Sortino Ratio, and Treynor Ratio with values of 0.8837, 0.9351, and 0.0537, which measure the risk of the portfolio. (Table 3)

Table 3. Backtesting result of portfolios.

| Portfolio | Accumulated Return | CAGR | Sharpe Ratio | Beta | Sortino Ratio | Jensen's Alpha | Treynor Ratio |
|---|---|---|---|---|---|---|---|
| Market | 1.154 | 1.071 | 0.426 | 1 | 0.412 | 0 | 0.018 |
| Markowitz | 0.870 | 0.936 | -0.253 | 0.637 | -0.236 | -0.004 | -0.015 |
| DNN | 1.377 | 1.16 | 0.6250 | 1.032 | 0.587 | 0.004 | 0.039 |
| GAT | **1.685** | **1.270** | **0.884** | 1.092 | **0.935** | **0.008** | **0.054** |

To validate the robustness of the results from our model, we conducted an additional ablation study on the effect of the complex system theory-based noise reduction methodology employed in the process of weighting the five assets selected through the GAT model. For each of the five assets selected through the GAT model, we compare the results of 1) all five assets are weighted equally, 2) the portfolio is weighted using the traditional Markowitz methodology, and 3) the portfolio is weighted using the correlation information learned through the GAT model and the correlation information reduced by the complex system theory-based noise reduction methodology. As shown in Fig. 2, we can see that the portfolio weight selection method proposed by our model yields higher returns than equally weighted and the traditional Markowitz method. This suggests that the method proposed by our model significantly impacts the portfolio's performance.

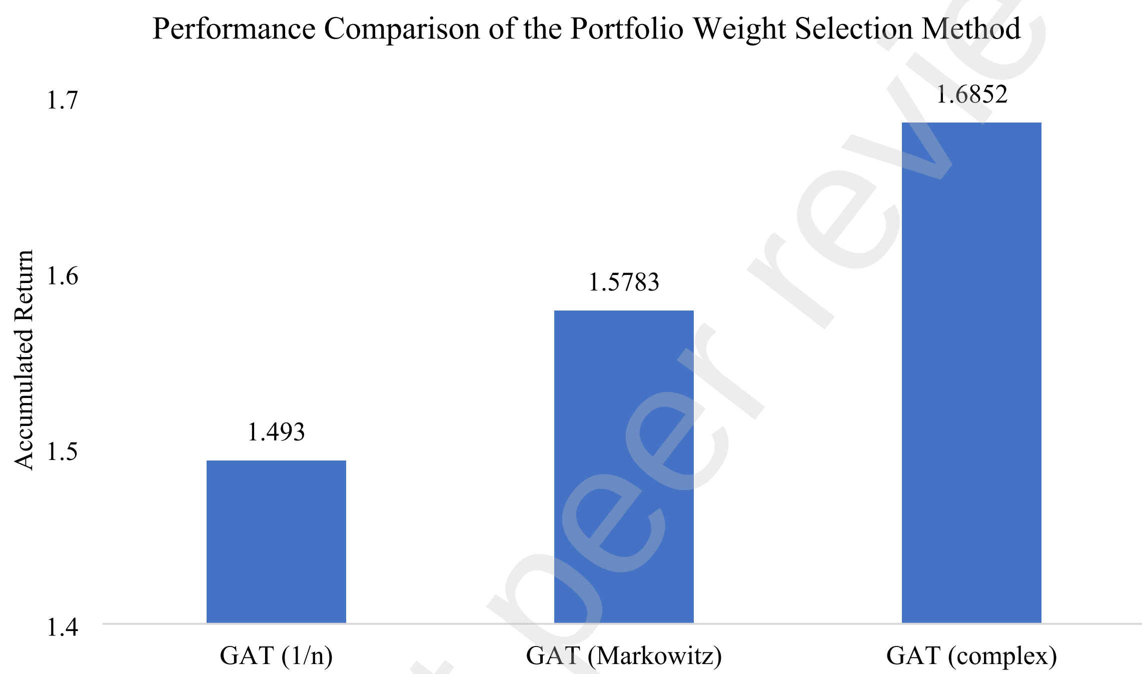


Fig. 2. Ablation study for portfolio weight selection method

As a parameter study, we also conducted an additional comparative experiment to see if the parameter of 2% is a significant value in the process of identifying stocks that will rise after the rebalancing period predicted by the model. We are interested in seeing if the conservative value of 2% is justified, as the stock market often fluctuates around 0%, which is why a simple division into rising and falling stocks based on 0% is not a practical option.

Table 4 displays the backtesting results of our proposed GAT model with weight thresholds set at 0, 0.01, 0.02, and 0.03. The findings indicate that the portfolio metrics, when parameterized with a 0.02 weight threshold, outperform in all aspects. Instead, specifying a threshold for labeling at a value higher than 0.02 results in relatively fewer assets increasing above that threshold after the rebalancing period, so backtesting simulation of portfolios with poor learning through the GAT model can lead to poor results. This demonstrates the effectiveness of a conservative weight parameterization strategy.

Table 4. Parameter study for weight thresholds.

| Weight | Accumulated Return | CAGR | Sharpe Ratio | Beta | Sortino Ratio | Jensen's Alpha | Treynor Ratio |
|---|---|---|---|---|---|---|---|
| 0 | 1.309 | 1.129 | 0.610 | 0.687 | 0.749 | 0.003 | 0.042 |
| 0.01 | 1.368 | 1.153 | 0.603 | 1.007 | 0.669 | 0.004 | 0.036 |
| 0.02 | **1.685** | **1.270** | **0.884** | 1.092 | **0.935** | **0.008** | **0.054** |
| 0.03 | 1.374 | 1.149 | 0.537 | 1.222 | 0.595 | 0.003 | 0.029 |

Collectively, these findings affirm the GAT model's enhanced performance and risk management over traditional and DNN model portfolios. It especially proves that portfolio diversification through complex system theory-based noise reduction methodology is effective by learning the relationship between individual assets to capture the meaningful volatility of the market rather than the short-term volatility of the stock market.

## 4. Conclusions

In our research, we introduce an effective approach to portfolio optimization that leverages a GAT model for identifying significant asset correlations through graph analysis, combined with a noise reduction technique grounded in complex system theory. Unlike previous artificial intelligence-based studies on portfolio optimization, which often overlook the intricate relationships between assets, our method acknowledges and intends to capture the significant interdependencies that exist among assets. This study aims to validate the hypothesis that GAT model can enhance portfolio optimization effectiveness. Moreover, we investigate whether integrating GAT-derived correlations with complex system theory for noise reduction further refines this optimization process. We present evidence in favor of our innovative model and methodology by conducting extensive backtesting simulations and comparing our GAT model's performance against traditional portfolios (Market, Markowitz) and DNN models.

Our contribution to the field of financial research is the confirmation of our approach that a novel methodology called GAT can effectively learn the interactions between assets that have been overlooked in existing artificial intelligence-assisted portfolio optimization studies, and combine it with a complex system theory-based noise reduction method extends existing traditional portfolio optimization methods and produces superior results in terms of return and risk.

A limitation of this study is that when determining correlations via graphs, each asset is placed in the same context, which could lead to inaccurate training of correlations for assets sharing similar characteristics as the number of assets increases. Therefore, to address this, we suggest a method for future research that groups assets with similar characteristics (e.g., the same sector or industry) and further reflects the relationship between similar characteristics using graph-based GATs.

**References**

1. Barua, R., and Sharma, A. K., 2022. Dynamic Black Litterman portfolios with views derived via CNN-BiLSTM predictions. Finance Research Letters 49.
2. Brody, S., Alon, U., and Yahav, E., 2021. How attentive are graph attention networks? arXiv preprint arXiv:2105.14491.
3. Chen, C. H., Su, X. Q., and Lin, J. B., 2016. The role of information uncertainty in moving-average technical analysis: A study of individual stock-option issuance in Taiwan. Finance Research Letters 18, 263-272.
4. Fernández, A., and Gómez, S., 2007. Portfolio selection using neural networks. Computers & Operations Research 34, 1177-1191.

5. Gorenc Novak, M., and Veluščck, D., 2016. Prediction of stock price movement based on daily high prices. Quantitative Finance 16, 793-826.

6. Harlow, W. V., 1991. Asset Allocation in a Downside-Risk Framework. Financial Analysts Journal 47, 28-40.

7. Henrique, B. M., Sobreiro, V. A., and Kimura, H., 2018. Stock price prediction using support vector regression on daily and up to the minute prices. Journal of Finance and Data Science 4, 183-201.

8. Jang, B. G., and Park, S., 2016. Ambiguity and optimal portfolio choice with Value-at-Risk constraint. Finance Research Letters 18, 158-176.

9. Jang, J., and Seong, N., 2023. Deep reinforcement learning for stock portfolio optimization by connecting with modern portfolio theory. Expert Systems with Applications 218.

10. Ji, G., Yu, J., Hu, K., Xie, J., and Ji, X., 2022. An adaptive feature selection schema using improved technical indicators for predicting stock price movements. Expert Systems with Applications 200, 116941.

11. Ma, C. Y., and Yan, S., 2022. Deep learning in the Chinese stock market: The role of technical indicators. Finance Research Letters 49.

12. Markowitz, H., 1952. Portfolio Selection. The Journal of Finance 7, 77-91.

13. Nam, K., and Seong, N., 2019. Financial news-based stock movement prediction using causality analysis of influence in the Korean stock market. Decision Support Systems 117, 100-112.

14. Paiva, F. D., Cardoso, R. T. N., Hanaoka, G. P., and Duarte, W. M., 2019. Decision-making for financial trading: A fusion approach of machine learning and portfolio selection. Expert Systems with Applications 115, 635-655.

15. Peng, Y., Albuquerque, P. H. M., Kimura, H., and Saavedra, C. A. P. B., 2021. Feature selection and deep neural networks for stock price direction forecasting using technical analysis indicators. Machine Learning with Applications 5, 100060.

16. Ramezanian, R., Peymanfar, A., and Ebrahimi, S. B., 2019. An integrated framework of genetic network programming and multi-layer perceptron neural network for prediction of daily stock return: An application in Tehran stock exchange market. Applied soft computing 82, 105551.

17. Seong, N., and Nam, K., 2022. Forecasting price movements of global financial indexes using complex quantitative financial networks. Knowledge-Based Systems 235.

18. Shynkevich, Y., McGinnity, T. M., Coleman, S. A., Belatreche, A., and Li, Y., 2017. Forecasting price movements using technical indicators: Investigating the impact of varying input window length. Neurocomputing 264**,** 71-88.

19. Silva, N. F., de Andrade, L. P., da Silva, W. S., de Melo, M. K., and Tonelli, A. O., 2024. Portfolio optimization based on the pre-selection of stocks by the Support Vector Machine model. Finance Research Letters 61**,** 105014.

20. Song, Q., Liu, A. Q., and Yang, S. Y., 2017. Stock portfolio selection using learning-to-rank algorithms with news sentiment. Neurocomputing 264**,** 20-28.

21. Thawornwong, S., Enke, D., and Dagli, C., 2003. Neural networks as a decision maker for stock trading: a technical analysis approach. International Journal of Smart Engineering System Design 5**,** 313-325.

22. Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., and Bengio, Y., 2017. Graph attention networks. arXiv preprint arXiv:1710.10903.

23. Wang, X., Ji, H., Shi, C., Wang, B., Ye, Y., Cui, P., and Yu, P. S., Heterogeneous graph attention network. The world wide web conference, 2019, pp. 2022-2032.

24. Weng, B., Ahmed, M. A., and Megahed, F. M., 2017. Stock market one-day ahead movement prediction using disparate data sources. Expert Systems with Applications 79**,** 153-163.

25. Zhou, Z. B., Song, Z. Y., Xiao, H. L., and Ren, T. T., 2023. Multi-source data driven cryptocurrency price movement prediction and portfolio optimization. Expert Systems with Applications 219.