DCAGAT: A Graph Attention-Based Model with Reconstruction Regularization for Dollar-Cost Averaging Investment Prediction

Zhongli Lv
School of Economics and Management, Guizhou Normal University, Guiyang,550001, China
E-mail: zhongliLv@outlook.com

Keywords: Graph Attention Network (GAT), Dollar-Cost Averaging (DCA), Financial Market Prediction, Multi-Feature Fusion.

Received: April 29, 2025

The financial market is characterized by high volatility and noisy data, making it a formidable challenge to forecast trends and design robust investment strategies. In this paper, we propose an innovative prediction model that integrates multi-feature fusion with graph attention mechanisms to address these challenges, specifically tailored for dollar-cost averaging (DCA) strategies. Our model, termed DCAGAT, leverages Graph Attention Networks (GATs) to dynamically assess the interdependencies among various financial assets. By incorporating multiple market features—historical price fluctuations and trading volumes—the model constructs a comprehensive representation of market dynamics. A key innovation is the inclusion of an autoencoder-inspired reconstruction verification mechanism, which mitigates overfitting by ensuring that the model focuses on persistent market trends rather than transient noise. We validate the effectiveness of DCAGAT using historical data from Yahoo Finance and ETF from 2012 to 2022. We benchmark DCAGAT against four neural baselines—DNN, Conv1D, GCN and ST-GCN—on three DCA-oriented metrics: one-, seven- and fourteen-day directional accuracy (Adir) and Top-5 hit-rate (Aselect). While DCAGAT matches the best baseline on Adir, it consistently improves Aselect across the evaluated scenarios, with every gain passing the paired statistical-significance test., underscoring its superior stock-selection capability. Overall, our research provides a robust framework for financial market forecasting by combining advanced graph-based learning techniques with feature-rich data integration, offering valuable insights for investors seeking to optimize multi-day investment decisions in unpredictable market environments.

Povzetek: Prispevek predstavi nov model za napovedovanje finančnih trgov, ki izboljša naložbene odločitve tudi v nestabilnih in šumnih razmerah.

1 Introduction

Finance is inherently a game of risk and reward, and since the inception of the financial industry, forecasting market trends, mitigating risks, and enhancing returns have remained paramount objectives for practitioners[1, 2]. The prevailing view posits that, to a certain extent, financial markets are predictable and that both risk and return can be systematically evaluated [3]. For example, armed conflict has been observed to drive up gold prices, technological breakthroughs can trigger increases in stock prices[4], and favorable weather conditions may lead to declines in agricultural commodity futures prices. Nonetheless, the process of evaluation is considerably more complex in practice; in actual trading environments, such clear-cut trends are often elusive, with the market itself frequently serving as the most reliable indicator of underlying economic forces[5, 6].

Dollar-cost averaging (DCA) represents a unique investment strategy characterized by its systematic and periodic allocation of fixed funds into selected financial assets, irrespective of market price fluctuations[7]. This method appeals to a diverse group of investors—including novice individuals seeking to minimize the emotional and financial risks associated with market timing, as well as more experienced investors who value a disciplined, multi-day approach. Among its advantages, DCA mitigates the adverse effects of market volatility and reduces the risk of mistimed investment decisions, although it may underperform during extended bull markets; nevertheless, its benefits often outweigh these limitations[8]. Fundamentally, the strategy relies on historical market performance as its primary informational input, thereby anchoring investment decisions in empirically observed trends rather than speculative forecasts[9].

Rapid advances in artificial intelligence (AI) have precipitated its widespread adoption across diverse domains such as entertainment[10], healthcare[11], manufacturing[12], robot[13], and transportation[14]. In each of these fields, AI's data-driven methodologies have demonstrated remarkable predictive capabilities, a strength that naturally extends to the realm of finance[15, 16]. This technological prowess has spurred financial practitioners to explore innovative approaches, including the ambitious pursuit of constructing a "Laplace demon"—a theoretical, data-driven model capable of processing vast amounts of market information to make highly informed predictions[17, 18].

However, the application of AI in finance presents unique challenges. Unlike machine vision or natural language processing, financial data is inherently more complex[19, 20]. The critical patterns and signals are deeply embedded within a noisy environment, making them more elusive and difficult to extract[21, 22]. Despite these hurdles, leveraging AI in financial analytics holds significant promise. By integrating historical market performance as a primary informational input, sophisticated computational models can gradually unravel these complexities, enhancing risk assessment and market forecasting capabilities[23]. Ultimately, the transformation of raw data into actionable insights not only reinforces AI's pivotal role in financial innovation but also sets the stage for more informed and strategic investment decisions[24, 25].

Existing studies center on short-horizon tasks—one-day price forecasts or binary direction probabilities—whereas multi-step prediction and strategies tailored to Dollar-Cost Averaging (DCA) receive scant attention. As Table 1 shows, most models optimize single-step direction accuracy, leaving the portfolio-level, multi-period allocation challenge intrinsic to DCA largely unaddressed.

Table 1: Representative graph/autoencoder models for financial forecasting

Year	Study / Model	Core Architecture	Public Dataset & Market Span	Pred. Horizon	Eval. Metrics	Reported Best Outcome
2025	**LSTMAE-Network** [26]	LSTM + Autoencoder	S&P 500 indices, 2000 – 2022	Next-day	Topological network metrics; crisis detection recall	Detected 5/6 major crises; F1 ↑ 10 % vs. correlation graph
2024	**SAE-MLP** [27]	Supervised Autoencoder + MLP	S&P 500, EUR/USD, BTC/USD, 2010 – 2022	5 – 30 min	Sharpe, Info-Ratio	Information Ratio ↑ 32 % vs. MLP baseline
2024	**DGDNN** [28]	Decoupled Graph Diffusion GNN	2 893 stocks (NASDAQ/NYSE/ SSE), 2012 – 2021	Next-day	Accuracy, MCC	Acc. ↑ 9.1 % / MCC ↑ 0.09 over 12 baselines
2023	**DGATS** [29]	Dynamic-Attributes GAT	U.S. equities, 2014 – 2020	Next-day	Accuracy	Acc. ↑ 5 % vs. dynamic-GCN
2021	**ST-Trader** [30]	VAE-GCN-LSTM	Minute-level NYSE stocks, 2018	30 min	Direction accuracy	Acc. ↑ 4 % vs. GCN-LSTM w/o VAE

According to Kraken's research depicted in Figure 1, 59% of cryptocurrency investors utilize DCA as their primary investment strategy for different reasons[31]. Considering the inherent complexity of the stock market and the potential interdependencies among individual stocks, this paper proposes a financial market prediction model that leverages a graph attention mechanism combined with multi-feature fusion to develop comprehensive DCA recommendations[32]. Moreover, whereas short-horizon signals may offer high-frequency cues, they also carry elevated noise and sudden-event risk, reducing confidence for DCA's medium-range allocations. Therefore, we integrate an autoencoder-inspired reconstruction verification mechanism to mitigate attention overfitting: this encourages the model to focus on stable, persistent patterns rather than transient spikes. In summary, our aims are:

1. To develop robust multi-step (e.g., 7–14 day) forecasts suited for DCA investors, balancing information sufficiency against noise in these horizons.
2. To mitigate overfitting in graph attention via an autoencoder-style reconstruction penalty, fostering durable feature representations.
3. To evaluate the proposed DCAGAT model on realistic portfolio sizes, demonstrating its practical value for DCA decision-making.

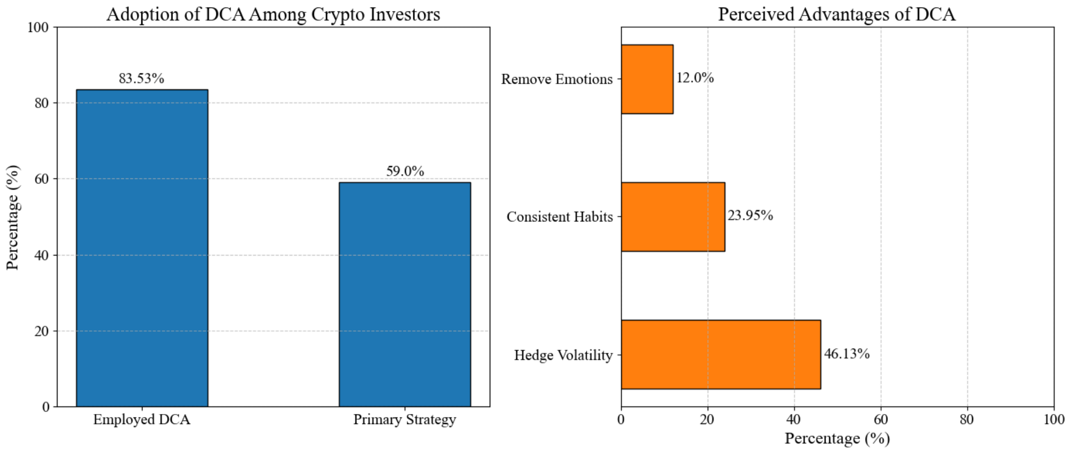

Figure 1: The visualization of Kraken's research on DCA

The remainder of this paper is organized as follows. Section II provides the necessary background on Graph Neural Networks and Graph Attention Networks. Section III introduces the proposed methodology in detail. Experimental setup and results are presented in Section IV. Section V offers a discussion of the findings, and Section VI concludes the paper.

2 Background

2.1 Graph neural networks

Graph Neural Networks (GNNs) have emerged as a pivotal tool for learning from graph-structured data by explicitly incorporating the inherent relational information into the learning process[33]. Consider a graph G defined as:

$$\mathrm{G} = (\mathrm{V}, \mathrm{E}) \quad (1)$$

where V represents the set of nodes and E denotes the set of edges, encapsulating the relationships among nodes.

Each node $v_i \in \mathrm{V}$ is associated with a feature vector $\mathbf{x}_i$ in $\mathbb{R}^d$:

$$\mathbf{x}_i \in \mathbb{R}^d, \quad v_i \in \mathrm{V} \quad (2)$$

The fundamental concept underlying GNNs is the iterative refinement of node representations through a message-passing mechanism, which leverages the graph structure as prior knowledge. Specifically, at the l-th layer, the hidden representation $\mathbf{h}_i^{(l)}$ of node v_i is updated according to:

$$\mathbf{h}_i^{(l)} = \sigma\left(\sum_{j \in \mathrm{N}(i)} \mathbf{W}^{(l)} \mathbf{h}_j^{(l-1)} + \mathbf{b}^{(l)}\right) \quad (3)$$

In this equation, $\mathrm{N}(i)$ denotes the neighborhood of node v_i, $\mathbf{W}^{(l)}$ is the learnable weight matrix at layer l, $\mathbf{b}^{(l)}$ is the bias vector, and σ is a non-linear activation function, such as ReLU. The initial hidden state is established by setting:

$$\mathbf{h}_i^{(0)} = \mathbf{x}_i \quad (4)$$

By recursively applying the update rule in (3) over multiple layers, GNNs are able to capture both local and global structures of the graph. The final node embeddings, $\mathbf{h}_i^{(L)}$ after L layers, serve as comprehensive representations for various downstream tasks, including node classification, link prediction, and graph-level inference. This explicit integration of structural information, as evidenced in (1)–(4), enables GNNs to effectively harness the relational context intrinsic to graph data.

2.2 Graph attention networks

Graph Attention Networks (GATs) extend the framework of GNNs by incorporating an attention mechanism, which allows the network to assign different weights to nodes in a neighborhood during the aggregation process[34]. This is particularly useful in scenarios where some neighbors contribute more significantly than others to a node's representation.

In GAT, each node v_i with feature vector $\mathbf{h}_i$ is first transformed by a linear mapping $\mathbf{W}$. To compute the attention coefficients, an attention mechanism is applied to every pair of connected nodes. Specifically, for nodes v_i and v_j, the unnormalized attention score is given by:

$$e_{ij} = \mathrm{LeakyReLU}\left(\mathbf{a}^T \left[\mathbf{W}\mathbf{h}_i \| \mathbf{W}\mathbf{h}_j\right]\right) \quad (5)$$

where $\mathbf{a}$ is a learnable weight vector. These scores are then normalized using the softmax function to obtain the attention coefficients:

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k \in \mathrm{N}(i)} \exp(e_{ik})} \quad (6)$$

The normalized coefficients α_{ij} indicate the importance of node v_i and v_j. Finally, the updated representation for node v_i is computed as a weighted sum of its neighbors' transformed features:

$$\mathbf{h}_{i'} = \sigma\left(\sum_{j \in \mathrm{N}(i)} \alpha_{ij} \mathbf{W}\mathbf{h}_j\right) \quad (7)$$

To further enhance the stability and performance of the model, multi-head attention is commonly employed, where multiple independent attention mechanisms operate in parallel and their outputs are either concatenated or averaged. This design allows GATs to dynamically modulate the influence of neighboring nodes, enabling the model to prioritize the most informative connections. As a result, the learned representations are more expressive and better suited for downstream tasks.

3 Methodology

Dollar-Cost Averaging (DCA) has emerged as an increasingly popular investment strategy, particularly among investors seeking to mitigate the inherent volatility of financial markets. By systematically investing a fixed amount at regular intervals, DCA reduces the risk associated with market timing, as it naturally smooths out the effects of short-term price fluctuations. This approach not only minimizes the potential for purchasing assets at their peak prices but also instills a disciplined, multi-day investment mindset that can be especially beneficial during periods of market instability.

However, the success of DCA extends beyond mere regularity in investments—it critically hinges on the selection of the right assets. While the strategy is inherently designed to counteract volatility, its effectiveness can be substantially enhanced when paired with predictive models that assess the multi-day potential of various stocks. Traditional predictive models tend to focus on short-term price movements, which may not align with the enduring goals of DCA. Therefore, there is a clear need to tailor these models specifically for DCA by incorporating a broader set of features that capture the sustained growth and resilience of financial assets over time. By customizing predictive frameworks to evaluate multi-day trends rather than ephemeral market noise, investors can maximize returns and improve the overall efficiency of their investment strategies. This tailored approach not only bolsters confidence in asset selection but also lays the groundwork for more stable and rewarding investment outcomes over the long haul.

3.1 The graph attention network for stock prediction

In predictive modeling, analyzing individual stocks in isolation is often inadequate due to the limited availability of reliable data and the presence of substantial noise, which can obscure meaningful patterns. A more comprehensive perspective that considers the broader market context is essential, as stocks do not operate independently but are inherently interrelated through various economic, industrial, and behavioral factors. Their collective dynamics often contain richer and more stable signals that can significantly improve prediction quality. GNNs offer a powerful framework for capturing these complex interdependencies. By modeling stocks as nodes within a graph and their relationships—such as correlations, sector affiliations, or co-movement trends—as edges, GNNs facilitate the development of an integrated market model. This graph-based representation not only addresses the limitations of sparse or noisy individual stock data but also capitalizes on the relational structure among assets to produce more robust and informative predictions, thereby enhancing the overall accuracy and reliability of financial forecasting.

While GNNs effectively capture relational structures by connecting stocks within a market, they typically aggregate information from all neighboring nodes in a uniform manner. This approach may overlook the varying degrees of influence that different stocks exert on one another. In contrast, GATs incorporate an attention mechanism that assigns adaptive weights to each neighboring node, thus highlighting the more informative relationships. This weighted aggregation enhances the model’s ability to focus on critical interdependencies and attenuate the impact of noise or less relevant connections. In fact, GATs dynamically learn these weights during training, thereby improving the model's capacity to capture subtle market dynamics. To exploit these advantages, we have integrated a GAT layer into our methodology. This addition not only refines the aggregation process by emphasizing pivotal inter-stock relationships but also facilitates a more nuanced and accurate representation of the overall market structure.

For each node in the graph, the model incorporates multi-dimensional historical data of the corresponding stock, which may include, but is not limited to, features such as daily high, low, opening prices, and trading volume. This rich set of inputs allows the graph attention mechanism to capture not only the structural correlations between different stocks but also the temporal patterns and performance trends of each individual asset. By simultaneously considering inter-stock relationships and time-series behavior, the model gains a more comprehensive understanding of the market dynamics. This dual perspective enhances its ability to identify meaningful patterns, reduce the impact of noise, and ultimately improve prediction accuracy. A schematic representation of the overall model architecture is provided in Figure 2. Given that the prediction task in this context resembles a generation problem—though with a lower level of complexity—Mean Squared Error (MSE) is employed as the training loss to measure the discrepancy between predicted and actual values, as equation (8).

$$\mathrm{L}_{\mathrm{pred}} = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{y}_i - y_i\right)^2 \quad (8)$$

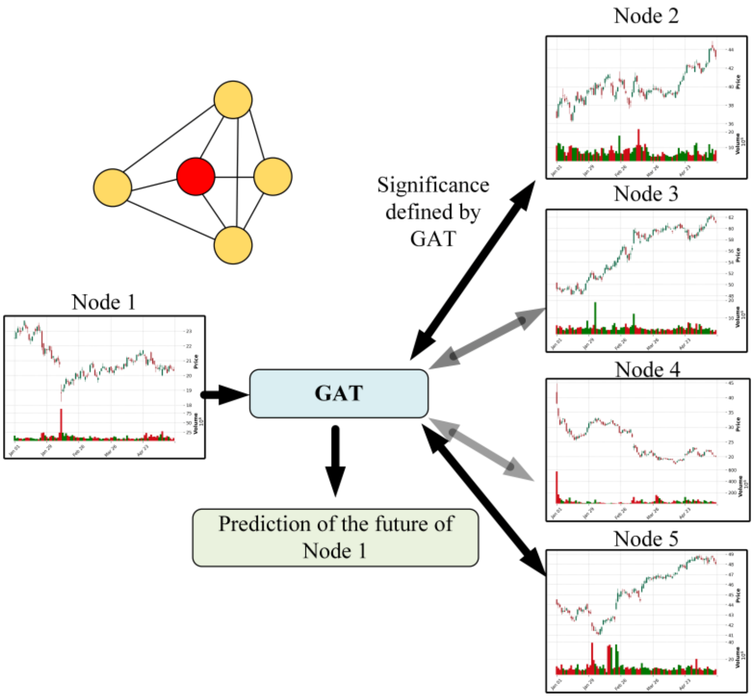

Figure 2. The diagram of GAT for stock prediction based on history of market.

3.2 Tailored design for dollar-cost averaging investment strategy

Financial market prediction poses a significant challenge due to the inherently volatile and nonlinear nature of market behavior. Sudden shifts driven by macroeconomic events, investor sentiment, or unforeseen external factors are commonplace, making consistent and reliable forecasting difficult. These challenges are particularly pronounced in the context of multi-day investment strategies such as DCA, which differs fundamentally from short-term or high-frequency trading. Unlike strategies that demand frequent recalibration based on market movements, DCA involves making periodic, fixed investments in selected assets over time—regardless of price fluctuations. This disciplined approach helps reduce emotional decision-making and cushions the impact of market volatility. However, its static nature limits the ability to respond to short-term opportunities, thereby placing greater emphasis on making strong, generalizable multi-day predictions at the outset.

Given these constraints, the design of predictive models for DCA must be approached with caution. Overreliance on narrowly defined features or recent market behavior can be detrimental, especially in volatile environments where transient signals may dominate the data. If the model concentrates too heavily on such patterns, it risks overfitting—capturing noise rather than structure, and failing to generalize in future scenarios. While GATs offer a powerful mechanism for modeling relational dependencies by assigning adaptive weights to neighboring nodes, they are not without shortcomings. In their standard form, GATs tend to focus attention disproportionately on nodes with strong short-term signals, which can amplify noisy or misleading correlations—particularly problematic in financial domains where causality is often ambiguous. Additionally, without proper regularization, the attention mechanism may become overly sensitive to minor variations in the input, leading to unstable representations. To counteract these issues, an effective model should minimize the risk of becoming overly influenced by any single factor or fleeting trend. Doing so encourages broader, more stable representations that better match the multi-day, noise-tolerant nature of DCA investment strategies.

Over-weighting a single factor or fleeting market blip creates an information bottleneck: when the latent representation fails to retain enough information to reconstruct the original inputs, essential signals are lost. Routine hyper-parameter tweaks—learning-rate schedules, deeper layers, extra attention heads—or even generic regularizes such as dropout and weight decay can temper, but seldom eliminate, this structural leak. To address it directly, we embed an autoencoder (AE) branch alongside the GAT. The requirement to compress each input sequence into a latent code and then faithfully rebuild it supplies a reconstruction loss that penalizes representations discarding informative dimensions, steering the attention mechanism toward broader, information-rich patterns. In the volatility-prone, noise-laden world of finance, this reconstruction-based regularization is particularly effective for preserving the robust multi-day signals that Dollar-Cost Averaging depends on.

To integrate an AE into the standard GAT workflow, we treat the GAT's node embeddings as the AE's latent code (Figure 3). Raw market features are first compressed by an encoder into a lower-dimensional vector, which is fed to the GAT. The GAT learns cross-asset dependencies and passes its embeddings to (i) a dense prediction head and (ii) a decoder that tries to reconstruct the encoder's original input. Because the decoder operates directly on these post-attention embeddings, the reconstruction loss regularizes the GAT's output. We minimize a joint objective that combines the prediction loss with a reconstruction loss measured by MSE as Eq. (9). This dual-loss scheme preserves the familiar "GAT → dense layer" paradigm while regularizing the attention mechanism, preventing it from overreacting to short-lived market shocks and fostering information-rich, noise-robust representations. The resulting architecture, DCAGAT, thus balances representation integrity with attention flexibility and achieves consistently higher predictive accuracy in the multi-day DCA setting.

$$L_{rec} = \frac{1}{N}\sum_{i=1}^{N}(\hat{x}_i - x_i)^2 \quad (9)$$

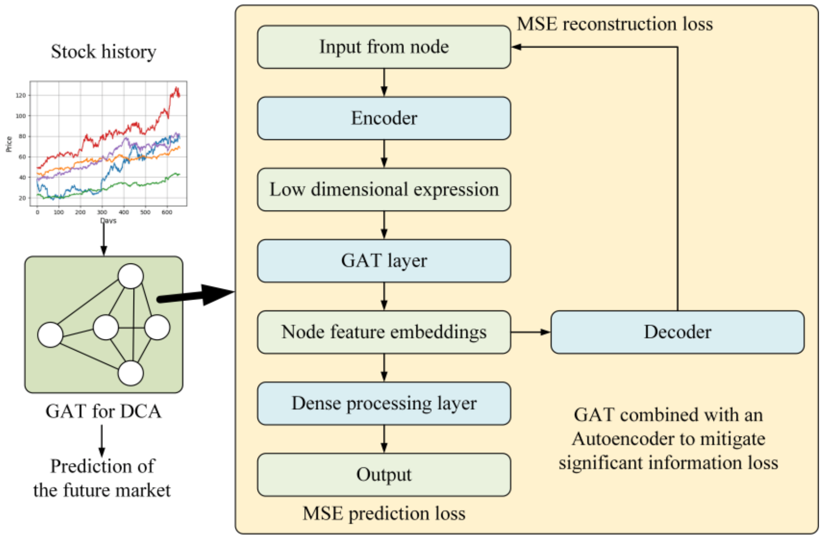

Figure: 3 The Diagram of DCAGAT.

Under this design, the model incurs two loss terms—one for prediction and one for reconstruction. Both terms are optimized jointly with user-defined weights, allowing the network to balance predictive accuracy against reconstruction fidelity. The overall objective is therefore the weighted sum:

$$L_{total} = \lambda_1 L_{pred} + \lambda_2 L_{rec} \quad (10)$$

promoting simultaneous improvement on both fronts.

In our model, both the prediction loss and the reconstruction loss are assigned a default weight of 1, ensuring that they contribute equally during the training process. This balanced weighting scheme is designed to optimize both the accuracy of the predictions and the quality of the reconstructed inputs. In the subsequent section, we conduct an ablation study to systematically evaluate the impact of the reconstruction component on overall model performance, shedding light on how this mechanism contributes to mitigating overfitting and enhancing the model's ability to generalize in complex market conditions. In summary, the modular algorithm description is as Algorithm 1:

Algorithm 1: DCAGAT Forward & Training Loop	
Input : node features X ∈ R^{N×T×F}, adjacency A ∈ {0,1}^{N×N}	
Output: predictions ŷ, reconstruction R	
H ← ENCODER(X)	# (N,T,F) → (N,d)
Z ← GAT (H, A)	# node embeddings
ŷ ← PREDICT_HEAD(Z)	# three targets per node
R ← DECODER(Z)	# reconstruct H
Lpred = MSE(ŷ, y)	# prediction loss
Lrecon = MSE(R, H)	# reconstruction loss
Ltotal = λp·Lpred + λr·Lrecon	
Update parameters with Adam	# minimise Ltotal

4 Experiments

4.1 Experimental setting

Yahoo Finance dataset provides historical financial data for a wide range of stocks and widely used for verification of stock prediction methods[35], as depicted in Figure 4. This dataset includes daily records of key features such as opening price, closing price, highest price, lowest price, and trading volume, spanning from January 1, 2012, to December 31, 2022.

Figure 4: A part of historical closing prices of stocks in yahoo finance dataset.

Besides, ETF dataset contains historical time-series records for one or more exchange-traded funds, restricted to daily market price fields (open, high, low, close) and trading volume over the period from 1993 to 2025. Such a dataset provides a consistent basis for benchmarking experimental studies. By focusing solely on these elements within the specified interval, the dataset remains well-defined and straightforward to process. Because the elements are identical, we omit the visualization here.

For our experiment, we focus on a selection of 20, 30, 50 stocks, chosen to represent a diverse and broad cross-section of the market. By incorporating this number of stocks, we create a comprehensive market picture that allows our model to capture inter-stock relationships and market-wide dynamics. For each stock, the past 30 days of records are provided. This dataset is particularly useful for our study, as it reflects the complex and noisy nature of real-world financial data, providing an ideal testbed for evaluating the predictive performance of DCAGAT. The wide variety of stocks ensures that our model is challenged to discern meaningful patterns across different sectors and market conditions.

It is important to acknowledge that due to the inherent volatility and randomness of the stock market, relying on market conditions from 2010 to predict outcomes in 2020 would be highly inappropriate. Instead, we have structured the dataset in a more meaningful manner by grouping the data into segments where three consecutive years of market information are used to forecast the market situation for the subsequent fourth year. This approach not only aligns more closely with the dynamic nature of financial markets but also ensures that the training and testing datasets reflect more coherent and temporally relevant trends. The detailed breakdown of these training and testing periods is presented in the Table II.

Table 2: The training and testing dataset

Subsets	Training dataset time period	Testing dataset time period
1	2012-01-01 to 2014-12-31	2015-01-01 to 2015-12-31
2	2015-01-01 to 2017-12-31	2018-01-01 to 2018-12-31
3	2018-01-01 to 2021-12-31	2022-01-01 to 2022-12-31

To conduct a comparative experiment, we reviewed several studies in the literature where GNNs and GATs have been used for stock prediction. However, there is limited research on applying these models to DCA investment strategies. Directly applying the methods from these existing studies would likely not yield good results, and such comparisons would have limited value. As such, our approach is to demonstrate the distinctions among the various mechanisms we employed—namely, traditional GNNs, the Graph Attention Mechanism, ST-GCN[36] (which is a SOTA combined convolutional layer in GNN and used for stock prediction), and our proposed DCAGAT model. This comparative analysis not only elucidates the unique advantages of each technique but also emphasizes the superior performance of DCAGAT in addressing the specific challenges associated with DCA strategies. By tailoring our methods to the multi-day, disciplined nature of DCA, we provide a more relevant and robust framework for evaluating stock selection and prediction in volatile financial markets.

Another important consideration is the evaluation criteria used to assess model performance. In conventional financial forecasting, metrics such as next-day price direction or pointwise prediction accuracy are commonly employed and serve as useful indicators of short-term forecasting capability. However, these metrics fall short when applied to DCA strategies, which emphasize multi-day consistency over immediate gains. DCA inherently downplays short-term fluctuations, making it necessary to adopt evaluation methods that reflect this investment philosophy. Therefore, beyond standard predictive accuracy, we introduce metrics that better align with DCA's temporal horizon. Specifically, we assess the model's ability to predict the direction of returns over medium-range windows—namely, 7-day and 14-day periods. These evaluation windows are designed to capture more stable market trends while reducing sensitivity to daily noise, thereby offering a more meaningful benchmark for DCA-oriented decision-making. The directional accuracy for these timeframes is computed using the following equation:

$$\mathrm{A}_{\mathrm{dir},k} = \frac{1}{N}\sum_{i=1}^{N} \mathrm{I}\left(\mathrm{sgn}(\hat{r}_{i,t+k}) = \mathrm{sgn}(r_{i,t+k})\right) \quad (11)$$

In equation (11), I is an indicator function that returns 1 if the predicted and actual directions match, and 0 otherwise. $\hat{r}_{i,t+k}$ and $r_{i,t+k}$ are the predicted and actual return sum for stock *i* on day *t+k*.

To ensure comparability, the stock prices in the original dataset are adjusted to reflect the percentage change. Furthermore, the primary role of a predictive model in stock market investment is typically to assist in stock selection, rather than predicting exact price values. Therefore, prediction accuracy alone is not of primary interest to end-users. Rather, what is of greater importance is the model's ability to recommend stocks. To this end, we introduce an additional evaluation metric: stock selection accuracy. This metric assesses the model's ability to correctly identify the top 5 performing stocks over the next 14 days, providing a more practical and realistic evaluation of its predictive power. This metric can be expressed as:

$$\mathrm{A}_{\mathrm{select}} = \frac{1}{N}\sum_{i=1}^{N} \mathrm{I}\left(\hat{S}_{t+14}^{(5)} \cap S_{t+14}^{(5)}\right) \quad (12)$$

$\hat{S}_{t+14}^{(5)}$ denote the set of the top 5 stocks by return on day 14, and $S_{t+14}^{(5)}$ denote the predicted set of the top 5 stocks.

Stability across time is essential, so we measure temporal consistency by computing the standard deviation of stock-selection accuracy $\mathrm{A}_{\mathrm{select}}$ across successive test windows. Each prediction is generated on inputs drawn from a different point in time; the resulting dispersion therefore captures how performance fluctuates under varying market conditions. All evaluation criteria for the model are calculated using testing datasets.

4.2 Experimental results

All experiments were run on an NVIDIA TITAN RTX GPU. Tables III–VIII summaries the results of the DNN, GNN, GAT, ST-GCN, and DCAGAT models on the Yahoo Finance and ETF datasets listed in Table II. Full architectural details of DCAGAT appear in Appendix I. Unless noted otherwise, each model was trained for 100 epochs with the Adam optimizer (learning rate 1×10^{-5}) and a mini-batch size of 32—settings that provided stable convergence in preliminary trials.

While interdependencies among stocks do exist, these relationships are not always immediately obvious. To mitigate the risk of inadvertently introducing overfitting factors due to hidden correlations, we opted to fully connect all nodes in the graph. In practice, this means that the adjacency matrix used as input to the network is an all-ones matrix with dimensions equal to the number of stocks, thereby ensuring that every possible inter-stock dependency is considered. This approach not only simplifies the network structure but also reinforces the model's ability to capture a broad spectrum of market dynamics by accounting for all potential interactions among stocks.

Table 3: The experimental results on 50 stocks in Yahoo

Dataset	Subset 1(%)					Subset 2(%)					Subset 3(%)				
Models	$A_{\mathrm{dir},1}$	$A_{\mathrm{dir},7}$	$A_{\mathrm{dir},14}$	A_{select}	*Std*	$A_{\mathrm{dir},1}$	$A_{\mathrm{dir},7}$	$A_{\mathrm{dir},14}$	A_{select}	*Std*	$A_{\mathrm{dir},1}$	$A_{\mathrm{dir},7}$	$A_{\mathrm{dir},14}$	A_{select}	*Std*
DNN	49.4	53.8	**55.2**	10.77	0.14	52.3	54.9	56.5	11.2	0.13	50.4	**50.8**	49.2	10.82	0.15
GNN	47.7	48.9	52.8	11.35	0.15	**52.7**	**55.0**	**56.6**	8.41	0.1	49.9	50.4	49.9	6.02	0.10
GAT	49.6	51.0	52.7	8.75	0.12	49.7	49.1	50.3	10.21	0.14	**52.8**	50.7	53.0	9.66	0.13
ST-GCN	50.	48.	47.	11.3	0.1	47.	48.	47.	8.41	0.1	50.	48.	**60.**	6.09	0.1

	6	1	5	5	5	6	7	8		0	1	8	**8**		0
DCAGAT	51.1	49.7	54.7	**12.12**	0.13	52.1	49.6	51.7	**11.43**	0.13	51.8	50.3	50.6	**11.82**	0.13

Table 4: The experimental results on 30 stocks in Yahoo

Dataset	Subset 1(%)					Subset 2(%)					Subset 3(%)				
Models	$A_{\mathrm{dir},1}$	$A_{\mathrm{dir},7}$	$A_{\mathrm{dir},14}$	A_{select}	*Std*	$A_{\mathrm{dir},1}$	$A_{\mathrm{dir},7}$	$A_{\mathrm{dir},14}$	A_{select}	*Std*	$A_{\mathrm{dir},1}$	$A_{\mathrm{dir},7}$	$A_{\mathrm{dir},14}$	A_{select}	*Std*
DNN	49.9	54.2	56.4	12.10	0.15	52.6	56.2	57.9	13.24	0.13	50.0	50.1	51.1	13.03	0.17
GNN	50.6	53.5	55.6	15.00	0.15	52.0	**56.2**	57.7	14.20	0.13	49.8	51.6	50.9	15.07	0.14
GAT	50.1	**49.9**	52.8	14.35	0.16	51.8	49.5	55.4	15.85	**0.15**	**49.0**	**49.7**	53.3	16.04	0.16
ST-GCN	49.3	**45.6**	56.4	15.00	0.15	**47.1**	56.2	42.2	14.20	**0.13**	**50.5**	**52.1**	51.4	15.07	0.14
DCAGAT	**50.3**	50.1	**55.1**	**15.29**	0.14	51.0	49.6	**52.9**	**17.39**	0.16	48.8	51.3	**52.2**	**19.52**	0.16

Table 5: The Experimental results on 20 stocks in Yahoo

Dataset	Subset 1(%)					Subset 2(%)					Subset 3(%)				
Models	$A_{\mathrm{dir},1}$	$A_{\mathrm{dir},7}$	$A_{\mathrm{dir},14}$	A_{select}	*Std*	$A_{\mathrm{dir},1}$	$A_{\mathrm{dir},7}$	$A_{\mathrm{dir},14}$	A_{select}	*Std*	$A_{\mathrm{dir},1}$	$A_{\mathrm{dir},7}$	$A_{\mathrm{dir},14}$	A_{select}	*Std*
DNN	49.4	52.8	54.1	20.92	0.17	53.0	54.4	**55.7**	21.57	0.17	51.4	**52.5**	50.9	24.93	0.16
GNN	**51.9**	**53.3**	**54.7**	21.71	0.16	**53.1**	**54.8**	54.3	22.30	0.16	51.4	51.4	50.3	25.12	0.16
GAT	49.3	51.6	52.0	23.27	0.16	49.4	48.6	51.5	21.20	0.17	49.3	52.1	52.1	26.30	0.18
ST-GCN	50.2	47.2	49.5	23.30	0.16	50.1	50.3	49.2	22.42	0.16	**51.7**	49.2	57.8	25.10	0.16
DCAGAT	46.4	52.2	53.4	**23.65**	0.15	49.0	50.1	53.8	**23.96**	0.17	51.3	47.7	**51.6**	**27.63**	0.17

Table 6: The Experimental results on 50 stocks in ETF

Dataset	Subset 1(%)					Subset 2(%)					Subset 3(%)				
Models	$A_{\mathrm{dir},1}$	$A_{\mathrm{dir},7}$	$A_{\mathrm{dir},14}$	A_{select}	*Std*	$A_{\mathrm{dir},1}$	$A_{\mathrm{dir},7}$	$A_{\mathrm{dir},14}$	A_{select}	*Std*	$A_{\mathrm{dir},1}$	$A_{\mathrm{dir},7}$	$A_{\mathrm{dir},14}$	A_{select}	*Std*
DNN	49.2	53.9	**55.6**	12.88	0.13	**52.5**	**55.3**	**57.2**	13.33	0.14	50.1	51.8	51.0	13.53	0.14
GNN	48.4	**54.1**	54.2	9.42	0.11	49.6	51.8	55.2	14.40	0.13	**52.5**	**52.1**	50.6	13.91	0.14
GAT	47.9	51.8	53.1	12.60	0.15	49.4	47.7	52.5	12.27	0.15	51.1	50.3	**55.8**	10.24	0.13
ST-GCN	**49.8**	45.9	55.7	9.42	0.11	51.3	55.0	51.2	14.40	0.13	50.2	51.9	51.0	13.91	0.14
DCAGAT	47.5	50.8	52.3	**13.46**	0.14	50.3	50.1	55.5	**14.75**	0.15	52.4	50.3	53.7	**14.62**	0.16

Table 7: The Experimental Results on 30 Stocks in ETF

Dataset	Subset 1(%)					Subset 2(%)					Subset 3(%)				
Models	$A_{\mathrm{dir},1}$	$A_{\mathrm{dir},7}$	$A_{\mathrm{dir},14}$	A_{select}	*Std*	$A_{\mathrm{dir},1}$	$A_{\mathrm{dir},7}$	$A_{\mathrm{dir},14}$	A_{select}	*Std*	$A_{\mathrm{dir},1}$	$A_{\mathrm{dir},7}$	$A_{\mathrm{dir},14}$	A_{select}	*Std*
DNN	50.0	54.0	55.7	17.79	0.15	52.7	**55.8**	57.6	18.65	0.15	50.3	51.9	51.4	18.65	0.16

GNN	47.1	**54.4**	**55.9**	15.48	0.14	49.6	54.7	**57.9**	19.23	0.15	49.7	51.7	50.9	17.10	0.15
GAT	51.7	49.0	54.4	17.46	0.15	49.8	50.4	53.3	17.29	0.15	**53.8**	49.8	51.2	20.29	0.16
ST-GCN	48.1	53.5	54.0	15.48	0.14	**53.9**	50.4	44.6	19.23	0.15	50.3	**52.0**	51.3	17.10	0.15
DCAGAT	**50.7**	49.7	52.8	**18.96**	0.14	51.5	49.6	53.9	**20.65**	0.16	48.6	48.6	**52.4**	**20.68**	0.17

Table 8: The experimental results on 20 stocks in ETF

Dataset	Subset 1(%)					Subset 2(%)					Subset 3(%)				
Models	$A_{\text{dir},1}$	$A_{\text{dir},7}$	$A_{\text{dir},14}$	A_{select}	*Std*	$A_{\text{dir},1}$	$A_{\text{dir},7}$	$A_{\text{dir},14}$	A_{select}	*Std*	$A_{\text{dir},1}$	$A_{\text{dir},7}$	$A_{\text{dir},14}$	A_{select}	*Std*
DNN	51.0	53.4	55.2	20.10	0.14	52.0	**55.0**	**56.7**	16.33	0.13	50.5	**52.3**	51.3	20.22	0.16
GNN	50.0	**54.3**	**55.3**	19.33	0.14	49.6	54.0	54.7	15.27	0.12	48.8	51.9	51.3	23.29	0.15
GAT	**51.9**	48.5	53.1	20.05	0.17	50.9	49.4	54.1	17.25	0.17	50.4	50.7	50.5	23.00	0.17
ST-GCN	49.7	50.5	47.4	19.33	0.14	**52.6**	53.1	49.7	15.27	0.12	50.4	52.2	49.2	23.29	0.15
DCAGAT	48.6	49.5	51.1	**21.35**	0.16	47.0	51.6	55.9	**22.61**	0.17	**50.6**	47.5	**53.3**	**25.60**	0.17

From the results, it is evident that the overall prediction accuracy remains relatively low—a reflection of the inherent randomness and volatility present in financial markets. Although increasing the number of stocks in the analysis makes the task of stock selection more challenging, it does not significantly affect the accuracy of directional predictions. This observation underscores the immense difficulty in forecasting stock prices, given that markets are influenced by a multitude of unpredictable factors, and it should always be clear that perfect prediction is a fantasy. Despite these modest performance levels, the results offer valuable insights into the comparative effectiveness of the various models tested, highlighting both the strengths and limitations of each approach in capturing the complex dynamics of the market.

First, we observe that the DCAGAT architecture yields a clear improvement in stock selection accuracy. As shown in Figure 5, the GAT-based models deliver notable improvements in stock selection—a critical and practical metric—while directional accuracy, though informative, is not the core concern; hence, subsequent figures focus only on selection accuracy. Therefore, subsequent figures present only selection accuracy. The proposed DCAGAT model outperforms alternative approaches in this respect, underscoring the benefit of integrating an autoencoder mechanism to preserve and reconstruct salient information. Although GNN and ST-GCN differ in directional prediction, their stock selection accuracies are comparable. While DCAGAT does not uniformly attain the highest prediction accuracy across all metrics (with DNN occasionally matching or exceeding performance), its superior selection accuracy highlights its enhanced practical value for investment decision-making, offering a more reliable strategy compared to competing models. The small standard-deviation values reported in Tables III–VIII indicate that the model's performance varies only marginally across successive evaluation windows, confirming that it remains stable throughout the entire test period.

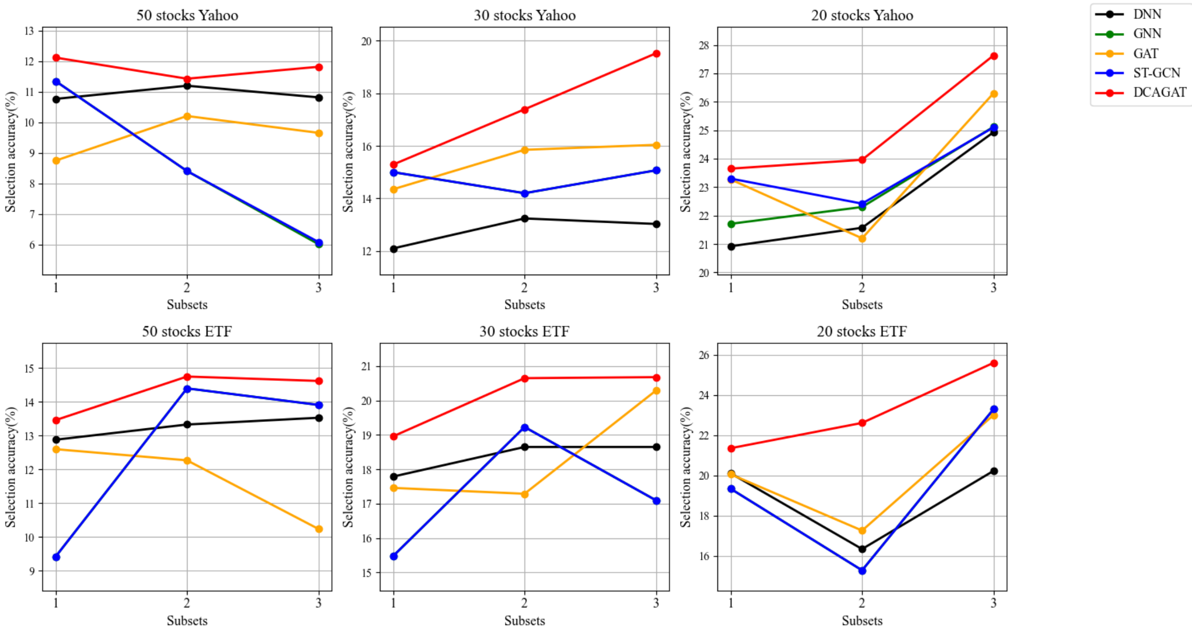

Figure 5: The visualization of aselect.

Although the performance gap in Figure. 5 is visually compelling, we verified its statistical validity with a paired t-test. Because the remaining baselines differ only marginally from one another, we compared our model solely with ST-GCN. For every experimental condition—two datasets and three portfolio sizes—we collected Top-5 hit-rates from the same 15 runs, yielding perfectly matched score pairs. The resulting p-values (Table IX) are all below the 0.05 threshold, allowing us to reject the null hypothesis of equal means and to conclude that DCAGAT's advantage in stock selection is statistically significant across all market settings.

Table 9: Paired t-test results demonstrating the statistical significance of DCAGAT vs. ST-GCN

p-values	Yahoo			ETF		
	Subset 1	Subset 2	Subset 3	Subset 1	Subset 2	Subset 3
50	0.021	0.002	0.002	0.000	0.023	0.013
30	0.032	0.012	0.001	0.000	0.002	0.000
20	0.035	0.002	0.012	0.001	0.000	0.015

Additionally, the number of stocks included in the prediction task plays a significant role in model performance. Empirically, we observe that models tend to achieve higher prediction accuracy when operating on a smaller set of stocks. This phenomenon is closely tied to the inherent randomness and noise present in financial markets. While expanding the stock pool may introduce more information and reveal broader market relationships, it simultaneously increases the complexity and noise within the data. The added noise can obscure useful patterns and make it more difficult for the model to generalize effectively. Therefore, it is reasonable that performance improvements may plateau—or even diminish—as the number of stocks increases. This observation highlights a key trade-off in financial modeling: expanding coverage can enhance representational richness, but it also demands more robust mechanisms to filter out irrelevant or misleading signals.

Although the observed improvements may not appear dramatic in absolute terms, they are nonetheless meaningful in the context of quantitative trading, where even small performance gains can translate into significant financial impact over time. This highlights the practical value of incorporating the proposed mechanism into the predictive model. In real-world trading systems, modest enhancements in stock selection accuracy or return direction prediction can compound across multiple trades and time periods, ultimately leading to notable returns. Therefore, the consistent edge provided by our model—however incremental—makes DCAGAT a valuable tool for financial applications, especially in strategies like DCA where multi-day consistency and robustness are prioritized. Such improvements reinforce the relevance of our approach in real trading environments, where every slight advantage can contribute to sustained strategic success.

5 Discussion

5.1 Weight configuration

Because the performance gaps we observe on the Yahoo and ETF datasets are statistically indistinguishable, the remainder of our analysis focuses on the Yahoo data set only. Here, we examine the weight configuration between prediction loss and reconstruction loss using three ratios: 1:1 (default), 1:2, and 2:1. These settings are designed to emphasize different objectives—1:2 leans toward reconstruction, while 2:1 favors prediction. The 1:1 ratio provides a balanced baseline, but varying the weights helps explore trade-offs and identify the optimal focus based on specific dataset characteristics.

As shown in Table X, the default 1:1 weight setting offers a competitive and balanced baseline, yielding strong performance across a variety of scenarios. This balanced approach ensures that the model pays equal attention to predictive accuracy and representational consistency. However, it is also evident that the 1:1 ratio is not universally optimal. In some cases, alternative weight combinations outperform the default setting, suggesting that the relative importance of prediction versus reconstruction can vary depending on dataset characteristics, noise levels, and task complexity. This highlights the need for adaptive or task-specific tuning of the loss weights rather than relying on fixed values. Overall, the findings suggest that while the 1:1 ratio serves as a solid default, a more nuanced calibration of loss components may be necessary to achieve optimal results in different financial contexts.

Table 10: The weight ratio experimentation and results on 50 stocks

DCAGAT	Subset 1(%)				Subset 2(%)				Subset 3(%)			
Weight ratio	$A_{dir,1}$	$A_{dir,7}$	$A_{dir,14}$	A_{select}	$A_{dir,1}$	$A_{dir,7}$	$A_{dir,14}$	A_{select}	$A_{dir,1}$	$A_{dir,7}$	$A_{dir,14}$	A_{select}
1:1	51.1	49.7	**54.7**	**12.12**	52.1	49.6	51.7	11.43	**51.8**	**50.3**	50.6	**11.82**
1:2	49.4	**50.4**	52.7	9.62	**52.5**	**50.9**	**52.6**	11.01	51.4	**50.3**	51.6	11.59
2:1	**54.4**	49.5	52.5	10.52	51.7	47.4	51.9	**12.83**	51.7	47.4	**51.9**	10.95

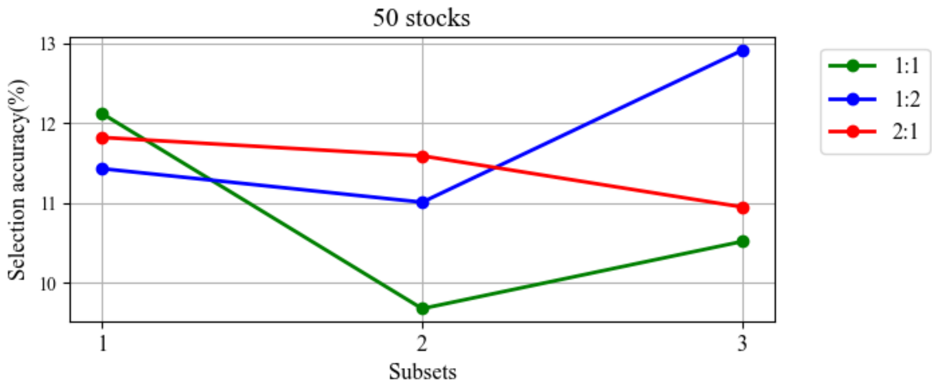

Figure 6: The visualization of Aselect under different weight configurations.

As shown in Figure 6, for stock selection accuracy, the 2:1 weight ratio outperforms the default 1:1 setting in Subset 2. This outcome is not unexpected, given that model performance is intimately linked to the underlying data characteristics. The results underscore a fundamental aspect of data-driven problems: the data itself plays a critical role in determining outcomes. This finding emphasizes the importance of customizing model parameters to align with the specific attributes of the dataset, reinforcing that data significantly influences the effectiveness of predictive models, particularly in stock market forecasting.

In the context of stock market forecasting, the inherent characteristics of stock data—such as volatility, trends, and interdependencies between stocks—play a pivotal role in determining model performance. Financial markets are dynamic and subject to constant fluctuations, meaning that model parameters, including the weight ratio between prediction and reconstruction, must be carefully calibrated to align with the underlying patterns and noise present in the data. Consequently, the success of stock prediction models is not solely reliant on the sophistication of their architectures but also on their ability to adapt effectively to the unique attributes of the training data. This adaptability ensures that models are better equipped to manage the complexities and uncertainties inherent in real-world financial environments.

5.2 Time configuration

In previous experiments, the historical window length was fixed at 30 days, which aligns with common investor routines. However, the length of historical data used for prediction can significantly impact model performance. To better understand this influence, we compare results using four different time windows: 15 days, 30 days, 45 days, and 60 days. This variation allows us to observe how the amount of past information affects the model's ability to capture trends and make accurate predictions. However, the results are presented in Table VI. In terms of prediction accuracy, there is no clear indication that any specific time window consistently outperforms the others. As shown in Figure 7, the 30-day time window appears to yield slightly better selection accuracy overall, but the advantage is marginal and not conclusive.

Table 11: the time window configuration experimentation and results on 50 stocks

DCAGAT	Subset 1(%)				Subset 2(%)				Subset 3(%)			
Time window	$A_{dir,1}$	$A_{dir,7}$	$A_{dir,14}$	A_{select}	$A_{dir,1}$	$A_{dir,7}$	$A_{dir,14}$	A_{select}	$A_{dir,1}$	$A_{dir,7}$	$A_{dir,14}$	A_{select}
15 days	**51.2**	53.4	**56.1**	10.76	51.8	**54.9**	56.4	9.82	**54.6**	50.8	47.3	**12.88**
30 days	51.1	49.7	54.7	**12.12**	52.1	49.6	51.7	**11.43**	51.8	50.3	50.6	11.82
45 days	49.6	**54.4**	55.7	11.71	50.2	51.0	53.0	10.52	48.3	**53.1**	53.0	10.94
60 days	47.9	54.0	55.9	11.91	**53.1**	51.7	53.7	11.41	49.9	49.2	**53.3**	10.73

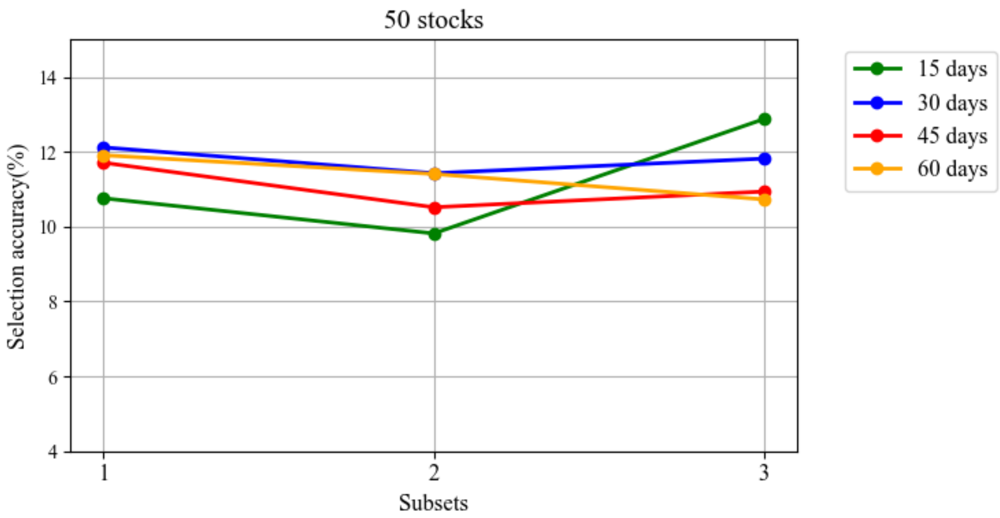

Figure 7: The visualization of Aselect under different time window configurations.

These findings suggest that there is no universal rule for determining the optimal length of historical data—the effectiveness of a given window largely depends on the characteristics of the underlying data. For certain stocks or market conditions, longer histories may be beneficial in capturing extended trends, delayed reactions, or cyclical patterns—such as those influenced by quarterly financial reports—making a 45- or 60-day window more effective. Conversely, shorter windows like 15 days may reduce noise and enhance responsiveness for stocks that are highly volatile or sensitive to recent news. This reinforces the importance of data-driven configuration, indicating that the ideal time window should be learned through experimentation and tailored to the specific context, rather than predefined.

5.3 Fully connected adjacency matrix graph

In our experiments we adopt a fully connected graph, thereby modelling every pair-wise stock interaction. While this choice guarantees that no potential dependency is omitted, it also admits many weak or spurious links, which can increase the risk of over-fitting. A natural next step is to repeat the analysis with a scarified graph—e.g., retaining only edges whose historical return correlation exceeds a threshold—and compare the resulting performance against the fully connected baseline.

Table 12: Impact of correlation-threshold sparsification on DCAGAT performance

CAGAT	Subset 1(%)				Subset 2(%)				Subset 3(%)			
threshold	$A_{dir,1}$	$A_{dir,7}$	$A_{dir,14}$	A_{select}	$A_{dir,1}$	$A_{dir,7}$	$A_{dir,14}$	A_{select}	$A_{dir,1}$	$A_{dir,7}$	$A_{dir,14}$	A_{select}
-	51.1	49.7	54.7	**12.12**	52.1	49.6	51.7	**11.43**	51.8	50.3	50.6	11.82
>0.3	49.9	55.6	57.0	11.35	52.4	54.3	56.3	9.51	49.7	49.9	48.3	13.04
>0.5	49.3	53.8	55.2	11.15	52.6	54.9	56.6	8.79	49.9	50.4	49.**9**	**13.62**

Table X shows that sparsifying the graph by correlation threshold has only a marginal effect on DCAGAT's overall performance. None of the differences in A _select across thresholds exceeds the standard-error band, indicating no statistically significant change in stock-selection accuracy. Nevertheless, the > 0.30 and > 0.50 graphs do yield small gains in specific subsets (e.g., A _select rises from 11.43 % to 13.04 % in Subset 3 at the 0.30 threshold). These modest, scenario-dependent improvements suggest that the correlation cut-off can be treated as a tunable hyper-parameter: selecting a threshold tailored to a particular market segment may provide incremental gains without materially altering the model's stability.

5.4 Practical deployment and limitations

In deployment, a single hyper-parameter sweep and minor architectural tweaks can lift metrics, demonstrating that practical adjustment does yield better performance. Yet two limitations remain clear: (i) for ultrashort horizons a vanilla GAT can still outscore DCAGAT, and (ii) even the best-tuned model identifies only 10 % of the top performers. These outcomes expose the market's chronically low signal-to-noise ratio, where external forces—especially investor sentiment—often drown out historical price patterns. Reliance on past prices alone therefore cannot sustain reliable near future forecasts. Lasting progress will hinge less on deeper model stacks and more on superior data pipelines: meticulous cleaning, richer temporal structures, and domain-aware features.

6 Conclusion

This paper presents a novel financial market prediction model, DCAGAT, which combines Graph Attention Networks with multi-feature fusion and an autoencoder-based reconstruction mechanism to address the challenges posed by the dollar-cost averaging investment strategy. The integration of GAT allows for adaptive weighting of stock relationships, enhancing the model's ability to capture meaningful market interdependencies. Our experiments demonstrate that DCAGAT significantly improves stock selection accuracy and prediction performance compared to traditional models, making it a promising tool for multi-day investment strategies. Despite the inherent volatility and randomness of financial markets, the results show that even modest improvements in prediction accuracy can have meaningful implications for quantitative trading, making DCAGAT a valuable approach for real-world financial applications.

Future work should place greater emphasis on data, as our experiments show that model performance is fundamentally constrained by the quality, structure, and relevance of the input information. Rather than relying solely on architectural enhancements, integrating richer data sources—such as macroeconomic indicators, industry reports, and market sentiment extracted from news or social media—could offer broader context and help mitigate noise. In parallel, more adaptive learning strategies like reinforcement learning could be explored to enable the model to dynamically adjust to evolving market conditions. Expanding the model's capability to perform multi-step predictions over extended time horizons would further improve its suitability for multi-day strategies such as DCA. Lastly, validating the model in real-time trading environments is critical to assessing its practical reliability and robustness, and to ensure that it can transition from a promising research framework to a viable tool for real-world financial decision-making.

Acknowledgement

The authors would like to thank all collaborators and reviewers for their valuable feedback and support throughout this research. This work was conducted with funding from the Department of Education of Guizhou Province, and the authors declare no conflicts of interest.

Appendix I

Table 1: Detail of DCAGAT

Layer	shape	Connected to
InputLayer	(None, market size, history length, 5)	/
Dense1	150, ReLU	InputLayer
Dense2	8, ReLU	Dense1
Adjacency Input	(None, market size, market size)	/
GAT	Channels = 8, Heads = 4, ReLU	[Dense2, Inputgraph]
Dense3	150, ReLU	GAT
Reconstruction Output	(None, market size, history length, 5)	Dense3
OutputLayer	(None, market size, 3)	GAT

Reference

[1] T. Marwala, "Opportunity Seeking Versus Risk Aversion," in The Balancing Problem in the Governance of Artificial Intelligence: Springer, 2024, pp. 51-67, DOI: 10.1007/978-981-97-9251-1_4.

[2] A. Rangapur, H. Wang, and K. Shu, "Investigating online financial misinformation and its consequences: A computational perspective," arXiv preprint arXiv:2309.12363, 2023.

[3] A. Zik-Rullahi, I. Jide, and E. Onuh, "Behavioural Finance: Exploring the Psychology and Economic Aspects of Financial Decision-Making," European Review in Accounting and Finance, 7 (3), 50-64. DOI: https://doi. org/10.5281/zenodo, vol. 8320116, 2023.

[4] N. Ahmad, L. Youjin, S. Žiković, and Z. Belyaeva, "The effects of technological innovation on sustainable development and environmental degradation: Evidence from China," Technology in Society, vol. 72, p. 102184, 2023, DOI: 10.1016/j.techsoc.2022.102184.

[5] T. Sugadev, N. S. Hameed, S. Vijayakumar, P. Tamilarasan, and M. S. Islam, "Portfolio optimization using machine learning techniques," in 2023 4th International Conference on Computation, Automation and Knowledge Management (ICCAKM), 2023: IEEE, pp. 1-7, DOI: 10.1109/ICCAKM58659.2023.10449598.

[6] Y. Li, "Essays in financial technology: banking efficiency and application of machine learning models in Supply Chain Finance and credit risk assessment," University of Glasgow, 2024.

[7] T. Kongkaew and C. Saengpao, "Investment Strategies in Covid-19 Pandemic: Comparative Analysis of Stock Investment Performance in the SETTHSI Index and the SET100 Index Utilizing Dollar Cost Averaging (DCA), Lump Sum (LS), and Value Averaging (VA) Strategies," Journal of Business, Innovation and Sustainability (JBIS), vol. 19, no. 2, 2024.

[8] E. Vasileiou and E. Hadad, "Benefits and drawbacks of the cost average plan as an alternative investment strategy in EMU countries," Cogent Economics & Finance, vol. 12, no. 1, p. 2426539, 2024, DOI: 10.1080/23322039.2024.2426539.

[9] H. C. Carvalho, "Riding the Waves A Comprehensive Study of 20/50 EMA and Dollar Cost Averaging in Equities, Stock Indexes, and Forex Market," Universidade NOVA de Lisboa (Portugal), 2024.

[10] T. Kaneko, H. Kameoka, K. Tanaka, and N. Hojo, "Cyclegan-vc2: Improved cyclegan-based non-parallel voice conversion," in ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019: IEEE, pp. 6820-6824, DOI: 10.1109/ICASSP.2019.8682897.

[11] J. Gu, T. S. Yang, J. C. Ye, and D. H. Yang, "CycleGAN denoising of extreme low-dose cardiac CT using wavelet-assisted noise disentanglement," Medical image analysis, vol. 74, p. 102209, 2021, DOI: 10.1016/j.media.2021.102209.

[12] Z. Dai, L. Zhao, K. Wang, and Y. Zhou, "Mode standardization: A practical countermeasure against mode collapse of GAN-based signal synthesis," Appl. Soft Comput., vol. 150, p. 111089, 2024, DOI: 10.1016/j.asoc.2023.111089.

[13] J. Deng, G. Pang, Z. Zhang, Z. Pang, H. Yang, and G. Yang, "cGAN based facial expression recognition for human-robot interaction," IEEE Access, vol. 7, pp. 9848-9859, 2019, DOI: 10.1109/ACCESS.2019.2891668.

[14] X. Hu, Y. Cao, T. Tang, and Y. Sun, "Data-driven technology of fault diagnosis in railway point machines: review and challenges," Transportation Safety and Environment, vol. 4, no. 4, 2022, doi: 10.1093/tse/tdac036.

[15] L. Cao, "Ai in finance: challenges, techniques, and opportunities," ACM Computing Surveys (CSUR), vol. 55, no. 3, pp. 1-38, 2022, DOI: 10.1145/3502289.

[16] Y. Liu and B. Pan, "Profit estimation model and financial risk prediction combining multi-scale convolutional feature extractor and BGRU model," Informatica, vol. 48, no. 11, 2024, DOI: 10.31449/inf.v48i11.5941.

[17] R. K. Dash, T. N. Nguyen, K. Cengiz, and A. Sharma, "Fine-tuned support vector regression model for stock predictions," Neural Computing and Applications, vol. 35, no. 32, pp. 23295-23309, 2023, DOI: 10.1007/s00521-021-05842-w.

[18] J. C. Tellez Gaytan et al., "AI‐based prediction of capital structure: Performance comparison of ANN SVM and LR models," Computational intelligence and neuroscience, vol. 2022, no. 1, p. 8334927, 2022, DOI: 10.1155/2022/8334927.

[19] S. Ahmed, M. M. Alshater, A. El Ammari, and H. Hammami, "Artificial intelligence and machine learning in finance: A bibliometric review," Research in International Business and Finance, vol. 61, p. 101646, 2022, DOI: 10.1016/j.ribaf.2022.101646.

[20] T. Šestanović and T. Kalinić Milićević, "Identification of the Optimal Neural Network Architecture for Prediction of Bitcoin Return," Informatica, vol. 36, no. 1, pp. 175-196, 2025, DOI: 10.15388/24-INFOR561.

[21] C. Milana and A. Ashta, "Artificial intelligence techniques in finance and financial markets: a survey of the literature," Strategic Change, vol. 30, no. 3, pp. 189-209, 2021, DOI: 10.1002/jsc.2403.

[22] S. Joshi, "Review of Gen AI models for financial risk management," International Journal of Scientific Research in Computer Science, Engineering and Information Technology, vol. 11, no. 1, pp. 709-723, 2025, DOI: 10.32628/CSEIT2511114.

[23] H. Pallathadka, E. H. Ramirez-Asis, T. P. Loli-Poma, K. Kaliyaperumal, R. J. M. Ventayen, and M. Naved, "Applications of artificial intelligence in business management, e-commerce and finance," Materials Today: Proceedings, vol. 80, pp. 2610-2613, 2023, DOI: 10.1016/j.matpr.2021.06.419.

[24] D. Mohanty, A. K. Parida, and S. S. Khuntia, "Financial market prediction under deep learning framework using auto encoder and kernel extreme learning machine," Appl. Soft Comput., vol. 99, p. 106898, 2021, DOI: 10.1016/j.asoc.2020.106898.

[25] Y. Cheng, G. Li, X. Zhou, and S. Ye, "Research on Time Series Forecasting Models Based on Hybrid Attention Mechanism and Graph Neural Networks," Informatica, vol. 49, no. 21, 2025, DOI: 10.31449/inf.v49i21.7580..

[26] K. H. Tuhin, A. Nobi, M. H. Rakib, and J. W. Lee, "Long short-term memory autoencoder based network of financial indices," Humanities and Social Sciences Communications, vol. 12, no. 1, p. 100, 2025/01/29 2025, doi: 10.1057/s41599-025-04412-y.

[27] B. Bieganowski and R. Slepaczuk, "Supervised Autoencoder MLP for Financial Time Series Forecasting," arXiv preprint arXiv:2404.01866, 2024.

[28] Z. You, Z. Shi, H. Bo, J. Cartlidge, L. Zhang, and Y. Ge, "DGDNN: Decoupled graph diffusion neural network for stock movement prediction," arXiv preprint arXiv:2401.01846, 2024.

[29] Q. Zhang, Y. Zhang, X. Yao, S. Li, C. Zhang, and P. Liu, "A Dynamic Attributes-driven Graph Attention Network Modeling on Behavioral Finance for Stock Prediction," ACM Trans. Knowl. Discov. Data, vol. 18, no. 1, p. Article 16, 2023, doi: 10.1145/3611311.

[30] X. Hou, K. Wang, C. Zhong, and Z. Wei, "ST-Trader: A Spatial-Temporal Deep Neural Network for Modeling Stock Market Movement," IEEE/CAA Journal of Automatica Sinica, vol. 8, pp. 1015-1024, 05/01 2021, doi: 10.1109/JAS.2021.1003976.

[31] Kraken, "59% of Crypto Investors Use Dollar-Cost Averaging as Their Primary Investment Strategy," 2024. [Online]. Available: https://www.kraken.com/features/dollar-cost-averaging.

[32] S. Miao, M. Liu, and P. Li, "Interpretable and generalizable graph learning via stochastic attention mechanism," in International Conference on Machine Learning, 2022: PMLR, pp. 15524-15543, DOI: 10.48550/arXiv.2201.12987.

[33] Z. Liu, T.-K. Nguyen, and Y. Fang, "Tail-gnn: Tail-node graph neural networks," in Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, 2021, pp. 1109-1119, DOI: 10.1145/3447548.3467276.

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

[35] P. N. Achyutha, S. Chaudhury, S. C. Bose, R. Kler, J. Surve, and K. Kaliyaperumal, "User Classification and Stock Market - Based Recommendation Engine Based on Machine Learning and Twitter Analysis," Mathematical Problems in Engineering, vol. 2022, no. 1, p. 4644855, 2022, DOI: 10.1155/2022/4644855.

[36] R. Feng, S. Jiang, X. Liang, and M. Xia, "STGAT: Spatial–Temporal Graph Attention Neural Network for Stock Prediction," Applied Sciences, vol. 15, no. 8, p. 4315, 2025, DOI: 10.3390/app15084315.