MEHGT-LKG: MULTIMODAL EDGE-ENHANCED HET-EROGENEOUS GRAPH TRANSFORMER WITH LLM-DRIVEN KNOWLEDGE GRAPH FOR STOCK TREND PRE-DICTION

Anonymous authors

000

001

002

004

006

008

009

010 011 012

013

015

016

017

018

019

021

024

025

026

027

028

029

031

032

033

034

037

040

041

042 043

044

046

047

048

051

052

Paper under double-blind review

ABSTRACT

Stock trend prediction plays a central role in optimal investment decision-making, and has attracted extensive research from both investors and institutions. Although recent studies have employed graph structures to model the complex relationships among financial entities, the corresponding models fail to efficiently capture semantically rich edge features across heterogeneous entities, thereby limiting the ability to fuse and align multimodal data such as market indicators, financial events, and heterogeneous graph structures. Therefore, in this paper, we propose a Multimodal Edge-Enhanced Heterogeneous Graph Transformer with LLMdriven Knowledge Graphs (MEHGT-LKG) for stock trend prediction. Specifically, we first fine-tune a large language model (LLM) by using instruction tuning datasets to design a financial event-centric knowledge extraction agent (FinEX). Subsequently, we encode the structured tuples generated from FinEX into financial event-centric knowledge graphs (FEKGs) and then construct multimodal heterogeneous graphs by incorporating multimodal information. Finally, we design a Multimodal Edge-Enhanced Heterogeneous Graph Transformer (MEHGT) to fully encode a series of semantically enriched multimodal heterogeneous graphs spanning different time horizons. MEHGT models edge-level features through type-specific encoders and integrates them into both multi-head attention and message propagation, significantly enriching the representation of relational semantics and target nodes. Extensive experimental results and trading simulations on multiple real-world datasets demonstrate the superior performance of the proposed approach beyond other state-of-the-art models.

1 Introduction

The stock market is a core component of the financial system, providing capital allocation functions for enterprises and investment opportunities for individuals. However, due to its high volatility, complex influencing factors, and nonlinear dynamics, forecasting stock trends remains a challenging and important research area.

Stock price fluctuations are typically driven by two major types of factors: intrinsic market signals and external shocks from related financial entities. The former includes trading behaviors and technical indicators; the latter involves financial events, company announcements, and policy changes. To better improve prediction accuracy and provide more informed investment decisions, integrating multimodal information and capturing complex market dynamics has become a key research direction Cheng et al. (2022); Liu et al. (2024b); Sheng et al. (2024); Huang et al. (2024).

Among various multimodal representation approaches, knowledge graphs have attracted increasing attention due to their ability to structurally represent relationships among financial events Zhao et al. (2022); Wang et al. (2023). However, traditional knowledge graph construction methods rely on predefined schemas and use deep learning models to extract entities and relations from financial texts. These approaches are heavily dependent on fixed rules and templates, making it difficult to capture the diverse and complex expression of financial events. Notably, with advances of Natural

 Language Processing (NLP), large language models (LLMs) have shown exceptional abilities in semantic understanding, knowledge reasoning, and information extraction in many fields Wang et al. (2024); Li & Sanna Passino (2024). Therefore, fine-tuning LLMs to extract key financial events and structured tuples is crucial for constructing accurate and reliable financial event-driven graphs (FEKGs).

To further learn expressive representations and model complex relationships from knowledge graphs, graph-based learning has been increasingly applied to stock trend prediction owing to its ability to capture the complex dependencies among entities Hsu et al. (2021); Li et al. (2024); Liao et al. (2024). Heterogeneous Graph Neural Networks (HGNNs) as advanced variants of Graph Neural Networks (GNNs), distinguish between node and edge types, enabling deeper modeling of complex financial interactions. However, existing HGNNs often overlook encoding edge features, which carry rich semantic information about financial events and capital flows Zhang et al. (2023); Ma et al. (2024); Liu et al. (2024a). Thus, we propose a multimodal edge-enhanced Heterogeneous Graph Transformer (MEHGT) that incorporates structured financial event tuples as edge features into both attention calculation and message passing, enabling more effective integration of multimodal data and improving stock trend prediction.

In summary, to address these issues, we propose a Multimodal Edge-Enhanced Heterogeneous Graph Transformer with LLM-driven Knowledge Graphs (MEHGT-LKG) for stock trend prediction. Specifically, we first fine-tune LLM by using instruction datasets to design a financial event-centric knowledge extraction agent (FinEX). Subsequently, we leverage FinEX to construct FEKGs and build multimodal heterogeneous graphs by incorporating multimodal information with sliding windows. Finally, we design MEHGT to encode edge features via type-specific encoders into attention and message passing, enhancing relational semantics and target node representations within multimodal heterogeneous graphs. The contributions are summarized as follows:

- We design the FinEX agent by finetuning LLM with instruction-based dataset. It generates financial events and structured tuples from financial texts accurately, supporting the automatic construction of FEKGs.
- Based on FEKGs, we construct multimodal heterogeneous graphs within a sliding time window by integrating trading data, market indicators, and other relevant information.
- We propose MEHGT, which explicitly incorporates edge-level features into both the attention computation and message passing process. By modeling financial relations and actions, the model leverages multimodal data to capture entity associations and information flow patterns, thereby enhancing stock trend prediction.
- To verify the superiority of the proposed model, extensive experiments are conducted with the state-of-the-art baselines on multiple real stock datasets.

2 PROBLEM DEFINITION

We formulate the problem of predicting the trend of stocks for excess return as a classification task. The objective of this research is to leverage constructed multimodal financial heterogeneous graphs during w days to predict the rise or fall of the target stock at trading day t+1 (w denotes the actual time window). We represent these dynamic heterogeneous graphs as $G_{1:T} = \{G_1, G_2, \ldots, G_T\}$. Each graph $G_t = (V_t, E_t, R_t)$ at time t contains five types of nodes: $V_t = \{V_t^{KS} \cup V_t^{OE} \cup V_t^{HK} \cup V_t^F \cup V_t^L\}$ and seven types of edges $E_t = \{E_t^{Corr} \cup E_t^{HI} \cup E_t^{Long} \cup E_t^{Short} \cup E_t^{SRE} \cup E_t^{ERS} \cup E_t^{KS} \cup E_t^{COE}\}$. Meanwhile, R_t denotes seven types of relations corresponding to edges. We have $(u \overset{\leftarrow}{\to} v, u \overset{\rightarrow}{\to} v) \in E_t$, where $t \in R_t$ is the relation type and $t \in V_t$. The input features of graph during time window includes node attributes $t \in V_t$ and edge attributes $t \in V_t$. The node attributes $t \in V_t$ of the node set $t \in V_t$, and the edge attributes $t \in V_t$ of the edge set $t \in V_t$, the node attributes to forecast the trend of a key stock in the next time point $t \in V_t$, using the proposed MEHGT model (denoted as $t \in V_t$):

$$\hat{Y}_{t+1}(s) = f_{\theta}(G_t, X_t^V, X_t^E), \tag{1}$$

3 THE PROPOSED METHOD

The architecture of the proposed methodology, MEHGT-LKG is shown in Figure 1. It comprises three main stages: fine-tuning LLM for knowledge extraction, multimodal heterogeneous graphs construction, and designing MEHGT for graph learning and stock trend prediction.

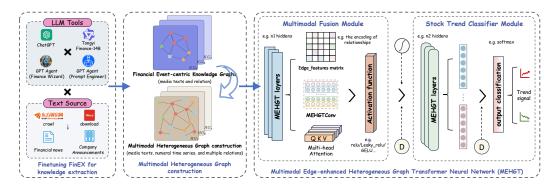


Figure 1: Graphical illustration of the proposed methodology MEHGT-LKG.

3.1 Designing FineX Agent for Knowledge Extraction

To extract financial events and structured tuples, we design an LLM Agent, namely FinEX.

High-quality instruction datasets are critical for LLM-based information extraction, yet remain scarce in the financial domain. To address this, we construct an instruction dataset (in Figure 2) by collecting financial news and company announcements, and use GPT tools to extract structured financial events. Detailedly, guided by optimized prompts, ChatGPT-4 serves as a financial analyst, generating key events and structured tuples. These outputs are further refined by the Finance Wizard Agent based on domain knowledge and validated by experts to ensure accuracy and completeness. Notably, the dataset preserves both triplets, such as \langle Kunlun Tech CO.,LTD — plans to acquire — YOOZOO GAMES CO.,LTD \rangle and event pairs like \langle CATL CO.,LTD — experiences a severe explosion \rangle allowing flexible representations for different event structures. By capturing multi-entity relations and single-entity events, this dual-format design improves semantic precision and completeness in extraction.

Building constructed on the instruction-based dataset, fine-tune Owen model to design FinEX agent. An overview of the fine-tuning process is shown in Figure 2. Each training sample includes both the event description and its corresponding structured tuples, which effectively reduces hallucination during large model reasoning and improves the reliability of outputs. We choose Tongyi-Finance-14B (TF-14B), a domain-specific variant of Qwen-14B pre-trained on extensive financial corpora, as the base model Bai et al. (2023). Fine-tuning is performed with LoRA in the Llama-Factory framework Zheng et al. (2024), updating a small subset of parameters while keeping the backbone frozen to greatly reduce

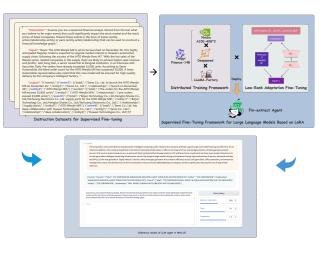


Figure 2: The procedure of fine-tuning LLM to build the FinEX Agent.

memory and computation costs. And training is performed with the DeepSpeed framework on NVIDIA A100 GPUs, ensuring efficient handling of long instruction texts. Finally, FinEX supports

accurate, large-scale extraction of key financial events and structured tuples from raw news and announcements, providing essential inputs for constructing the financial event-centric knowledge graphs (FEKGs).

3.2 Multimodal Heterogeneous Graphs Construction

Financial event-centric knowledge graphs are the foundation to build multimodal Heterogeneous Graphs. We leverage the FinEX Agent to process financial news and announcements from 2021 to 2024 about selected stocks to construct FEKG. The structured tuples include entities such as companies, products, and key individuals, along with relations such as mergers, investments, and short-selling events. To ensure data quality, entities and relationships were standardized to avoid duplication caused by inconsistent naming.

Multimodal heterogeneous graphs extend conventional heterogeneous graphs by incorporating multimodal sources. Additional node types, such

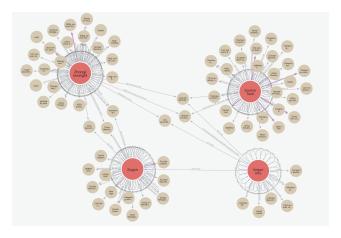


Figure 3: The procedure of fine-tuning LLM to build the FinEX Agent.

as Northbound Trading, margin financing, and securities lending, are introduced to represent key financial activities. Edge types are enriched with stock co-movement correlations and capital flow information to reflect long-short market dynamics. Finally, a temporal sequence of multimodal heterogeneous graphs is generated via sliding time windows. These graphs effectively capture intricate financial interactions across modalities, laying a strong foundation for multimodal fusion, graph learning, and trend prediction. The demo of the subgraph developed with Neo4j is shown in Figure 3

3.3 MEHGT FOR STOCK TREND PREDICTION

We design MEHGT to better capture relational patterns and information dynamics in financial markets. MEHGT explicitly integrates edge-level features into both the attention computation and message passing process, and further conducts a more comprehensive and effective multimodal fusion across both nodes and edges during graph representation learning. The overall architecture is illustrated in Figure. 4.

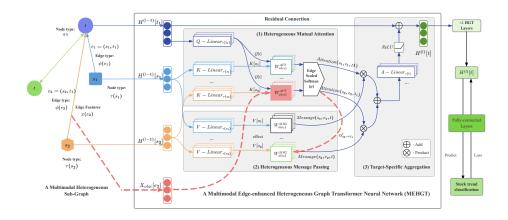


Figure 4: Overview of the MEHGT framework.

Heterogeneous mutual attention: Given a target node t_1 of type $\tau(t_1)$ (i.e., key stock) and a source node $s \in N(t_1)$ in a heterogeneous subgraph, we compute their attention via a multi-head mechanism. To better capture relation semantics, edge features are incorporated through a type-specific transformation and scaling function. Specifically, each target node t_1 and its neighbor s are transformed into a Query vector and a Key vector, respectively.

$$Q_i(t_1) = Q\text{-}Linear_{\tau(t_1)}\left(H^{(l-1)}[t_1]\right),$$
 (2)

$$K_i(s_1) = K\text{-}Linear_{\phi(s_1)}\left(H^{(l-1)}[s_1]\right),$$
 (3)

$$K_i(s_2) = K\text{-}Linear_{\phi(s_2)}\left(H^{(l-1)}[s_2]\right),$$
 (4)

where $Q\text{-}Linear_{\tau(t_1)}$, $K\text{-}Linear_{\tau(s_1)}$, and $K\text{-}Linear_{\tau(s_2)}$ represent the linear projection functions for the Query and Key vectors. $H^{(l)}$ denotes the node embedding of the l-th layer, with $H^{(0)}$ being the initial node embedding.

The similarity between $Q_i(t_1)$ and $K_i(s)$ is calculated as the attention weight between them:

$$ATT-head^{i}(t_{1}, e_{1}, s_{2}) = \left(K_{i}(s_{2})W_{\phi(e_{2})}^{ATT}Q_{i}(t_{1})^{T}\right) \cdot \frac{\mu(\tau(t_{1}), \phi(e_{2}), \tau(s_{2}))}{\sqrt{d}} \cdot f(X_{edge}[e_{2}]) \quad (5)$$

where $W_{\phi(e_2)}^{ATT}$ is the edge-based transformation matrix for capturing the semantic information of different types of edges between t_1 and s_2 , and $f(X_{\text{edge}}[e_2])$ represents how the edge feature X_{edge} influences the attention score. The edge feature $X_{\text{edge}}[e_1]$ is used to scale the attention weight with the specific edge type, which helps refine the attention mechanism by considering financial relationships and actions or short events from graphs. The representation of $ATT - head^i(t_x, e_y, s_z)$ is similar to the above.

Finally, the attention vectors between each node pair are obtained by concatenating the h attention heads. Then, for each target node t, the attention vectors from all its neighboring nodes N(t) are gathered:

$$Attention(s, e, t) = \underset{\forall s \in N(t)}{\operatorname{Softmax}} \left(\underset{i \in [1, h]}{\parallel} ATT\text{-head}^{i}(s, e, t) \right), \tag{6}$$

where $\mathop{\parallel}\limits_{i\in[1,h]}$ is the concatenating function.

Heterogeneous message passing: A message operator is employed to pass messages between various nodes such as stocks, financial entities, and stock markets. The multi-head message is computed by the following process. The source node s_2 is projected into a message vector using a linear transformation:

$$MSG-head^{i}(s_{2}, e_{2}, t_{1}) = \alpha_{s_{2} \to t_{1}}^{i} \cdot \left(M-Linear_{\phi(s_{2})}^{i} H^{(l-1)}(s_{2}) \times W_{\phi(e_{2})}^{MSG}\right), \tag{7}$$

where the function M-Linear $_{\phi(v_s)}^i$ is the linear projection function corresponding to the ith message head, and $W_{\phi(e_{s,t})}^{\mathrm{MSG}}$ is the edge-type transformation matrix. The final step is to concatenate h message headers to get the $Message(s_2,e_2,t_1)$ for each node pair:

$$Message(s_2, e_2, t_1) = \| MSG-head^i(s_2, e_2, t_1).$$
(8)

The representation of MSG-head $^i(s_1,e_1,t_1)$ and $Message(s_1,e_1,t_1)$ is similar to the aforementioned.

Target-specific aggregation: To update embedding of a key stock node t, this module uses multiheaded attention and message passing to refine its representation from neighboring nodes.

$$\widetilde{H}^{(l)}[t] = \oplus_{\forall s \in N(t)} \mathsf{Attention}(s, e, t) \cdot \mathsf{Message}(s, e, t). \tag{9}$$

where $\widetilde{H}^{(l)}[t]$ denotes the updated embedding of the target node, which aggregates the information of all neighboring nodes. The updated embedding of the key stock node is projected to its type-specific distribution:

$$H^{(l)}[t] = \text{A-Linear}_{\phi(t)}\left(\sigma(\widetilde{H}^{(l)}[t])\right) + H^{(l-1)}[t], \tag{10}$$

where A-Linear $_{\phi(t)}$ is a type-aware linear projection, $\sigma(\cdot)$ denotes the nonlinear activation function. The module incorporates the residual structure, where $H^{(l-1)}[t]$ is the node embedding of the target node in the $(l-1)^{th}$ layer.

(4) **Target forecasting network and optimization:** Given the learned representations of target nodes from the MEHGT model, we employ a shallow neural network to predict stock trend. The output is defined as:

$$\hat{Y}_s = \operatorname{softmax}(\operatorname{NN}_f(W_n H^{(l)}[t] + b_n)), \tag{11}$$

where NN_f represents a shallow neural network with two fully connected layers, and $W_n \in \mathbb{R}^{d_s \times 2}$, $b_n \in \mathbb{R}^2$ are the weight matrix and bias, respectively.

The model is trained with cross-entropy loss: d_l is the number of target categories. In this work, we set the $d_l = 2$.

$$\mathcal{L} \mathcal{H}_{\text{target}} = -\sum_{s \in \mathcal{V}_T} \sum_{c=1}^{d_l} \hat{Y}_{t+1}(s) \ln \hat{Y}_{t+1}(s), \tag{12}$$

where $\hat{Y}_{t+1}(s)$ is the ground-truth label of c_{th} price movement category for stock s, which is marked as 1 for the "up" price movements, 0 for the "down" movement, respectively. \mathcal{V}_T denotes the set of target nodes.

Hence, after forecasting networks, MEHGT can effectively leverage the learned node representations from multimodal heterogeneous graph to predict stock trend.

4 EXPERIMENTS

4.1 EXPERIMENT SETTINGS

4.1.1 Datasets.

We select stocks from CSI300 and CSI500 to construct datasets Duan et al. (2025); Zhou et al. (2025), which include AI and renewable energy sectors. Data are collected from Wind (numerical) and Eastmoney (textual), spanning Jan 5, 2021 – Mar 29. We split the datasets into mutually exclusive training/validation/testing sets in the ratio of 7:2:1. Moreover, the original dataset, weights of FinEX, code of MEHGT-LKG, and implementation details will be provided in our GitHub.

4.1.2 Compared Methods.

To show the performance of our proposed model, we compare MEHGT-LKG with SOTA methods. We select the following models as the baseline for comparison: (1) Time series modeling methods: Informer Zhou et al. (2021), TCN Bai et al. (2018), and CNN-LSTM Vidal & Kristjanpoller (2020); (2) Graph-based modeling methods: GAT Velickovic et al. (2018), HGT Hu et al. (2020), MAC Ma et al. (2023), and MDGNN Qian et al. (2024)

4.1.3 EVALUATION METRICS.

Following previous study Zeng et al. (2018); Wadden et al. (2019), we use Precision, Recall, and F1 score for two NLP tasks (events extraction and tuples extraction). And we select Accuracy (ACC), Matthew's Correlation Coefficient (MCC), precision, recall, F1 score and Area Under Curve (AUC) to evaluate stock

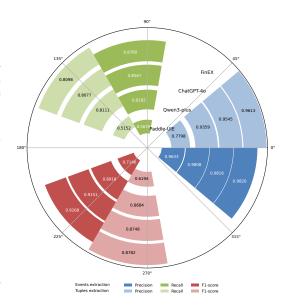


Figure 5: The procedure of fine-tuning LLM to build the FinEX Agent.

trend prediction performance, and choose Cumulative Return Rate (CRR), Maximum Drawdown (MDD), and Sharpe Ratio (Sharpe) to assess the profitability Liu et al. (2024a); Ma et al. (2024).

4.2 Comparison Results of NLP models

We conduct experiments on a manually annotated news dataset to evaluate FinEX agent in event and tuple extraction by comparing it with baselines including Paddle-UIE (0.5B), Qwen3-plus (235B), and ChatGPT-4o (200B). The results are shown in Figure 5.

FinEX outperforms all baselines in both tasks. It reaches precision scores of 0.9820 and 0.9613, respectively, and reaches F1 score of 0.9268 and 0.8782. This may be attributed to fine-tuning the LLM with instruction data, which significantly enhances its capacity to comprehend lengthy financial news and announcements deeply, accurately extract key financial events, and generate structured tuples in JSON format aligned with standard requirements for constructing knowledge graphs. ChatGPT-40 and Qwen3-plus, though competitive, occasionally generate malformed or redundant JSON outputs. Paddle-UIE struggles with long, unstructured financial texts due to its schema-constrained decoding, leading to low performance.

All LLM-based models show strong extraction ability due to their language understanding and generalization. FinEX, with only 14B parameters, offers high performance and fast inference, making it suitable for deployment. We will open-source FinEX with full pipelines.

4.3 MAIN COMPARISON RESULTS

We compare MEHGT-LKG with a range of state-of-the-art baselines, including time-series models and graph-based models. As shown in Table 1, MEHGT-LKG consistently achieves superior performance across all datasets, with notable MCC scores of 0.3718 on Inspur, 0.3681 on IFLYTEK, and 0.3337 on Sungrow.

Compared with time-series methods, our model significantly outperforms them, as these methods focus on single-stock sequences and fail to capture inter-stock dependencies. Among graph-based models, MAC leverages sentiment features via a GCN framework, but has weak expressive power due to simplistic feature aggregation. HGT demonstrates stronger performance by employing type-aware multi-head attention over heterogeneous graphs. MEHGT-LKG builds upon this by further injecting edge-level features into the attention calculation and message-passing processes, enhancing its capacity to model complex financial relations. MDGNN benefits from dynamic multi-relational modeling with Transformers, but lacks multimodal integration, particularly of financial events. Overall, compared with the baselines, our method has the following advantages.

- We design FinEX by finetuning an LLM to accurately extract structured tuples of financial events, enabling the construction of FEKGs with rich domain-specific semantics.
- We construct multimodal heterogeneous graphs within a sliding time window by integrating trading data, market indicators, and other relevant information.
- The MEHGT model explicitly incorporates edge-level features into both the attention computation and the message passing process, enabling deeper multimodal fusion and enhancing stock trend prediction performance.

4.4 MARKET TRADING SIMULATION

To further evaluate the profitability of our method, we conduct an investment simulation. Figure 6 presents the cumulative return curves on six representative stocks during the backtesting. MEHGT-LKG outperforms all baselines, remaining the highest equity curve throughout the trading period under bullish and bearish conditions. Especially, on Zhongji Innolight, it attains a CRR of 274.49% and on Inspur, it attains 104.50%. And MDGNN also exhibits competitive performance.

Detailedly, MEHGT-LKG achieves the highest return among all baseline models, and consistently delivers positive returns throughout the entire backtesting period. Notably, on upward-trending stocks (e.g., Zhongji Innolight), MEHGT-LKG captures strong buy signals driven by accurate trend prediction and achieves a remarkable cumulative return. Meanwhile, for stocks experiencing down-

Table 1: Prediction performance of different methods across selected stock datasets (The results on other datasets are shown in the Appendix).

		Inspur (000977)							CATL (300750)						
Methods	ACC	MCC	Precision	Recall	F1	AUC	ACC	MCC	Precision	Recall	F1	AUC			
Time-series models															
Informer	0.6104	0.2195	0.6154	0.5333	0.5714	0.5894	0.6104	0.2372	0.5357	0.6818	0.6000	0.5808			
TCN	0.5844	0.1669	0.5846	0.5067	0.5429	0.5515	0.5909	0.1800	0.5200	0.5909	0.5532	0.5913			
CNN-LSTM	0.6234	0.2462	0.6349	0.5333	0.5797	0.6645	0.5455	0.1913	0.4828	0.8485	0.6154	0.5248			
Graph-based models															
GAŤ	0.6169	0.2865	0.5690	0.8800	0.6911	0.6019	0.6234	0.2439	0.5541	0.6212	0.5857	0.5754			
HGT	0.6169	0.2630	0.5755	0.8133	0.6740	0.5516	0.6364	0.2737	0.5658	0.6515	0.6056	0.6030			
MAC	0.6039	0.2171	0.5761	0.7067	0.6347	0.5954	0.6234	0.2104	0.5833	0.4242	0.4912	0.5559			
MDGNN	0.6299	0.2605	0.6500	0.5200	0.5778	0.6226	0.6558	0.2960	0.6000	0.5909	0.5954	0.6288			
MEHGT-LKG (ours)	0.6234	0.3718	0.6883	0.7713	0.7184	0.6322	0.6039	0.2748	0.5243	0.8182	0.6391	0.6307			
	IFLYTEK (002230)							EVE (300014)							
Time-series models															
Informer	0.5779	0.1962	0.5288	0.7746	0.6286	0.5280	0.6429	0.2517	0.4769	0.5962	0.5299	0.5850			
TCN	0.5974	0.1883	0.5652	0.5493	0.5571	0.5305	0.5909	0.1922	0.4286	0.6346	0.5116	0.5786			
CNN-LSTM	0.6234	0.2438	0.5890	0.6056	0.5972	0.5843	0.5260	0.2302	0.4054	0.8654	0.5521	0.5754			
Graph-based models															
GAŤ	0.6234	0.2456	0.5867	0.6197	0.6027	0.6129	0.6039	0.2567	0.4471	0.7308	0.5547	0.6655			
HGT	0.6104	0.2544	0.5556	0.7746	0.6471	0.6359	0.6494	0.2858	0.4857	0.6538	0.5574	0.6158			
MAC	0.6234	0.2382	0.6000	0.5493	0.5735	0.5959	0.4545	0.2350	0.3806	0.9808	0.5484	0.5430			
MDGNN	0.6299	0.2562	0.6842	0.5662	0.5771	0.5418	0.6169	0.2383	0.4533	0.6538	0.5354	0.5911			
MEHGT-LKG (ours)	0.6818	0.3681	0.6375	0.7183	0.6755	0.6930	0.6883	0.3166	0.5357	0.5769	0.5556	0.6640			
Zhongji Innolight (300308)								Sungrow (300274)							
Time-series models															
Informer	0.6558	0.3155	0.6857	0.6076	0.6443	0.6297	0.6234	0.2489	0.5584	0.6418	0.5972	0.5891			
TCN	0.6169	0.2369	0.6000	0.7595	0.6704	0.5690	0.5974	0.1810	0.5373	0.5373	0.5373	0.5963			
CNN-LSTM	0.5779	0.1538	0.5795	0.6456	0.6108	0.4928	0.6234	0.2874	0.5474	0.7761	0.6420	0.6056			
Graph-based models															
GAT	0.6104	0.2217	0.5979	0.7342	0.6591	0.6175	0.6494	0.3085	0.5802	0.7015	0.6351	0.6447			
HGT	0.6169	0.3076	0.5758	0.9620	0.7204	0.6046	0.6558	0.2910	0.6207	0.5373	0.5760	0.5656			
MAC	0.5974	0.2342	0.7073	0.3671	0.4833	0.5332	0.6169	0.2144	0.5645	0.5224	0.5426	0.5805			
MDGNN	0.6364	0.2992	0.7255	0.4684	0.5692	0.6117	0.6039	0.2665	0.5294	0.8060	0.6391	0.5656			
MEHGT-LKG (ours)	0.6688	0.3375	0.6591	0.7324	0.6946	0.6485	0.6753	0.3337	0.6393	0.5821	0.6094	0.6523			

ward trends (e.g., CATL), the model generates timely exit signals and flexibly adjusts positions, resulting in a solid and positive return.

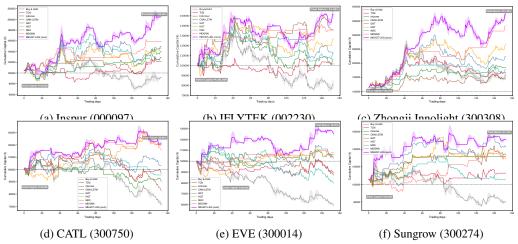


Figure 6: Simulated trading performance of all models during backtesting.

4.5 Ablation Study

In this section, several ablation experiments are performed to examine the effectiveness of each component of MEHGT-LKG.

As shown in Table 2, MEHGT-LKG model achieves the best performance. The method removing financial text data performs worse, verifying that structured tuples extracted from financial texts

enrich graph semantics and enhance trend prediction. Notably, the variant without edge features performs comparably to the one w/o financial text data, as the edge features in MEHGT-LKG are primarily constructed from event-centric financial text. And the variant w/o numerical indicators underperforms, confirming that market indicators and trading features serve as essential signals for stock-level inference.

Table 2: Comparison results of ablation analysis across selected stock datasets (The results on other datasets are shown in the Appendix).

	Inspur (000977)							CATL (300750)						
Methods	ACC	MCC	Precision	Recall	F1	AUC	ACC	MCC	Precision	Recall	F1	AUC		
w/o events	0.6169	0.2328	0.6250	0.5333	0.5755	0.5967	0.6104	0.2728	0.5306	0.7879	0.6341	0.6121		
w/o edge feats	0.6175	0.2570	0.6270	0.5500	0.5860	0.6175	0.6227	0.2744	0.5608	0.7734	0.6502	0.6286		
w/o indicators	0.6364	0.2757	0.6145	0.6800	0.6456	0.6409	0.6104	0.2614	0.5319	0.7576	0.6250	0.5989		
MEHGT-LKG	0.6234	0.3718	0.6883	0.7713	0.7184	0.6322	0.6039	0.2748	0.5243	0.8182	0.6391	0.6307		
IFLYTEK (002230)								EVE (300014)						
w/o events	0.6364	0.2893	0.5843	0.7324	0.6500	0.6302	0.6688	0.2492	0.5102	0.4808	0.4950	0.6110		
w/o edge feats	0.6287	0.2799	0.5724	0.7269	0.6405	0.6186	0.6457	0.2601	0.5187	0.4995	0.5089	0.6222		
w/o indicators	0.6169	0.2546	0.5652	0.7324	0.6380	0.5817	0.6623	0.2595	0.5000	0.5385	0.5185	0.6283		
MEHGT-LKG	0.6818	0.3681	0.6375	0.7183	0.6755	0.6930	0.6883	0.3166	0.5357	0.5769	0.5556	0.6640		
Zhongji Innolight (300308)								Sungrow (300274)						
w/o events	0.6494	0.3136	0.7119	0.5316	0.6087	0.6508	0.6429	0.3307	0.5625	0.8060	0.6626	0.6174		
w/o edge feats	0.6589	0.3351	0.7204	0.5562	0.6277	0.6625	0.6234	0.2873	0.5512	0.7815	0.6465	0.6159		
w/o indicators	0.6299	0.2615	0.6146	0.7468	0.6743	0.6518	0.6299	0.2710	0.5610	0.6866	0.6174	0.6021		
MEHGT-LKG	0.6688	0.3375	0.6591	0.7324	0.6946	0.6485	0.6753	0.3337	0.6393	0.5821	0.6094	0.6523		

4.6 Hyperparameter analysis

We use Sankey diagram to analyze the impact of key layer hyperparameters in MEHGT-LKG (Figure. 7). The analysis covers the hidden dimensions of three MEHGT layers and one Linear layer. Each path denotes a specific hyperparameter combination, with darkness indicating the MCC performance.

The best combination—128–128–64–16—achieves the highest MCC of 0.372, suggesting a

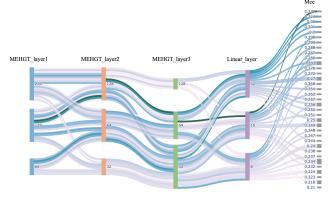


Figure 7: Hyperparameter analysis based on Sankey graph.

balanced architecture with good learning and classification abilities. In contrast, over-parameterized settings (e.g., 256–128–128–16, MCC=0.218) and under-parameterized ones (e.g., 64–32–32–8, MCC=0.240) perform worse, likely due to overfitting and limited representation capacity, respectively. Additionally, configurations with sharp reductions between layers (e.g., 256–32–32–32, 256–32–32–16, and 256–32–32–8) also perform poorly, possibly because abrupt dimensional drops hinder the model's ability to extract high-level features effectively.

5 Conclusion

In this work, we propose MEHGT-LKG for stock trend prediction. By fine-tuning Qwen with custom instruction datasets, we design a financial event-centric knowledge extraction agent (FinEX), and build financial event-centric knowledge graphs. Then, with sliding windows, these graphs are integrated with numerical indicators from multiple sources to form a sequence of multimodal heterogeneous graphs. Finally, we design MEHGT to learn a series of semantically enriched multimodal heterogeneous graphs spanning different time horizons. MEHGT models edge-level features through type-specific encoders and integrates them into both multi-head attention and message propagation, significantly enriching the representation of relational semantics and target nodes. In the future, we plan to integrate MEHGT-LKG with reinforcement learning frameworks to enable end-to-end portfolio-level trading.

REFERENCES

- Jinze Bai, Shuai Bai, Yunfei Chu, Zeyu Cui, Kai Dang, Xiaodong Deng, Yang Fan, Wenbin Ge, Yu Han, Fei Huang, Binyuan Hui, Luo Ji, Mei Li, Junyang Lin, Runji Lin, Dayiheng Liu, Gao Liu, Chengqiang Lu, Keming Lu, Jianxin Ma, Rui Men, Xingzhang Ren, Xuancheng Ren, Chuanqi Tan, Sinan Tan, Jianhong Tu, Peng Wang, Shijie Wang, Wei Wang, Shengguang Wu, Benfeng Xu, Jin Xu, An Yang, Hao Yang, Jian Yang, Shusheng Yang, Yang Yao, Bowen Yu, Hongyi Yuan, Zheng Yuan, Jianwei Zhang, Xingxuan Zhang, Yichang Zhang, Zhenru Zhang, Chang Zhou, Jingren Zhou, Xiaohuan Zhou, and Tianhang Zhu. Qwen technical report. *arXiv preprint arXiv:2309.16609*, 2023.
- Shaojie Bai, J Zico Kolter, and Vladlen Koltun. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv* preprint arXiv:1803.01271, 2018.
- Dawei Cheng, Fangzhou Yang, Sheng Xiang, and Jin Liu. Financial time series forecasting with multi-modality graph neural network. *Pattern Recognition*, 121:108218, 2022.
- Yitong Duan, Weiran Wang, and Jian Li. Factorgcl: A hypergraph-based factor model with temporal residual contrastive learning for stock returns prediction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 173–181, 2025.
- Yi-Ling Hsu, Yu-Che Tsai, and Cheng-Te Li. Fingat: Financial graph attention networks for recommending top-*k* k profitable stocks. *IEEE transactions on knowledge and data engineering*, 35(1): 469–481, 2021.
- Ziniu Hu, Yuxiao Dong, Kuansan Wang, and Yizhou Sun. Heterogeneous graph transformer. In *Proceedings of the web conference 2020*, pp. 2704–2710, 2020.
- Hongbin Huang, Minghua Chen, and Xiao Qiao. Generative learning for financial time series with irregular and scale-invariant patterns. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=CdjnzWsQax.
- Shuqi Li, Yuebo Sun, Yuxin Lin, Xin Gao, Shuo Shang, and Rui Yan. Causalstock: Deep end-to-end causal discovery for news-driven multi-stock movement prediction. *Advances in Neural Information Processing Systems*, 37:47432–47454, 2024.
- Xiaohui Victor Li and Francesco Sanna Passino. Findkg: Dynamic knowledge graphs with large language models for detecting global trends in financial markets. In *Proceedings of the 5th ACM international conference on AI in finance*, pp. 573–581, 2024.
- Sihao Liao, Liang Xie, Yuanchuang Du, Shengshuang Chen, Hongyang Wan, and Haijiao Xu. Stock trend prediction based on dynamic hypergraph spatio-temporal network. *Applied Soft Computing*, 154:111329, 2024.
- Mengpu Liu, Mengying Zhu, Xiuyuan Wang, Guofang Ma, Jianwei Yin, and Xiaolin Zheng. Echogl: Earnings calls-driven heterogeneous graph learning for stock movement prediction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 13972–13980, 2024a.
- Ruirui Liu, Haoxian Liu, Huichou Huang, Bo Song, and Qingyao Wu. Multimodal multiscale dynamic graph convolution networks for stock price prediction. *Pattern Recognition*, 149:110211, 2024b.
- Yu Ma, Rui Mao, Qika Lin, Peng Wu, and Erik Cambria. Multi-source aggregated classification for stock price movement prediction. *Information Fusion*, 91:515–528, 2023.
- Yu Ma, Rui Mao, Qika Lin, Peng Wu, and Erik Cambria. Quantitative stock portfolio optimization by multi-task learning risk and return. *Information Fusion*, 104:102165, 2024.
- Hao Qian, Hongting Zhou, Qian Zhao, Hao Chen, Hongxiang Yao, Jingwei Wang, Ziqi Liu, Fei Yu, Zhiqiang Zhang, and Jun Zhou. Mdgnn: Multi-relational dynamic graph neural network for comprehensive and dynamic stock investment prediction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 14642–14650, 2024.

- Yankai Sheng, Yuanyu Qu, and Ding Ma. Stock price crash prediction based on multimodal data machine learning models. *Finance Research Letters*, 62:105195, 2024.
 - Petar Velickovic, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph attention networks. In 6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 May 3, 2018, Conference Track Proceedings. OpenReview.net, 2018. URL https://openreview.net/forum?id=rJXMpikCZ.
 - Andrés Vidal and Werner Kristjanpoller. Gold volatility prediction using a cnn-lstm approach. *Expert Systems with Applications*, 157:113481, 2020.
 - David Wadden, Ulme Wennberg, Luheng He, and Hannaneh Hajishirzi. Entity, relation, and event extraction with contextualized span representations. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pp. 5784–5789, 2019. URL https://aclanthology.org/D19-1585.
 - Jiapu Wang, Sun Kai, Linhao Luo, Wei Wei, Yongli Hu, Alan Wee-Chung Liew, Shirui Pan, and Baocai Yin. Large language models-guided dynamic adaptation for temporal knowledge graph reasoning. *Advances in Neural Information Processing Systems*, 37:8384–8410, 2024.
 - Ting Wang, Jiale Guo, Yuehui Shan, Yueyao Zhang, Bo Peng, and Zhuang Wu. A knowledge graph–gcn–community detection integrated model for large-scale stock price prediction. *Applied Soft Computing*, 145:110595, 2023.
 - Datong Zeng, Daojian Zeng, Jun He, Kang Liu, and Jun Zhao. Extracting relational facts by an end-to-end neural model with copy mechanism. In *Proceedings of ACL 2018*, pp. 214–220, 2018.
 - Zeyang Zhang, Ziwei Zhang, Xin Wang, Yijian Qin, Zhou Qin, and Wenwu Zhu. Dynamic heterogeneous graph attention neural architecture search. In *Proceedings of the AAAI conference on artificial intelligence*, volume 37, pp. 11307–11315, 2023.
 - Yu Zhao, Huaming Du, Ying Liu, Shaopeng Wei, Xingyan Chen, Fuzhen Zhuang, Qing Li, and Gang Kou. Stock movement prediction based on bi-typed hybrid-relational market knowledge graph via dual attention networks. *IEEE transactions on knowledge and data engineering*, 35(8): 8559–8571, 2022.
 - Yaowei Zheng, Richong Zhang, Junhao Zhang, Yanhan Ye, Zheyan Luo, Zhangchi Feng, and Yongqiang Ma. Llamafactory: Unified efficient fine-tuning of 100+ language models. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 3: System Demonstrations)*, Bangkok, Thailand, 2024. Association for Computational Linguistics. URL http://arxiv.org/abs/2403.13372.
 - Haoyi Zhou, Shanghang Zhang, Jieqi Peng, Shuai Zhang, Jianxin Li, Hui Xiong, and Wancai Zhang. Informer: Beyond efficient transformer for long sequence time-series forecasting. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, pp. 11106–11115, 2021.
 - Quanshi Zhou, Lifen Jia, and Wei Chen. A deep learning-based model for stock price prediction under consideration of financial news publishers. *International Journal of General Systems*, 54 (4):499–539, 2025.