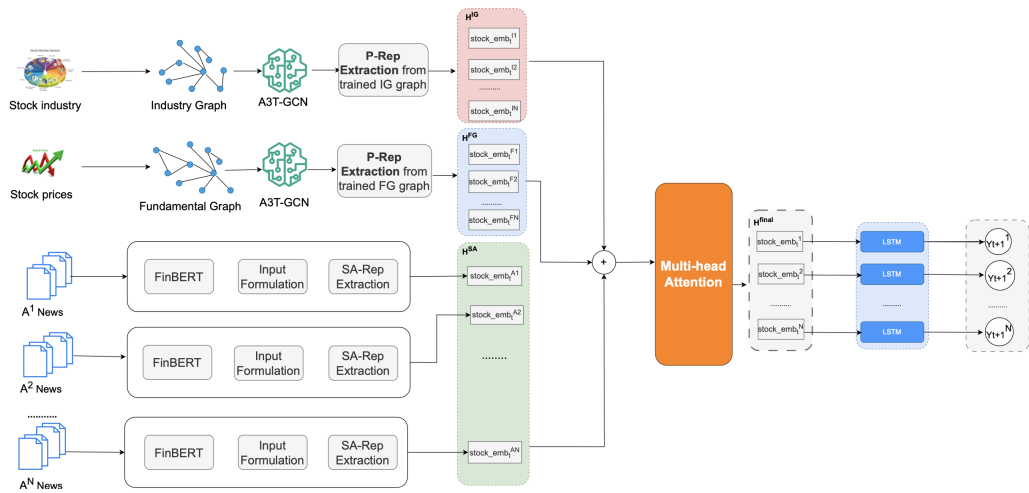
Graphical Abstract

GASF-Net: A Graph-Aware Sentiment Fusion Framework for Stock Forecasting

Nguyen Thi Thu, Le Ky Anh, Nguyen Khac Thai Binh, Vilayvanh Kenmany, Nguyen Nhat Hai



Overview of the proposed GASF-Net framework, which fuses graph-based structural and sentiment-aware representations for stock price forecasting.

Highlights

GASF-Net: A Graph-Aware Sentiment Fusion Framework for Stock Forecasting

Nguyen Thi Thu, Le Ky Anh, Nguyen Khac Thai Binh, Vilayvanh Kenmany, Nguyen Nhat Hai

- Propose **GASF-Net**, a graph-aware sentiment fusion framework for stock forecasting.
- Encode structural signals via **A3T-GCN**-based graph representations.
- Extract semantic sentiment vectors from financial news using **FinBERT**.
- Integrate heterogeneous modalities via **Multi-Head Attention** fusion.
- Demonstrate superior performance over baselines on S&P 500 stocks.

GASF-Net: A Graph-Aware Sentiment Fusion Framework for Stock Forecasting

Nguyen Thi Thu[a], Le Ky Anh[a], Nguyen Khac Thai Binh[a], Vilayvanh Kenmany[a], Nguyen Nhat Hai[a,*]

[a]*Hanoi University of Science and Technology, Hanoi, Vietnam*

Abstract

Forecasting stock prices is a challenging task due to the heterogeneous and asynchronous nature of financial signals, such as structural interdependencies among companies and sentiment conveyed through textual news. Existing approaches often struggle to effectively integrate such diverse data sources into a unified representation. In this work, we propose GASF-Net (Graph-Aware Sentiment Fusion Network), a novel deep learning framework that jointly models structural and sentiment information for robust stock price prediction. GASF-Net consists of two parallel encoders: a graph-based structural encoder, which employs an Attention-Augmented Adaptive Temporal Graph Convolutional Network (A3T-GCN) to capture relational dynamics across stocks, and a sentiment-aware encoder, which uses FinBERT to extract semantic and emotional features from financial news. These representations are integrated using a Multi-Head Attention (MHA) fusion module, enabling adaptive cross-modal learning. The fused embedding is then passed to a temporal forecasting module to predict future stock movements. We evaluate GASF-Net on 12 major S&P 500 stocks over a four-year period. Results show that GASF-Net outperforms a wide range of baselines, highlighting the effectiveness of modality-aware fusion. This study demonstrates the potential of combining structural graph features and sentiment representations in financial time series forecasting.

*Corresponding author

Email address: `hai.nguyenhat@hust.edu.vn` (Nguyen Nhat Hai)

URL: `https://scholar.google.com/citations?user=zmtZGk0AAAAJ&hl=vi` (Nguyen Nhat Hai)

1. Introduction

Predicting stock market fluctuations remains a complex yet vital challenge due to the noisy, nonlinear, and dynamic characteristics of financial markets. Stock prices are influenced by a diverse array of signals—including historical prices, company fundamentals, inter-stock relationships, and investor sentiment—each operating on different time scales and modalities. Capturing and fusing these heterogeneous, asynchronous sources in a unified and context-aware manner is a core difficulty that traditional methods often fail to address.

Classical statistical models such as ARIMA [1] and GARCH [2] are effective in modeling trend and volatility but struggle with nonlinear, cross-sectional dependencies. Machine learning models like SVMs [3], RFs [4], and gradient boosting [5] add flexibility but do not explicitly model relational or temporal structures. Deep learning techniques, particularly RNNs [6] and LSTMs [7, 8], improve time-series learning yet typically treat financial signals in isolation and ignore structured interdependencies between companies.

To overcome these limitations, graph-based learning has gained traction due to its strength in representing structured relationships. Techniques such as GNNs [9], RGNNs [10], Con-GNNs [11], and GATs [12] have achieved success in various domains, including social networks [13], recommendation systems [14], and bioinformatics [15]. In the financial domain, modeling stocks as nodes and defining edges based on industry sectors or pricing similarities allows for richer structural encoding.

However, integrating structured graph information with unstructured, semantic sentiment from financial news remains underexplored. Many existing models rely on static fusion schemes or neglect the adaptive interplay between structure and sentiment. To address this challenge, we propose GASF-Net—a unified framework that fuses graph-aware price representations and sentiment embeddings through an attention-driven mechanism.

Specifically, we introduce dual structural encoders that extract price-based representations (P-Rep) from two complementary graphs: a fundamental graph (FG) and an industry graph (IG), using A3T-GCN [16]. In parallel, we construct sentiment-aware representations (SA-Rep) by processing financial news through FinBERT [17], followed by a dimensionality reduction module. These representations capture both the relational and emotional context affecting market movement.

To unify these multi-modal features, we introduce a Multi-Head Attention

(MHA) module [18], which learns adaptive weights across modalities. Unlike static concatenation, MHA enables the model to dynamically emphasize more informative signals. The fused representation is then passed through temporal LSTM predictors for future stock price estimation.

Our contributions are summarized as follows:

- We propose **GASF-Net**, a novel graph-aware sentiment fusion network that integrates dual graph encoders with FinBERT-based sentiment modeling via a trainable attention mechanism.
- We construct dual price-based representations (P-Rep) from both fundamental and industry graphs using A3T-GCN, and a sentiment representation (SA-Rep) from FinBERT-encoded news embeddings.
- We adopt Multi-Head Attention to fuse modalities, enabling adaptive, context-aware integration. Experimental results suggest 16 heads yield optimal performance.
- We perform sensitivity analysis to explore the effect of historical and sentiment window sizes, identifying $(T_p, T_n) = (20, 1)$ as optimal.
- Extensive experiments on 12 S&P 500 stocks over four years show that GASF-Net consistently outperforms baselines in accuracy, robustness, and generalization.

The remainder of the paper is organized as follows. Section 2 reviews relevant literature. Section 3 details the proposed framework. Section 4 outlines experimental settings. Section 5 presents results and analysis. Section 6 concludes the study.

2. Related Works

This section reviews key developments in stock price prediction, covering traditional and deep learning models, graph-based approaches, sentiment-driven forecasting, and multimodal fusion strategies. We emphasize how these works motivate the design of our proposed GASF-Net framework.

2.1. Traditional and Deep Learning Models for Stock Price Prediction

Early models for stock forecasting were based on statistical time series techniques such as ARIMA and GARCH, which capture linear dependencies and volatility clustering but struggle with the inherent non-linearity of financial markets [1, 2]. To address non-linear patterns, machine learning models including SVMs [3], Random Forests [4], and Gradient Boosting [5] have been applied, though they rely heavily on feature engineering and do not explicitly model temporal dependencies.

Deep learning models such as LSTM [7, 8] and BiLSTM [6] provide improved sequence modeling by learning latent temporal dynamics. More recent work by Zhang et al. [19] combined CNNs with BiLSTM and attention mechanisms to capture both spatial and sequential features. However, these architectures remain unimodal, typically using only price sequences and ignoring external information such as market sentiment or inter-stock relationships—factors that are critical in real-world financial decision-making.

2.2. Graph-Based Learning in Financial Forecasting

Graph-based learning has emerged as a promising direction for modeling the structural dependencies between stocks. GNNs and their variants [9, 10, 11, 12] are effective in representing stock relationships through graphs constructed based on industry sectors, pricing correlations, or multi-view knowledge. For example, MGAR [20] integrates multiple graphs including industry, wiki-based knowledge, and price similarity, while Melody-GCN [21] incorporates multimodal inputs within a spatio-temporal GCN design.

Despite their strength in structural modeling, these approaches often neglect unstructured sources like textual sentiment, or they adopt static fusion strategies that lack adaptability. Fu and Zhang [22] combined sentiment and price information but used simple concatenation, which limits the model's ability to capture cross-modal interactions. In contrast, our proposed GASF-Net employs Multi-Head Attention to enable adaptive, modality-aware fusion of heterogeneous inputs, enhancing expressiveness and learning capacity.

2.3. Sentiment Analysis and NLP-Based Forecasting

The integration of Natural Language Processing (NLP) into financial forecasting has grown significantly, supported by the emergence of pretrained language models (PLMs). FinBERT [17] has been widely adopted to extract sentiment from financial news, as in Kim et al. [23], who paired FinBERT with LSTM to forecast SP 500 stocks. However, their method treated text features independently from structured market signals.

Other works have explored attention-based sentiment modeling. For instance, Cui et al. [24] and Zhang et al. [25] proposed attention-enhanced GCNs and bimodal fusion networks, respectively, for sentiment classification. These models, however, are not tailored for financial time series forecasting and do not incorporate graph-based structural reasoning.

Our framework addresses this gap by constructing sentiment-aware embeddings (SA-Rep) from FinBERT outputs and fusing them with graph-based price representations (P-Rep) using an MHA-based fusion module. This design enables joint modeling of structure and sentiment, a crucial feature for robust financial forecasting.

2.4. Multimodal Fusion and Feature Integration

Multimodal learning aims to integrate heterogeneous data sources to improve model generalization and robustness. Many existing studies apply basic fusion strategies like feature concatenation or pooling (e.g., max, average), which can overlook deeper dependencies between modalities [26]. Huang et al. [27] combined FinBERT sentiment vectors with LSTM-based price models, yet did not account for stock relationships or adaptive fusion mechanisms.

Our proposed GASF-Net distinguishes itself through three innovations: (i) a modular dual-representation architecture that encodes both structural (P-Rep) and semantic (SA-Rep) signals; (ii) an attention-based fusion module using Multi-Head Attention to integrate modalities adaptively; and (iii) a temporal sensitivity design that tunes historical windows for both price and news signals. Unlike prior work that applies attention within a single stream, GASF-Net applies attention across modalities, resulting in a more robust and context-aware forecasting system.

Vit li phn method sau cho logic vi introduction và related works. Gi tt c format và công thc theo format latex, ch thay i content cho phù hp:

3. Method

This section provides a detailed of problem definition and overview of our proposed framework, including the necessary mathematical formulations and notations.

3.1. Problem Definition

Given a dataset comprising a set of N stocks, we formulate the stock price prediction problem as a *regression task*, where the objective is to forecast the future price of each stock at time $t+1$ based on its current state at time t. Formally, the prediction for stock i is defined as:

$$Y_{t+1}^{i} = F(X_t^i), \tag{1}$$

where X_t^i denotes the input feature vector for stock i at time t, and Y_{t+1}^i is the corresponding predicted price at time $t+1$.

Collectively, the feature matrix $X_t = [x_t^1, x_t^2, \ldots, x_t^N] \in \mathbb{R}^{N\times M}$ and label matrix $Y_t = [y_t^1, y_t^2, \ldots, y_t^N] \in \mathbb{R}^{N\times 1}$ represent the stock trading attributes and their associated future prices, respectively, over a continuous time series. Here, M denotes the dimensionality of each input feature vector x_t^i, and y_t^i is the corresponding scalar label for stock i, where $i \in \{1, 2, \ldots, N\}$.

In this study, we focus on enhancing the input representation X_t by incorporating two critical aspects of the stock market: (1) *inter-stock relationships*, captured through price-based representations, and (2) *sentiment-based representation*, which reflect market perceptions (positive, negative, or neutral) at time t.

3.2. Proposed Framework

The proposed framework comprises four major components: (1) multi-relational graph construction, (2) extraction of graph-based and sentiment-based representations (P-Reps and SA-Reps), (3) feature integration via a Multi-Head Attention (MHA) module, and (4) temporal modeling using a sequence learning network. The overall architecture is illustrated in Figure 1.

Our framework is designed to integrate structured inter-stock dependencies and semantic cues from financial news into a unified and adaptive representation for stock forecasting. It leverages both domain-informed graph structures and pretrained language models to extract heterogeneous features, which are fused by attention mechanisms to enhance prediction robustness and interpretability.

3.2.1. Multi-Relational Graph Construction

To represent inter-stock relationships, we construct two complementary graphs that encode both sector-level and behavioral similarities:

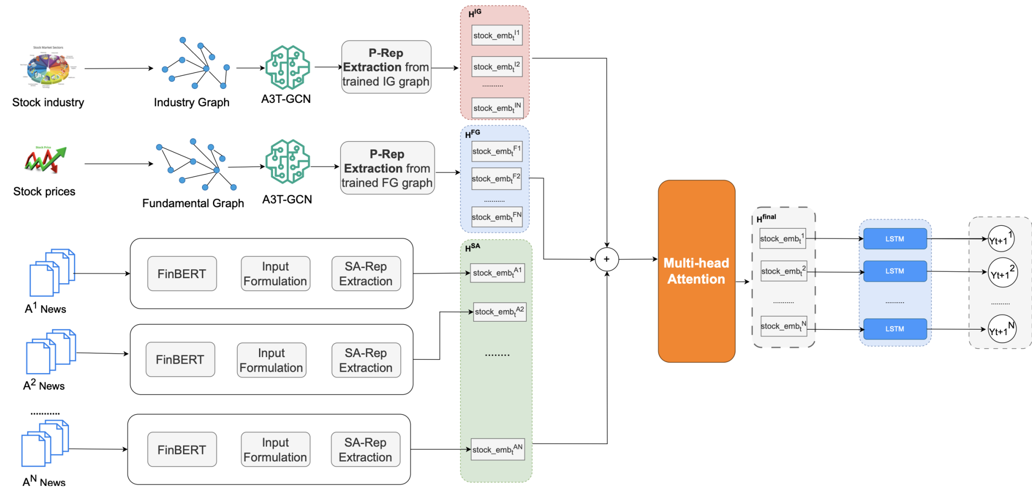


Figure 1: The overall architecture of GASF-Net. The model integrates graph-based structural features from two sources—industry graph and fundamental graph—via A3T-GCN encoders to produce P-Reps, and extracts sentiment-aware representations (SA-Reps) from financial news using FinBERT. All representations are adaptively fused through a Multi-Head Attention module and then passed through LSTM-based forecasting modules to predict future stock prices.

- **Industry Graph (IG):** Captures static connections among companies based on shared industry classification. If two companies belong to the same industry or are registered in the same state, an edge is established:

$$e_{ij}^{IG} = \begin{cases} 1 & \text{if same industry or region} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

- **Fundamental Graph (FG):** Encodes dynamic similarities in stock movement using cosine similarity of historical return vectors. Let $n^i \in \mathbb{R}^{T_p}$ be the return sequence of stock i, where each return is defined as:

$$r_t^i = \frac{close_t^i - close_{t-1}^i}{close_{t-1}^i}$$

Then the edge weight is defined as:

$$e_{ij}^{FG} = \left|\cos(n^i, n^j)\right| = \left|\frac{\langle n^i, n^j\rangle}{\|n^i\| \cdot \|n^j\|}\right| \tag{3}$$

Each graph $G^g = (V, E^g)$, where $g \in \{\text{IG}, \text{FG}\}$, is associated with node features n^i and adjacency matrix $[e^g_{ij}] \in \mathbb{R}^{N\times N}$, where N is the number of stocks. The dual-graph formulation enables the model to exploit both categorical and price-behavior relationships.

3.2.2. P-Reps and SA-Rep Extraction

To obtain price-based stock representations (P-Reps), we apply the Attention-Augmented Adaptive Temporal Graph Convolutional Network (A3T-GCN) [16] to both IG and FG. A3T-GCN jointly captures temporal trends using gated recurrent units (GRUs) [28] and spatial dependencies using graph convolutional layers [29]. It outputs node embeddings for each graph:

$$H^g = [h^{g,1}, h^{g,2}, \ldots, h^{g,N}] \in \mathbb{R}^{N\times M} \tag{4}$$

where $h^{g,i} \in \mathbb{R}^M$ is the embedding of stock i from graph $g \in \{\text{IG}, \text{FG}\}$, and $M = 32$ is the embedding dimension. These two embeddings serve as complementary views of stock behavior.

For sentiment-based representation (SA-Rep), we process daily financial news associated with each stock. Let $A^s_t = [a^s_1, ..., a^s_{Z_t}]$ be a set of news articles for stock s on day t, where Z_t is the number of articles. Each article a^s_z is passed through FinBERT to yield a scalar sentiment score:

$$sa^s_z = \text{Prob}_{pos}(a^s_z) - \text{Prob}_{neg}(a^s_z) \tag{5}$$

The set of scores forms a raw sentiment vector $SA^s_t = [sa^s_1, ..., sa^s_{Z_t}]$. This variable-length sequence is then compressed into a 5-dimensional feature vector using basic statistics:

$$\text{Stat}^s_t = [\min, \max, \text{mean}, \sigma, Z_t]$$

To incorporate temporal dynamics, we define a sliding window of length T_n to aggregate sentiment vectors over the past T_n days. The aggregated input for stock s becomes:

$$I^s_t \in \mathbb{R}^{T_n\times 5}$$

This matrix is flattened and passed through a fully connected layer to produce a fixed-length embedding:

$$h^{SA,s} \in \mathbb{R}^{32}, \quad \text{for each stock } s$$

Collectively, we obtain a sentiment-based representation matrix $H^{SA} \in \mathbb{R}^{N\times 32}$, which is later fused with the P-Reps from IG and FG using a Multi-Head Attention mechanism, as described in Section 3.2.3.

3.2.3. Feature Fusion via Multi-Head Attention

Since the proposed model constructs three distinct types of embeddings—two graph-based (P-Rep from IG and FG) and one sentiment-based (SA-Rep)—it is essential to integrate them in a manner that captures their complementary contributions. Each of these embeddings encodes stock information from a different perspective: sectoral structure, behavioral correlation, and sentiment dynamics. Direct concatenation or simple pooling would fail to model the intricate dependencies and relevance between them.

To address this, we introduce a Multi-Head Attention (MHA) mechanism, which allows the model to adaptively learn both the importance and interaction of each embedding stream. Unlike single-head attention such as the traditional Self-Attention mechanism [30], MHA employs multiple parallel attention heads, each focusing on different subspaces of the input features. This approach improves the model's expressiveness while maintaining computational efficiency.

Let $H = [H^{IG}, H^{FG}, H^{SA}] \in \mathbb{R}^{3\times N\times M}$ denote the concatenation of the three embedding matrices, where each $H^* \in \mathbb{R}^{N\times M}$ corresponds to a modality-specific representation (industry, fundamental, sentiment), and N is the number of stocks.

We then linearly project H into the query, key, and value spaces for each attention head using learned parameter matrices:

$$Q = HW_Q \tag{6}$$

$$K = HW_K \tag{7}$$

$$V = HW_V \tag{8}$$

Here, $W_Q, W_K, W_V \in \mathbb{R}^{M\times d_k}$, and $d_k = M/n_{heads}$ is the dimensionality per head. For each attention head $h \in \{1, \ldots, n_{heads}\}$, we compute the scaled dot-product attention as:

$$\text{head}_h = \text{softmax}\left(\frac{Q_h K_h^\top}{\sqrt{d_k}}\right) V_h \tag{9}$$

All attention heads are then concatenated and linearly transformed to obtain the fused representation:

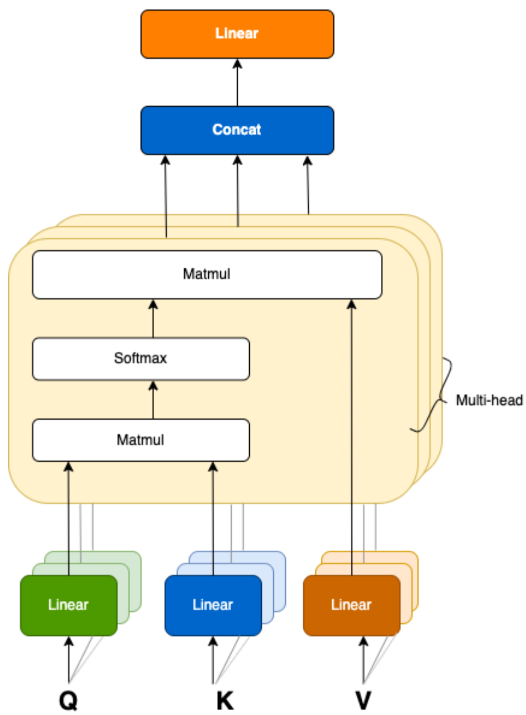


Figure 2: Architecture of the Multi-Head Attention fusion module.

$$H^{final} = \text{Concat}(\text{head}_1, \ldots, \text{head}_{n_{heads}})W_O \tag{10}$$

where $W_O \in \mathbb{R}^{(n_{heads} \cdot d_k) \times M}$ is the output projection matrix.

This fusion layer enables the model to capture both intra-modality and inter-modality interactions dynamically. The attention weights reflect the relevance of each modality to the prediction task, allowing the model to suppress irrelevant signals while enhancing critical features. The resulting fused embedding $H^{final} \in \mathbb{R}^{N \times M}$ is then passed to the next stage for temporal modeling and prediction.

3.2.4. Sequence Modeling

After the heterogeneous features from the industry graph (IG), fundamental graph (FG), and sentiment embeddings (SA-Rep) are fused through the Multi-Head Attention mechanism, the resulting unified representation $H^{final} \in \mathbb{R}^{N \times M}$ encodes rich multi-modal information for each stock. However, to fully leverage the temporal dynamics of market behavior, it is necessary to model sequential dependencies over time.

Conventional feedforward neural networks are inadequate for this task, as they lack mechanisms to retain temporal context. Recurrent Neural Networks (RNNs) [31] were developed to address this by maintaining hidden states across time steps, enabling the model to learn from historical sequences. Nevertheless, RNNs are known to suffer from vanishing gradients, making them ineffective for modeling long-range dependencies.

To overcome these limitations, Long Short-Term Memory (LSTM) networks [32] introduce gated memory units that selectively retain, update, or discard information over time. These gates allow the network to preserve relevant information from earlier time steps, making LSTM particularly suitable for financial forecasting, where market behavior may be influenced by events or trends occurring days or weeks earlier.

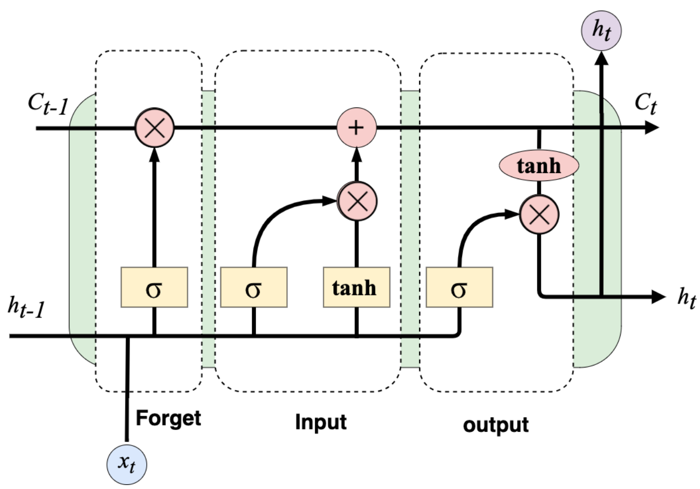


Figure 3: Overview of the LSTM architecture.

Each LSTM cell includes the following components:

- **Forget Gate:**

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{11}$$

Determines which information from the previous cell state C_{t-1} should be forgotten.

- **Input Gate:**

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{12}$$
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{13}$$

Controls how much new information is added to the memory cell.

- **Output Gate:**

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{14}$$
$$h_t = o_t \cdot \tanh(C_t) \tag{15}$$

Determines what information is output from the current memory cell.

The cell state C_t is updated by combining the retained memory and new information:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{16}$$

This gated mechanism equips LSTM with the ability to model long-term dependencies across time, which is critical for accurately forecasting stock price movements based on past trends, sentiment, and inter-company relationships.

3.2.5. Training Process

The complete training process of the proposed GASF-Net framework is summarized in Algorithm 1. The model takes as input historical stock price sequences and news articles, processes them into graph-based and sentiment-based representations, fuses them via Multi-Head Attention, and performs forecasting using LSTM.

4. Experiments

In this section, we will introduce the overview of dataset, the evaluation metrics, the comparison models, the parameters setting.

To conduct experiments on a large-scale stock dataset, we use the public dataset named Financial News and Stock Price Integration Dataset (FNSPID)[33] define and discuss several error metrics commonly used to evaluate the performance of stock price prediction models. The dataset comprises 29.7 million stock prices and 15.7 million time-aligned financial news records for 4,775

Algorithm 1 Training Procedure for GASF-Net

Input: Historical prices P_t, news articles A_t;
Output: Predictions $\hat{Y}_{t+1}$

Construct Industry Graph (IG) and Fundamental Graph (FG)
for each graph $G^g \in \{IG, FG\}$ **do**
    Compute stock return vectors for all stocks
    Construct adjacency matrix E^g based on similarity metric
    Apply A3T-GCN on G^g to get $H^g \in \mathbb{R}^{N \times M}$
end for
Retrieve financial news A_t^i for each stock i
for each article in A_t^i **do**
    Encode sentiment using FinBERT $\rightarrow$ score sa_z^i
end for
Compute 5-dim statistics for each stock and aggregate over sliding window
Pass statistics through fully connected layer $\rightarrow H^{SA} \in \mathbb{R}^{N \times 32}$
Fuse H^{IG}, H^{FG}, H^{SA} via Multi-Head Attention $\rightarrow H^{final}$
for each stock $i = 1$ to N **do**
    Predict next price $\hat{y}_{t+1}^i = \text{LSTM}(H_i^{final})$
end for
Compute MSE loss: $\mathcal{L} = \text{MSE}(\hat{Y}_{t+1}, Y_{t+1})$
Update parameters: $\theta \leftarrow \theta - \eta \cdot \nabla_\theta \mathcal{L}$ =0

S&P500 companies, covering the period from 1999 to 2023, sourced from 4 stock market news websites.

In this study, we extracted the dataset from 01-Jan-2020 to 31-Dec-2023. Among the periods, the training dataset, validation dataset, and test dataset are defined in Table 2.

4.1. Evaluation Metrics

The **Mean Squared Error (MSE)** is defined as the average of the squared differences between the predicted values ($\hat{y}_i$) and the actual values (y_i). This metric evaluates the difference between the ground truth stock price compared to the predicted prices for each day.

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

Table 1: Description of Dataset

Dataset	Time Period	Number of Days	Number of News Articles
Training	01/Jan/2020 – 31/Dec/2022	1008	172,784
Validation	01/Jan/2023 – 30/Jun/2023	125	73,416
Test	01/Jul/2023 – 31/Dec/2023	125	83,814

The **Mean Absolute Error (MAE)** calculates the average of the absolute differences between the predicted and actual stock prices:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i|$$

where:

- y_i is the actual stock price at time i,
- $\hat{y}_i$ is the predicted stock price at time i.

4.2. Baseline Selection and Parameters

In this chapter, we presented the baseline models for the stock price prediction compared to our proposal.

- **LSTM+CNN** [34]: A deep learning architecture that integrates multiple pipelines of Convolutional Neural Networks (CNNs) with Bidirectional Long Short-Term Memory (BiLSTM) units to capture both spatial and temporal features in stock price data.
- **Multi-GCGRU** [35]: A hybrid model that combines Graph Convolutional Networks (GCNs) and Gated Recurrent Units (GRUs) to exploit structural relationships among stocks while modeling temporal dependencies.
- **Sentiment+LSTM** [36]: An LSTM-based approach that incorporates sentiment features extracted from financial news along with historical price data to enhance stock trend prediction.
- **MGAR** [37]: A framework that applies traditional fusion strategies to integrate embeddings derived from four distinct graph structures, each representing a different type of inter-stock relation.

4.3. Primary Settings

We implemented the proposed framework using the PyTorch library. The detailed configurations of the model's hyperparameters are presented in Table 2. For all baseline models, we retained the original hyperparameter settings as reported in their respective publications to ensure a fair comparison.

Table 2: Framework Configurations

Configuration	**Value**
Dimension of H^{SA}, H^{IG}, H^{FG}	32
Dimension of Node Features	32
Dimension of Egde Features	1
Number A3TGCN layers	2
Hidden size of LSTM	[32, 16]

5. Results and Analysis

5.1. Prediction Accuracy Across Different Forecast Horizons

Table 3 presents the quantitative results of 1-day, 2-day, and 3-day stock price forecasting across several competitive baselines and ablated variants of our framework. Evaluation is conducted using two widely adopted regression metrics: Mean Squared Error (MSE) and Mean Absolute Error (MAE). Lower values on these metrics indicate better predictive accuracy.

Table 3: Comparison of stock price prediction performance over different numbers of trading days. Bold values indicate the best performance.

Model	**1-day**		**2-day**		**3-day**	
	MSE	MAE	MSE	MAE	MSE	MAE
LSTM+CNN [34]	1.9×10^{-3}	3.0×10^{-2}	2.0×10^{-3}	3.3×10^{-2}	2.0×10^{-3}	3.6×10^{-2}
Multi-GCGRU [35]	3.1×10^{-2}	1.1×10^{-1}	4.2×10^{-2}	1.3×10^{-1}	4.8×10^{-2}	1.4×10^{-1}
Sentiment+LSTM [36]	7.2×10^{-3}	3.2×10^{-2}	1.3×10^{-2}	4.5×10^{-2}	1.7×10^{-2}	5.3×10^{-2}
MGAR [37]	4.5×10^{-3}	1.6×10^{-2}	6.5×10^{-3}	2.1×10^{-2}	1.1×10^{-2}	2.3×10^{-2}
Our (IG only)	7.3×10^{-4}	2.1×10^{-2}	1.6×10^{-3}	3.2×10^{-2}	2.4×10^{-3}	3.9×10^{-2}
Our (FG only)	3.4×10^{-3}	4.3×10^{-2}	3.0×10^{-3}	3.9×10^{-2}	2.7×10^{-3}	4.1×10^{-2}
Our (IG + FG)	4.6×10^{-4}	1.8×10^{-2}	9.5×10^{-4}	2.6×10^{-2}	1.4×10^{-3}	2.8×10^{-2}
GASF-Net	$\mathbf{3.8 \times 10^{-4}}$	$\mathbf{1.5 \times 10^{-2}}$	$\mathbf{8.2 \times 10^{-4}}$	$\mathbf{2.1 \times 10^{-2}}$	$\mathbf{1.1 \times 10^{-3}}$	$\mathbf{2.4 \times 10^{-2}}$

Across all horizons, our proposed GASF-Net framework consistently outperforms the baselines. For 1-day prediction, it achieves an MSE of 3.8×10^{-4}, corresponding to relative reductions of 91.6% against MGAR [37], 94.7% against Sentiment+LSTM [36], and 80.0% against LSTM+CNN [34]. MAE improvements follow a similar trend, confirming the robustness of our method in capturing both magnitude and directional accuracy.

Compared to our own ablations, GASF-Net demonstrates substantial performance gains over price-only models. Specifically, the full model reduces MSE by 17.0% compared to IG+FG fusion, and by 88.8% compared to FG-only. This highlights the strong complementarity between the sentiment-based representation (SA-Rep) and graph-based price representations (P-Reps).

In longer-term forecasts (3-day), GASF-Net maintains its superiority with an MAE of 2.4×10^{-2}, compared to 3.9×10^{-2} for IG-only and 4.1×10^{-2} for FG-only, showing relative improvements of 38.5% and 41.5%, respectively. These results underscore the importance of integrating sentiment signals alongside structural information, particularly in volatile or delayed market responses.

The observed gains can be attributed to the combination of three key innovations: (i) multi-relational graph learning (IG + FG), (ii) context-aware sentiment encoding using FinBERT, and (iii) adaptive fusion via Multi-Head Attention. Collectively, these components enable the model to interpret technical, relational, and emotional aspects of the market in a unified, end-to-end trainable architecture.

In summary, the proposed GASF-Net framework sets a new state-of-the-art for multi-horizon stock price forecasting by jointly modeling multi-source market signals with high precision and adaptability.

5.2. Ablation Study: Component-Wise Effectiveness

To assess the contribution of each component in our GASF-Net framework, we conduct an ablation study by systematically removing or altering individual modules and evaluating the performance degradation. All models are trained and evaluated under identical settings using 1-day forecasting on 12 representative S&P 500 stocks. The results are reported in Table 4 and visualized in Figure 4.

From Table 4, we observe that removing the sentiment-based representation (SA-Rep) leads to a noticeable degradation in performance: MSE increases by 21.1% and MAE by 20.0%. This highlights the complementary nature of textual sentiment in modeling short-term stock fluctuations. Additionally, excluding either the industry graph (IG) or fundamental graph

Table 4: Ablation study results (1-day prediction).

Configuration	**MSE**	**MAE**
w/o SA-Rep	4.6×10^{-4}	1.8×10^{-2}
w/o IG	6.1×10^{-4}	2.2×10^{-2}
w/o FG	7.3×10^{-4}	2.1×10^{-2}
Only SA-Rep	1.4×10^{-3}	2.9×10^{-2}
Only IG+FG (P-Reps only)	4.6×10^{-4}	1.8×10^{-2}
P-Reps + SA-Rep (w/o MHA, concat)	5.5×10^{-4}	1.9×10^{-2}
GASF-Net	$\mathbf{3.8 \times 10^{-4}}$	$\mathbf{1.5 \times 10^{-2}}$

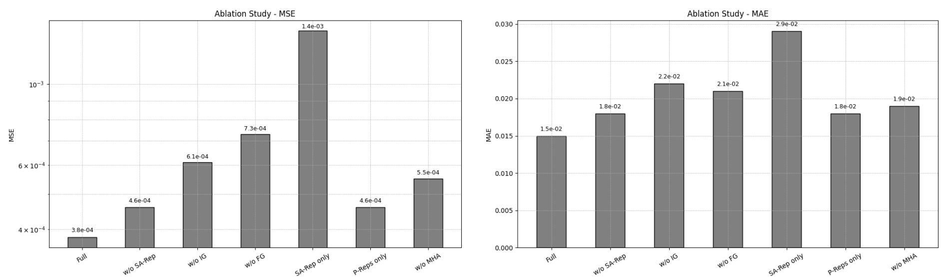


Figure 4: Impact of each component on 1-day forecasting (lower MSE and MAE are better).

(FG) also deteriorates accuracy, underscoring the value of multi-relational structural information.

When replacing the Multi-Head Attention (MHA) fusion with simple concatenation, performance declines by 44.7% in MSE compared to the full model, revealing the necessity of adaptive, modality-aware feature integration.

Interestingly, using SA-Rep alone yields significantly worse results than graph-based models, suggesting that sentiment alone is insufficient. However, when used in combination, it adds a non-trivial and orthogonal signal that improves prediction substantially.

These findings collectively validate the design choices of our framework and demonstrate the synergy between graph-based and sentiment-based features when fused adaptively.

5.3. Effect of Sentiment Window Size on SA-Rep Quality

To analyze the sensitivity of the model to the temporal scope of sentiment aggregation, we vary the window size T_n, which denotes the number of past trading days used to construct the sentiment-based representation (SA-Rep). We evaluate performance across five representative technology stocks: GOOG, AAPL, NVDA, META, and MSFT.

Table 5 reports the average MSE and MAE over these stocks for different values of T_n, and Figure 5 visualizes the results to illustrate trends more clearly.

Table 5: Effect of sentiment window size T_n on prediction performance (average over five stocks).

T_n	1	3	5	10	20
MSE	3.2×10^{-4}	3.5×10^{-4}	4.7×10^{-4}	8.3×10^{-4}	1.1×10^{-3}
MAE	1.3×10^{-2}	1.4×10^{-2}	1.6×10^{-2}	2.1×10^{-2}	2.6×10^{-2}

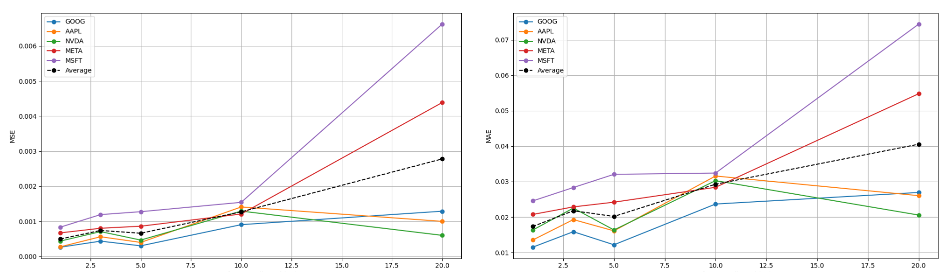


Figure 5: Model performance with varying sentiment aggregation window T_n across five selected stocks.

Both the tabular and visualized results consistently show that increasing T_n leads to performance degradation. The best results are obtained at $T_n = 1$, confirming that short-term sentiment is more predictive for stock price forecasting. Larger windows ($T_n = 10$ or $T_n = 20$ introduce outdated or irrelevant sentiment signals that obscure actionable market reactions.

Interestingly, a modest window such as $T_n = 3$ occasionally achieves comparable performance by smoothing noisy sentiment, but further expansion generally leads to diminishing returns or overfitting. This finding reinforces the necessity of using a compact and responsive window for constructing effective sentiment-based representations.

5.4. Effect of the Number of Attention Heads and Fusion Strategies

To evaluate the role of attention-based fusion mechanisms in our framework, we systematically examine the impact of (i) varying the number of attention heads in the Multi-Head Attention (MHA) module and (ii) comparing it to alternative static fusion strategies. Specifically, we consider:

- **Self-Attention**: a single-head attention mechanism that lacks the ability to learn diverse relational perspectives.
- **MHA-n**: multi-head attention with $n \in \{2, 4, 8, 16\}$ heads to allow distributed representation learning across different subspaces.
- **Mean-Pooling**: uniform averaging across feature representations.
- **Max-Pooling**: selection of dominant features without contextual adaptivity.

Table 6: Comparison of fusion strategies and number of attention heads (averaged over 12 S&P 500 stocks).

Fusion Strategy	MSE	MAE
Mean-Pool	9.1×10^{-4}	2.3×10^{-2}
Max-Pool	1.0×10^{-3}	2.5×10^{-2}
Self-Attn	1.3×10^{-3}	2.7×10^{-2}
MHA-2	8.2×10^{-4}	2.2×10^{-2}
MHA-4	5.4×10^{-4}	1.9×10^{-2}
MHA-8	4.2×10^{-4}	1.6×10^{-2}
MHA-16	$\mathbf{3.8 \times 10^{-4}}$	$\mathbf{1.5 \times 10^{-2}}$

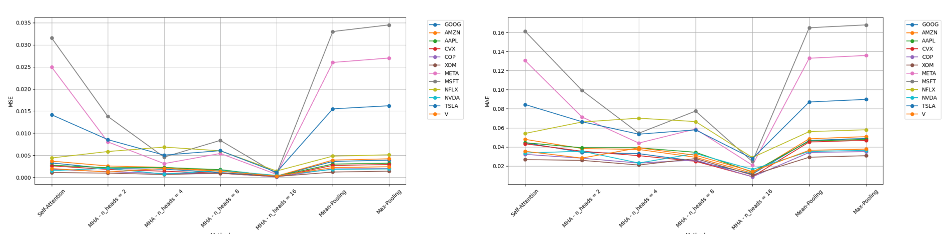


Figure 6: Performance comparison across attention-based and pooling-based fusion strategies.

As shown in Table 6 and Figure 6, the number of attention heads significantly influences the quality of feature fusion. The MHA configuration with 16 heads consistently yields the best performance, achieving an MSE of 3.8×10^{-4} and an MAE of 1.5×10^{-2}, substantially outperforming Self-Attention and static pooling baselines.

Notably, increasing the number of heads from 2 to 16 progressively reduces the error, indicating that the model benefits from attending to multiple subspaces in parallel. This richer representation allows the network to better model complex interactions between price-based and sentiment-based features. While improvements taper off beyond 8 heads, MHA-16 still provides marginal gains, suggesting its effectiveness in capturing fine-grained cross-modal dependencies.

In contrast, static fusion approaches—Mean-Pooling and Max-Pooling—are markedly less effective. These methods apply uniform or fixed aggregation, which cannot adaptively emphasize contextually important signals. For instance, Max-Pooling achieves an MSE of 1.0×10^{-3}, which is more than twice the error of MHA-16. Similarly, Mean-Pooling performs better than Max-Pooling but still lags behind even MHA-2.

These findings reinforce that:

- Learnable fusion strategies significantly outperform fixed ones in modeling heterogeneous financial features.
- The use of multiple attention heads provides complementary views of data, enabling more accurate and robust predictions.
- MHA-16 strikes the best balance between model complexity and predictive accuracy in our GASF-Net framework.

6. Conclusion

In this study, we introduced **GASF-Net**, a novel and unified framework for stock price forecasting that seamlessly integrates heterogeneous financial information—namely, inter-stock relational structures and sentiment signals—through graph-based representation learning and adaptive attention-based fusion. The framework incorporates two core representation modules: (i) *P-Reps*, which capture multi-relational structural dependencies using an Attention-Augmented Temporal Graph Convolutional Network (A3T-GCN); and (ii) *SA-Rep*, which encodes sentiment dynamics from financial news using FinBERT and a temporal aggregation mechanism.

These complementary views are fused using a Multi-Head Attention (MHA) module, which enables the model to adaptively weight modality-specific features and capture fine-grained cross-modal interactions. Through extensive experiments on 12 SP 500 stocks over a five-year period, GASF-Net consistently demonstrates superior performance in both short-term and multi-day forecasting tasks, outperforming a wide range of existing baselines across standard metrics such as MSE and MAE.

Ablation and sensitivity analyses further validate the contribution of each component. In particular, we find that shorter sentiment windows ($T_n = 1$ or 3) are optimal for immediate trend capture, while employing a larger number of attention heads (e.g., MHA-16) leads to improved representation fusion and overall accuracy. Comparisons with traditional fusion techniques (e.g., mean and max pooling) underscore the benefits of trainable, dynamic integration in enhancing robustness and expressiveness.

Looking ahead, future research will explore the integration of *graph diffusion mechanisms* to model long-range dependencies and richer relational paths, as well as the incorporation of *large language models (LLMs)* to enhance sentiment and event understanding at the document and narrative levels.

References

[1] M. Khashei, M. Bijari, G. A. Raissi Ardali, Improvement of autoregressive integrated moving average models using fuzzy logic and artificial neural networks (anns), Neurocomput. 72 (4–6) (2009) 956–967. doi:10.1016/j.neucom.2008.04.017.
URL `https://doi.org/10.1016/j.neucom.2008.04.017`

[2] T. Bollerslev, Generalized autoregressive conditional heteroskedasticity, Journal of Econometrics 31 (3) (1986) 307–327. doi:https://doi.org/10.1016/0304-4076(86)90063-1.

[3] Y. Lin, H. Guo, J. Hu, An svm-based approach for stock market trend prediction, in: The 2013 International Joint Conference on Neural Networks (IJCNN), 2013, pp. 1–7. doi:10.1109/IJCNN.2013.6706743.

[4] L. Khaidem, S. Saha, S. R. Dey, Predicting the direction of stock market prices using random forest, arXiv preprint arXiv:1605.00003 (2016).

[5] S. S. Roy, R. Chopra, K. C. Lee, C. Spampinato, B. Mohammadi-ivatlood, Random forest, gradient boosted machines and deep neural network for stock price forecasting: a comparative analysis on south korean companies, International Journal of Ad Hoc and Ubiquitous Computing 33 (1) (2020) 62–71.

[6] Y. Zhu, Stock price prediction using the rnn model, in: Journal of Physics: Conference Series, Vol. 1650, IOP Publishing, 2020, p. 032103.

[7] S. Selvin, R. Vinayakumar, E. Gopalakrishnan, V. K. Menon, K. Soman, Stock price prediction using lstm, rnn and cnn-sliding window model, in: 2017 international conference on advances in computing, communications and informatics (icacci), IEEE, 2017, pp. 1643–1647.

[8] K. Pawar, R. S. Jalem, V. Tiwari, Stock market price prediction using lstm rnn, in: Emerging Trends in Expert Applications and Security: Proceedings of ICETEAS 2018, Springer, 2019, pp. 493–503.

[9] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, S. Y. Philip, A comprehensive survey on graph neural networks, IEEE transactions on neural networks and learning systems 32 (1) (2020) 4–24.

[10] L. Fan, Q. Cheng, W. Yu, D. Griffith, N. Golmie, Survey of graph neural networks and applications, Wireless Communications & Mobile Computing (Online) 2022 (2022).

[11] S. Zhang, H. Tong, J. Xu, R. Maciejewski, Graph convolutional networks: a comprehensive review, Computational Social Networks 6 (1) (2019) 1–23.

[12] C. Sun, C. Li, X. Lin, T. Zheng, F. Meng, X. Rui, Z. Wang, Attention-based graph neural networks: a survey, Artificial Intelligence Review 56 (Suppl 2) (2023) 2263–2310.

[13] J. Z. Pan, E. Edelstein, P. Bansky, A. Wyner, A knowledge graph based approach to social science surveys, Data Intelligence 3 (4) (2021) 477–506.

[14] S. Wu, F. Sun, W. Zhang, X. Xie, B. Cui, Graph neural networks in recommender systems: a survey, ACM Computing Surveys 55 (5) (2022) 1–37.

[15] Z. Gao, C. Jiang, J. Zhang, X. Jiang, L. Li, P. Zhao, H. Yang, Y. Huang, J. Li, Hierarchical graph learning for protein–protein interaction, Nature Communications 14 (1) (2023) 1093.

[16] J. Zhu, Y. Song, L. Zhao, H. Li, A3t-gcn: Attention temporal graph convolutional network for traffic forecasting (2020). arXiv:2006.11583.
URL https://arxiv.org/abs/2006.11583

[17] D. Araci, Finbert: Financial sentiment analysis with pre-trained language models (2019). arXiv:1908.10063.
URL https://arxiv.org/abs/1908.10063

[18] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, I. Polosukhin, Attention is all you need (2023). arXiv:1706.03762.
URL https://arxiv.org/abs/1706.03762

[19] J. Zhang, L. Ye, Y. Lai, Stock price prediction using cnn-bilstm-attention model, Mathematics 11 (9) (2023) 1985. doi:10.3390/math11091985.

[20] G. Song, T. Zhao, S. Wang, H. Wang, X. Li, Stock ranking prediction using a graph aggregation network based on stock price and stock relationship information, Information Sciences 643 (2023) 119236. doi:https://doi.org/10.1016/j.ins.2023.119236.

[21] R. Liu, H. Liu, H. Huang, B. Song, Q. Wu, Multimodal multiscale dynamic graph convolution networks for stock price prediction, Pattern Recognition 149 (2024) 110211.

[22] K. Fu, Y. Zhang, Incorporating multi-source market sentiment and price data for stock price prediction, Mathematics 12 (10) (2024) 1572. doi:10.3390/math12101572.

[23] J. Kim, H.-S. Kim, S.-Y. Choi, Forecasting the s&p 500 index using mathematical-based sentiment analysis and deep learning models: A finbert transformer model and lstm, Axioms 12 (9) (2023) 835. doi:10.3390/axioms12090835.

[24] X. Cui, W. Tao, X. Cui, Affective-knowledge-enhanced graph convolutional networks for aspect-based sentiment analysis with multi-head attention, Applied Sciences 13 (7) (2023) 4458. doi:10.3390/app13074458.

[25] R. Zhang, C. Xue, Q. Qi, L. Lin, J. Zhang, L. Zhang, Bimodal fusion network with multi-head attention for multimodal sentiment analysis, Applied Sciences 13 (3) (2023) 1915. doi:10.3390/app13031915.

[26] F. Bieder, R. Sandkühler, P. C. Cattin, Comparison of methods generalizing max- and average-pooling (2021). arXiv:2103.01746.
URL `https://arxiv.org/abs/2103.01746`

[27] Y. Huang, J. Lin, X. Sun, Z. Ding, Finbert-stock: A sentiment-aware model for stock movement prediction using financial news, Information Processing & Management 59 (6) (2022) 103083.

[28] J. Chung, C. Gulcehre, K. Cho, Y. Bengio, Empirical evaluation of gated recurrent neural networks on sequence modeling (2014). arXiv:1412.3555.
URL `https://arxiv.org/abs/1412.3555`

[29] T. N. Kipf, M. Welling, Semi-supervised classification with graph convolutional networks (2017). arXiv:1609.02907.
URL `https://arxiv.org/abs/1609.02907`

[30] A. Vaswani, Attention is all you need, Advances in Neural Information Processing Systems (2017).

[31] A. Sherstinsky, Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network, Physica D: Nonlinear Phenomena 404 (2020) 132306. doi:10.1016/j.physd.2019.132306.
URL `http://dx.doi.org/10.1016/j.physd.2019.132306`

[32] J. Schmidhuber, S. Hochreiter, et al., Long short-term memory, Neural Comput 9 (8) (1997) 1735–1780.

[33] Z. Dong, X. Fan, Z. Peng, Fnspid: A comprehensive financial news dataset in time series (2024). arXiv:2402.06698.
URL `https://arxiv.org/abs/2402.06698`

[34] J. Eapen, D. Bein, A. Verma, Novel deep learning model with cnn and bi-directional lstm for improved stock market index prediction, in: 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC), 2019, pp. 0264–0270. doi:10.1109/CCWC.2019.8666592.

[35] J. Ye, J. Zhao, K. Ye, C. Xu, Multi-graph convolutional network for relationship-driven stock movement prediction, in: 2020 25th International Conference on Pattern Recognition (ICPR), IEEE, 2021, p. 6702–6709. doi:10.1109/icpr48806.2021.9412695.
URL `http://dx.doi.org/10.1109/ICPR48806.2021.9412695`

[36] Z. Jin, Y. Yang, Y. Liu, Stock closing price prediction based on sentiment analysis and lstm, Neural Computing and Applications 32 (2020) 9713–9729.

[37] D. Cao, Y. Wang, J. Duan, C. Zhang, X. Zhu, C. Huang, Y. Tong, B. Xu, J. Bai, J. Tong, Q. Zhang, Spectral temporal graph neural network for multivariate time-series forecasting, in: H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, H. Lin (Eds.), Advances in Neural Information Processing Systems, Vol. 33, Curran Associates, Inc., 2020, pp. 17766–17778.