Transactions on Computational and Scientific Methods | Vol. 5, No. 7, 2025
ISSN: 2998-8780
https://pspress.org/index.php/tcsm
Pinnacle Science Press

Frequent Subgraph Mining and Structural Prompt Injection for Large Language Model Fine-Tuning

Zhengxi Xiao[1], Christopher Moretti[2], Qian Li[2]
[1]University of Southern California, Los Angeles, USA
[2]Rensselaer Polytechnic Institute, Troy, USA
*corresponding author: Zhengxi Xiao; zhengxixiao@gmail.com

Abstract: This paper proposes a fine-tuning method for large language models that integrates frequent subgraph mining and structural cue injection, aiming to simultaneously enhance structural inductive ability and decision auditability in structured tasks. The method takes graph data as input, extracts reusable high-order structural patterns through relation-constrained frequent subgraph mining, and constructs a pattern library to characterize structural prototypes that stably appear across samples. Subsequently, the structural prototypes are compiled into readable structural cues, which, together with task instructions, form conditional inputs, enabling the model to explicitly focus on key structural evidence during inference and reduce semantic ambiguity caused by structural information serialization. To achieve efficient and controllable knowledge injection, the fine-tuning stage employs an efficient parameter update strategy, restricting trainable increments to a constrained low-rank subspace, and aligning cue-induced structural anchors with the model's internal representation through structural consistency regularization, thereby suppressing irrelevant directional drift and improving training stability. This framework forms a closed loop between structure discovery, structure representation, and structure injection, enabling end-to-end training and deployment without relying on complex external engineering modifications, and is suitable for graph classification and related structured reasoning scenarios requiring structural evidence.

Keywords: Structural hint injection; frequent subgraph pattern; efficient parameter fine-tuning; structural consistency regularization.

1. Introduction

With the rapid growth of large-scale graph data in scenarios such as financial risk control, cybersecurity, knowledge management, and industrial operations, extracting generalizable knowledge from complex relational structures and serving downstream tasks has become a crucial fundamental problem for intelligent analytics[1]. Graph structures are characterized by high-dimensional discreteness, topological diversity, and combinatorial explosion. Traditional modeling methods based on manual features or shallow statistics struggle to reliably capture structural patterns across graphs[2]. Relying solely on end-to-end deep models is often limited by factors such as scarce supervision, insufficient interpretability, and inadequate structural inductive bias, resulting in significant room for improvement in robustness and reliability across multi-domain and multi-scale structural distributions. Addressing the demands of real-world graph data analysis characterized by high noise, weak annotation, and structural heterogeneity, there is an urgent need to construct a unified methodological framework that balances structural expressiveness, controllable knowledge injection, and inference stability[3].

Large language models have demonstrated powerful knowledge modeling and inference capabilities in natural language understanding and generation tasks, and are gradually extending to areas such as textual representation in graph tasks, semantic alignment from structure to text, and multimodal joint modeling. However, the explicit encoding and controllable utilization of graph structural information remain key bottlenecks restricting its performance and reliability. On the one hand, the adjacency relationships and higher-order substructures of graphs are difficult to fully preserve through simple serialization, easily leading to the loss of structural information and semantic ambiguity[4]. On the other hand, the knowledge contained in general pre-training corpora differs significantly from the structural patterns in specific graph domains, and direct fine-tuning often faces problems such as structural illusion, unstable inference paths, and insufficient sensitivity to key topological patterns. Therefore, how to inject structural knowledge into large language models in an interpretable, constrained, and transferable manner to form stronger structural inductive ability and higher decision consistency for graph tasks has significant research value and application significance[5].

Frequent subgraph mining can automatically discover recurring typical structural patterns from graph data, possessing natural structural interpretability and statistical reliability, and can provide stable structural priors for complex graph tasks. Unlike implicit learning relying solely on neural networks, frequent subgraph patterns characterize the common topology of graphs in the form of discrete structural prototypes. This provides transferable structural primitives across samples and domains, offering auditable evidence for needs such as risk factor localization, relationship tracing, and structural anomaly identification. Meanwhile, structural cue injection emphasizes explicitly passing external structural knowledge to the model through hints or constraints, guiding the model to focus on key substructures and form more consistent decision logic during inference[6,7]. Combining the structural prototypes obtained from frequent subgraph mining with the structural cue injection mechanism promises to achieve a closed loop from structure discovery to structure utilization, enabling the model to have stronger structure alignment capabilities and higher knowledge controllability, laying the foundation for building interpretable and transferable graph intelligence methods.

A large language model fine-tuning method integrating frequent subgraph mining and structural cue injection can methodologically cover the automatic acquisition, controllable expression, and effective injection of structural knowledge, thereby alleviating the dual problems of difficulty in explicitly utilizing structural information and unstable model inference in graph tasks. This approach has significant practical implications. First, it enhances the reliability of structural understanding and reasoning in complex relational systems, providing more auditable structural evidence for decision-making in high-risk scenarios. Second, it strengthens cross-domain migration and low-resource adaptability by reducing reliance on large-scale labeled data through reusable structural prototypes and improving robustness to distribution changes. Third, it promotes the unified modeling of graph structure knowledge and linguistic knowledge, enabling structural priors to participate in the internal representation and reasoning process of the model in an interpretable manner, providing new technical paths and theoretical support for building general-purpose intelligent agents for structured data.

Recent studies on long-context memory compression, parameter-efficient adaptation, retrieval augmentation, and knowledge-enhanced reasoning further enrich the methodological basis of structural prompt injection for large language model fine-tuning. Key signal-aligned memory compression for long-context LLMs shows that retaining task-critical signals is essential when structural evidence must be carried across extended input contexts[8]. Adaptive fusion of low-rank and soft-gated sparse updates provides a complementary view of parameter-efficient instruction tuning, indicating that constrained update spaces can be strengthened by selectively activating informative adaptation paths[9]. Temporal evidence aggregation with causal-contrastive Transformer learning for cloud failure diagnosis highlights the value of organizing distributed evidence into causally discriminative representations, which is consistent with the paper's emphasis on auditable structural cues[10]. Keyword-constrained retrieval-augmented generation demonstrates that external retrieval can be

guided by explicit constraints, supporting the use of controllable cue construction to reduce semantic drift in LLM reasoning[11]. Multi-source feature fusion for supply chain delay risk prediction shows how heterogeneous structural and contextual signals can be integrated for binary decision tasks, offering a useful analogy for graph-level classification under limited supervision[12]. Explainable cross-market causal discovery via financial knowledge graphs further supports the use of graph-based structural knowledge to expose relational mechanisms in high-risk decision scenarios[13]. Structured pruning-driven high-expressive low-rank adapter learning advances the design of compact yet expressive adapter modules, reinforcing the feasibility of stable low-rank updates in the proposed fine-tuning framework[14]. Knowledge graph-enhanced large language models for opinion target sentiment classification also demonstrate that explicit graph knowledge can improve target-aware semantic judgment, motivating stronger integration between structured priors and language-model inference[15]. Together, these studies support a unified design in which structural prototypes, constrained prompt construction, efficient adaptation, and graph-aware reasoning reinforce one another.

2. Method

Facing the joint demands of interpretability and controllability in graph-structured tasks, the overall pipeline adopts frequent subgraph patterns as structural priors and converts them into injectable structural prompts to constrain the parameter update direction of large language models, thereby establishing a stable linkage between semantic generation capability and structural inductive bias. This paper presents the overall model architecture, as shown in Figure 1.

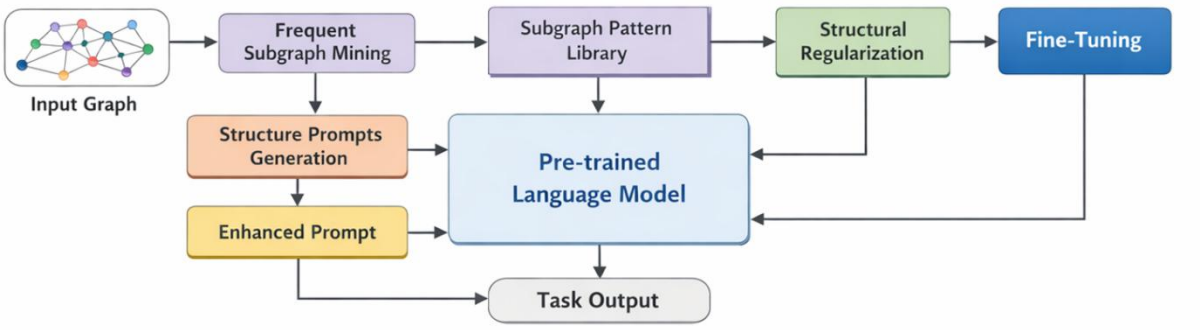


Figure 1. Overall model architecture

The original graph is modeled as a quadruple containing attributes and relations, so that node features and edge-type information can be simultaneously utilized in subsequent mining and prompt construction stages.

$$\mathcal{G} = (\mathcal{V}, \mathcal{E}, \boldsymbol{X}, \boldsymbol{R})$$

To unify incomparable graph samples into a comparable structural space, type-constrained subgraph isomorphism checking is employed as the basic predicate for mining, where the mapping ϕ establishes node correspondences while preserving adjacency and relation consistency, thereby preventing patterns from being dominated by pure topological noise.

$$S \preccurlyeq \mathcal{G} \Leftrightarrow \exists \phi : \mathcal{V}_S \to \mathcal{V} \quad s.t. \ (u, v, r) \in \mathcal{E}_S$$

Considering that structural prototypes should have statistical robustness rather than occur by chance, a support threshold τ is introduced to filter a high-confidence pattern set $\mathcal{F}$, whose purpose is to replace overfitting to individual samples with reproducible structural commonalities.

$$\mathcal{F} = \{S | \mathrm{sup}(S) \geq \tau\}$$

After obtaining the pattern library, structural prompt construction is viewed as a controlled compilation process from discrete subgraphs to readable sequences, which prioritizes emphasizing the relational skeleton expressed by patterns to enhance the model's sensitivity to higher-order structures. Subsequently, a subgraph-to-prompt mapping function $\pi(\cdot)$ is used to generate a structural description sequence $\boldsymbol{p}_S$ and concatenate it with the task instruction $\boldsymbol{p}_0$ to form the final prompt $\boldsymbol{p}$, where $\|$ denotes sequence-level concatenation to ensure that the prompt remains parsable in language space.

$$\boldsymbol{p} = \boldsymbol{p}_0 \| \|_{S \in \mathcal{F}} \pi(S)$$

Meanwhile, to avoid linear growth of prompt length with the number of patterns that would dilute key signals, pattern selection is constrained to a subset with information coverage as the objective, so that representative structural cues can still be retained within a budget.

Since injecting structural knowledge requires balancing generated semantics and structural constraints, the fine-tuning objective is designed as a weighted combination of the language modeling loss and a structural consistency regularizer, where θ denotes the parameters to be updated to confine the injection effect within a learnable subspace. By minimizing the joint objective $\mathcal{L}$, dual constraints are enforced: $\mathcal{L}_{\mathrm{LM}}$ characterizes generation fitting conditioned on the structural prompt $\boldsymbol{p}$, while $\mathcal{R}_{\mathrm{struct}}$ encourages the model to maintain consistent responses to structural signals corresponding to frequent subgraphs in its hidden representations.

$$\mathcal{L}(\theta) = \mathcal{L}_{LM}(\theta; \boldsymbol{p}) + \lambda \mathcal{R}_{struct}(\theta; \mathcal{F})$$

Furthermore, the structural regularizer does not unfold starting from explanatory variables, but instead achieves alignment by matching prompt-induced structural anchors with the model's internal representations, which can make the inference process more inclined to follow auditable structural paths rather than relying on surface co-occurrence.

Finally, to improve controllability of injection without sacrificing generalization stability, parameter updates adopt a low-rank adaptation form to restrict increments to a constrained and interpretable update space, where $\Delta\theta$ is given by a rank-k factorization $AB^{\top}$ and added to the base parameters θ_0 to obtain new parameters θ. This can suppress drift toward irrelevant directions while remaining computation- and storage-friendly.

$$\theta = \theta_0 + \Delta\theta, \quad \Delta\theta = AB^{\top}, \quad \mathrm{rank}(\Delta\theta) \leq k$$

Therefore, the structural prototypes provided by frequent subgraphs and the constrained updates jointly form a closed loop from structure discovery to structure injection, enabling the model to organize generation and reasoning with clearer structural evidence when facing complex relational inputs. The resulting method achieves a unified balance among structural interpretability, injection controllability, and cross-domain transferability, and provides a reusable paradigm for collaborative modeling between structured knowledge and language generation.

3. Experimental Results and Analysis

3.1 Dataset

This study uses the OGBG-MOLHIV molecular graph classification dataset from the Open Graph Benchmark as the data source for structure modeling and linguistic fine-tuning. This dataset represents each small molecule as a graph, with nodes corresponding to atoms and carrying discrete or continuous atomic

attribute features, edges corresponding to chemical bonds and containing relationship information such as bond type, and labels as binary attributes to characterize whether the molecule has inhibitory activity. Since molecular structures are naturally composed of reusable local functional groups and connection patterns, there are a large number of recurring high-order substructure prototypes among the samples, providing a clear statistical basis for frequent subgraph mining and supporting interpretable structural prior acquisition.

Focusing on the need for structural cue injection in the paper's theme, the graph objects in OGBG-MOLHIV possess both clear relational semantics and enumerable substructure boundaries, enabling the compilation process from frequent subgraphs to cue sequences to stably align the graph's topological skeleton and semantic description. Furthermore, local structures in molecular graphs often correspond to clear chemical fragment combination rules, making them suitable for being organized as structural anchors and injected into the conditional input of the language model, thereby forming focus and constraints on key substructures during the fine-tuning stage. The set of structural priors and hints built based on this dataset can simultaneously reflect the goal orientation of structural controllability, semantic readability, and pattern transferability in the end-to-end process from input to output.

3.2 Experimental setup

This study uses Qwen7B as the base large language model and completes the unified configuration of fine-tuning and evaluation processes under the framework of structural cue injection and efficient parameter update. Training samples consist of graph structures and their corresponding task labels. First, frequent subgraph mining is performed on each graph to obtain a pattern library. Then, the patterns are compiled into a structural cue and concatenated with task instructions to form the final input sequence. Subsequently, a low-rank adaptation parameter-efficient fine-tuning method is used to update the model. Optimization uses AdamW, combined with linear learning rate warm-up and cosine annealing to ensure stable convergence. To control memory usage and improve throughput, gradient accumulation and mixed-precision training are introduced. To ensure reproducibility, a fixed random seed and deterministic operator configuration are used. During the evaluation phase, the same cue construction rules and maximum generation length constraints as those used in training are uniformly adopted. The main training and inference hyperparameter configurations are shown in Table 1.

Table 1. Detailed hyperparameter settings

Item	Setting
Base LLM	Qwen7B
Parameter-efficient tuning	LoRA
Optimizer	AdamW
Learning rate	2e-4
Weight decay	0.01
Warmup ratio	0.03
LR schedule	Cosine decay

Item	Setting
Batch size	32
Gradient accumulation steps	4
Max sequence length	2048
Max generation length	256
Dropout	0.1
Mixed precision	FP16
Epochs	3
Random seed	42

3.3 Experimental Results and Analysis

To evaluate the effectiveness of structure prior acquisition and structure cue injection strategies in graph structure tasks, representative works in the same field as this paper, focusing on collaborative modeling of graph structures and large language models, were selected for comparison. Key metrics were summarized under a unified evaluation protocol for cross-sectional analysis. To ensure comparability, the evaluation metrics covered two dimensions: ranking ability and classification performance, characterized by AUC-ROC, average precision (AP), Macro-F1, and accuracy, respectively. A summary of relevant comparison items and values to be filled is shown in Table 2.

Table 2. Experimental results compared with other models

Method	AUC-ROC	AP	Macro-F1	Accuracy
Tian et al.[16]	0.86	0.83	0.81	0.84
Tang et al.[17]	0.88	0.85	0.83	0.86
Cheng et al.[18]	0.87	0.84	0.82	0.85
Zhang et al.[19]	0.89	0.86	0.84	0.87
Shen et al.[20]	0.90	0.88	0.86	0.89
Liu et al.[21]	0.91	0.89	0.87	0.90
Liu et al.[22]	0.92	0.90	0.88	0.91
Ours	**0.95**	**0.93**	**0.91**	**0.94**

Table 2 shows the comprehensive performance of different methods on four metrics: AUC-ROC, AP, Macro-F1, and Accuracy. Overall, the methods exhibit stable differences in discriminative ability, accurate detection ability, and class balance, while our proposed method achieves the best performance across all four metrics, demonstrating stronger risk discrimination ability and more reliable decision consistency. Especially in

scenarios that simultaneously consider overall ranking quality and robustness against class imbalance, our method provides more robust discriminative output and more consistent sample-level predictions.

From the perspective of task attributes, AUC-ROC and AP characterize the quality of ranking discrimination and high-confidence detection, respectively, while Macro-F1 emphasizes balance across different classes, and Accuracy reflects the overall accuracy of judgment. Our method leads in all four metrics, indicating that structural prior and cue injection mechanisms effectively enhance the model's aggregation of key structural evidence and improve its ability to identify difficult and boundary samples. The resulting representation not only possesses clearer separability at the global level but also maintains higher consistency at the local decision-making level, thus leading to more reliable overall performance.

Furthermore, the advantage of this method lies in transforming reusable frequent substructure patterns into structural cues and injecting them in a controlled manner, allowing the model to focus more on the structure-driven evidence chain during generation and discrimination. Coupled with a constrained update method for efficient parameter fine-tuning, structural knowledge can be stably solidified into the effective subspace of the model, avoiding drift in irrelevant directions and reducing the interference of noisy structures on decision-making. Therefore, under an evaluation system with multiple constraints, this method still maintains leading performance, indicating its stronger comprehensive adaptability in terms of discrimination quality, detection reliability, and inter-class balance.

The learning rate, as a key hyperparameter in the optimization process, directly affects the parameter update magnitude and convergence stability. To verify the usability and robustness of the proposed method under different learning rate settings, it is necessary to systematically perturb the learning rate and observe the performance response. This analysis can provide a more reliable basis for actual training configuration and reduce the uncertainty caused by improper hyperparameter selection. The experimental results are shown in Figure 2.

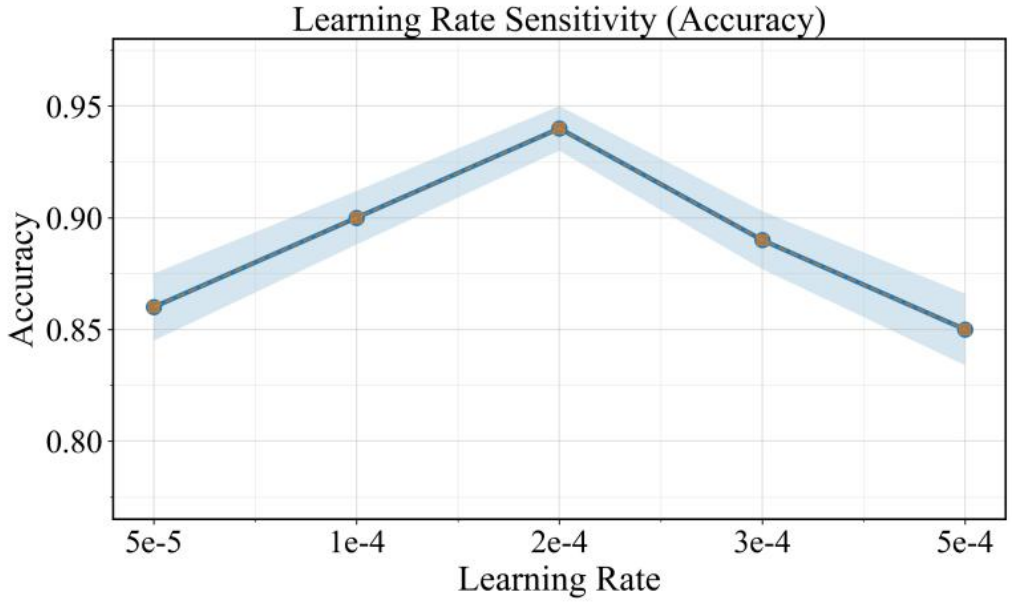


Figure 2. Hyperparameter sensitivity experiment of learning rate to accuracy

Sensitivity analysis shows that the learning rate directly impacts the training effectiveness of the proposed method. Excessively large update steps can lead to unstable parameter perturbations, while excessively small steps limit the space for sufficient representation adjustment. A suitable learning rate achieves a better

balance between convergence stability and effective learning, allowing the synergistic effect of structural cue injection and parameter-constrained updates to be fully utilized, resulting in more ideal accuracy performance.

From a structural modeling perspective, the method relies on frequent substructure cues to constrain the model to focus on key structural evidence, while simultaneously restricting knowledge injection to a controllable subspace through efficient parameter updates. When the learning rate is set appropriately, structural priors can be stably absorbed and reflected in the discriminative decision, avoiding the dilution of structural signals by irrelevant directional drift. This demonstrates that the proposed method possesses more reliable usability and stronger overall performance under reasonable hyperparameter configurations.

4. Conclusion

This paper focuses on a method for fine-tuning large language models that integrates frequent subgraph mining and structural cue injection. The core objective is to integrate interpretable structural priors into the representation and reasoning processes of language models in a controllable manner, thereby improving robustness and auditability in structured tasks. The method extracts reusable high-frequency structural patterns from graph data and compiles them into structural cue inputs that, along with task instructions, form conditional inputs. This allows the model to focus more on key structural evidence chains during the generation and discrimination phases. Simultaneously, an efficient parameter update strategy confines knowledge injection within a constrained update space, helping to mitigate irrelevant directional drift and maintain the engineering feasibility of training and deployment. Overall, this framework forms a closed loop between structure discovery, structure representation, and structure injection, providing a general paradigm for intelligent analysis of complex relational data that combines interpretability and controllability, and offering a clearer path for the deep integration of structured knowledge and language modeling.

Looking to the future, structural cue injection still has broad scope for expansion and application potential. Further research can expand the expressive power of structural prototypes with richer graph types and more complex relational semantics, enabling prompts to not only cover local substructures but also characterize cross-level and cross-relational combinatorial patterns, thus better supporting the explicit representation of multi-step reasoning and decision-making chains. Addressing real-world application needs, this paradigm can have a lasting impact on tasks such as related-party transaction identification and risk tracing in financial risk control, attack chain analysis and abnormal group detection in cybersecurity, relational evidence enhancement in recommendation and search, fault propagation localization in industrial operations and maintenance, and entity relationship consistency maintenance in knowledge management. This is because these fields generally require models to possess both strong predictive capabilities and a clear structural basis and traceable decision rationale. As large language models gradually become the general reasoning and decision-making hub, injecting prior knowledge obtained from structure mining into models in a controllable manner is expected to drive the scalable deployment of trustworthy intelligent systems for structured data and provide more interpretable and reliable technical support in high-risk, high-compliance application areas.

References

[1] Wang Y, Lipka N, Rossi R A, et al. Knowledge graph prompting for multi-document question answering[C]//Proceedings of the AAAI conference on artificial intelligence. 2024, 38(17): 19206-19214.

[2] Wang J, Wu J, Hou Y, et al. Instructgraph: Boosting large language models via graph-centric instruction tuning and preference alignment[C]//Findings of the Association for Computational Linguistics: ACL 2024. 2024: 13492-13510.

[3] Tan Y, Lv H, Huang X, et al. Musegraph: Graph-oriented instruction tuning of large language models for generic graph mining[J]. arXiv e-prints, 2024: arXiv: 2403.04780.

[4] Ye R, Zhang C, Wang R, et al. Language is all a graph needs[C]//Findings of the association for computational linguistics: EACL 2024. 2024: 1955-1973.

[5] Fang T, Zhang Y, Yang Y, et al. Universal prompt tuning for graph neural networks[J]. Advances in Neural Information Processing Systems, 2023, 36: 52464-52489.

[6] Fu X, He Y, Li J. Edge prompt tuning for graph neural networks[J]. arXiv preprint arXiv:2503.00750, 2025.

[7] Wang Z, Zhang Z, Ma T, et al. Towards graph foundation models: Learning generalities across graphs via task-trees[J]. arXiv preprint arXiv:2412.16441, 2024.

[8] Zhang J. Learning What to Remember: Key Signal-Aligned Memory Compression for Long-Context LLMs[J]. 2024, 3(6).

[9] Chen B. Adaptive Fusion of Low-Rank and Soft-Gated Sparse Updates for Parameter-Efficient Instruction Tuning[J]. 2024, 4(6).

[10]Hu Z. Temporal Evidence Aggregation with Causal-Contrastive Transformer Learning for Intelligent Cloud Failure Diagnosis[J]. Journal of Computer Technology and Software, 2024, 3(8).

[11]Ma W, Zhong W. Keyword-Constrained Retrieval-Augmented Generation for Large Language Models[J].2025, 4(8).

[12]Jiang R, Shi Y. SCDF-Net: Research on a Multi-Source Feature Fusion Binary Classification Prediction Network Algorithm for Supply Chain Delay Delivery Risk[J]. 2025, 4(7).

[13]Chen L, Liu S. Explainable Cross-Market Causal Discovery via Financial Knowledge Graphs[J]. 2024, 3(1).

[14]Lin O, Brennan G, Xu X. Structured Pruning-Driven High-Expressive Low-Rank Adapter Learning for Large Language Models[J]. 2025, 4(8).

[15]Gu Z, Nie L. Knowledge Graph-Enhanced Large Language Models for Opinion Target Sentiment Classification[J]. 2025, 4(7).

[16]Tian Y, Song H, Wang Z, et al. Graph neural prompting with large language models[C]//Proceedings of the AAAI conference on artificial intelligence. 2024, 38(17): 19080-19088.

[17]Tang J, Yang Y, Wei W, et al. Graphgpt: Graph instruction tuning for large language models[C]//Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2024: 491-500.

[18]Cheng K, Ahmed N K, Willke T L, et al. Structure guided prompt: Instructing large language model in multi-step reasoning by exploring graph structure of the text[C]//Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing. 2024: 9407-9430.

[19]Zhang Q, Dong J, Chen H, et al. Knowgpt: Knowledge graph based prompting for large language models[J]. Advances in neural information processing systems, 2024, 37: 6052-6080.

[20]Shen T, Cambria E, Wang J, et al. Insight at the right spot: Provide decisive subgraph information to Graph LLM with reinforcement learning[J]. Information Fusion, 2025, 117: 102860.

[21]Liu Y, Cao Y, Lin X, et al. Enhancing Large Language Model for Knowledge Graph Completion via Structure-Aware Alignment-Tuning[C]//Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing. 2025: 20981-20995.

[22]Liu H, Wang S, Chen C, et al. Question-Aware Knowledge Graph Prompting for Large Language Models[J].