Meta Relation Assisted Explanatory Model for Heterogeneous Graph Neural Networks

Yibowen Zhao, Yonghui Xu✉, Yu Zhang, Qingzhong Li, Xudong Lu, Wei He, and Lizhen Cui✉

School of Software & Joint SDU-NTU Centre for Artificial Intelligence Research (C-FAIR), Shandong University, Shandong, China
{ybw.zhao, z.yu}@mail.sdu.edu.cn, xu.yonghui@hotmail.com, {lqz, dongxul, hewei, clz}@sdu.edu.cn

Abstract. Heterogeneous Graph Neural Networks (HGNNs) have become essential for modeling complex systems with diverse, interacting components. Meta relations, which represent distinct interaction types between source and target nodes, are crucial for HGNNs to capture rich, context-specific semantic information. However, existing GNN's explanatory models struggle to effectively capture the unique semantic aspects of HGNNs, often leading to incomplete or misleading explanations. Moreover, current explanatory models tend to assess edge significance globally, overlooking finer-grained differences among edges with shared semantic content, thus limiting their ability to provide context-specific insights. To address these limitations, we propose MR-Explainer, a Meta Relation-Assisted Explanatory Model for HGNNs. MR-Explainer incorporates a heterogeneous information fusion module to integrate structural and semantic data, generating comprehensive edge representations. Additionally, a multiplex salience-aware module computes salience values both globally and within each meta relation, ensuring that explanations are context-sensitive and precise. Evaluations on classification tasks demonstrate MR-Explainer's effectiveness in delivering accurate, nuanced explanations for HGNN's outcomes.

Keywords: Explainable Graph Neural Network · Heterogeneous Graph Neural Network · Post-hoc Explanation.

1 Introduction

Heterogeneous Graph Neural Networks (HGNNs) are crucial for modeling complex systems composed of diverse, interacting components with intricate architectures and rich semantic information. HGNNs leverage meta relations, which represent distinct interaction categories between node types, to capture the nuanced relationships between different nodes and edges. These meta relations provide essential type-specific semantic information [5]. While HGNNs are highly parameterized to model complex interdependencies, they face a significant challenge in terms of explainability, which is especially critical in high-stakes fields like healthcare and finance, where understanding decision-making is vital [4].

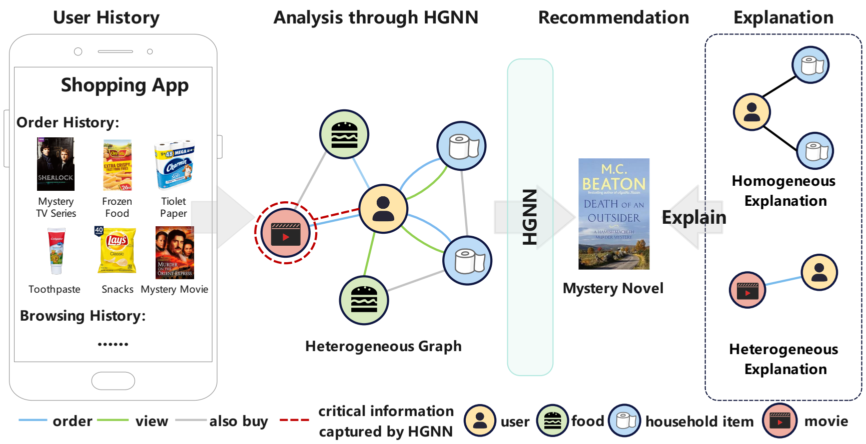


Fig. 1: An instance of an HGNN-based recommendation system involves modeling user history as a heterogeneous graph and analyzing it through HGNN to achieve efficient recommendations. However, current explanatory model might overlook the distinct interactions among various heterogeneous information, leading to erroneous homogeneous explanation.

Existing GNN's explanatory models primarily focus on identifying important substructures (e.g., edges) [3]. However, they often overlook the intricate interactions and semantic information within heterogeneous graphs that directly influence HGNN's outcomes. These interactions, defined by meta relations and the distinct attributes of nodes and edges, are essential for understanding the factors driving predictions. For instance, in a recommendation system (Figure 1), an HGNN might recommend a mystery novel based on thematic similarities with a user's interactions with TV shows and movies. However, explanatory models focusing on structural factors may erroneously highlight frequent interactions, such as groceries, instead of semantically relevant items. Thus, HGNN-specific explanatory models are needed to capture both structural and semantic nuances.

Moreover, while existing explanatory models assess edge significance relative to the entire graph, they often overlook the finer distinctions among edges within the same meta relation. When evaluating edge importance on a global scale, these explanatory models rely on a single significance score for each edge relative to the entire graph, which fails to capture the context-specific roles that edges play within particular meta relations [8]. For example, in the meta relation "user buys items", edges like "user buys toilet paper" and "user buys a movie" may have very different contributions to predicting a user's preference for a mystery novel. A global score that treats all edges in the same way across the entire graph does not account for the diverse contextual roles that these edges may have within the specific meta relation. As a result, such models cannot fully recognize the varying importance of edges in different meta relations, limiting their ability to provide meaningful and precise explanations.

To address this, we propose MR-Explainer, a Meta Relation-Assisted Explanatory Model for HGNNs. MR-Explainer integrates structural and semantic information through a heterogeneous information fusion module that generates comprehensive edge representations by considering multiple sources within meta relations. It also introduces a multiplex salience-aware module, computing salience values globally and within each meta relation to assess edge importance more accurately. We validate MR-Explainer on two widely used HGNN backbones, RGCN [10] and RGAT, across multiple datasets, demonstrating its effectiveness in providing nuanced explanations for HGNN's outcomes. Our contributions include:

- We introduce MR-Explainer, an innovative explanatory model that enhances HGNN's explainability by integrating structural and semantic information, offering detailed explanations that capture the complex relationships in heterogeneous graphs.
- MR-Explainer introduces a multiplex salience-aware module to assess edge significance both globally and within each meta relation, enabling more context-sensitive and precise explanations of HGNN's outcomes.
- We conduct extensive experiments with MR-Explainer using two HGNN backbones (RGCN and RGAT) across various datasets, demonstrating its effectiveness in providing meaningful explanations for HGNN's outcomes.

2 Related Works

Several explanatory models have been developed to enhance the explainability of GNNs by capturing graph structure. GNNExplainer [15] identifies key subgraphs via perturbations but risks introducing "evidence artifacts." PGExplainer [7] addresses this by learning a discrete mask through mutual information, providing more controlled subgraph selection. PGM-Explainer [11] uses probabilistic graphical models to generate explanations via conditional probabilities, offering reliability in synthetic data but lacking flexibility in real-world applications. In chemistry, SME [13] uses molecular segmentation to assess the impact of substructures on GNN predictions. While SME aligns with chemical knowledge, its domain-specific nature limits its generalizability. FANS [2] evaluates feature importance via counterfactual samples, offering causal insights, but struggles with heterogeneous data. SAME [14] identifies impactful substructures through Monte Carlo tree search, while GNNShap [1] improves efficiency through parallelization. However, most existing explanatory models fail to account for the diverse semantic context within heterogeneous graphs that significantly influences HGNN's outcomes. This limitation motivates the development of MR-Explainer, an explanatory model specifically designed for HGNN.

3 Preliminaries

Heterogeneous graphs, unlike homogeneous graphs, contain multiple types of nodes and edges, which enhances their ability to model complex relationships in real-world systems. However, this added complexity poses significant challenges for the explainability of HGNNs. In this section, we define heterogeneous graphs and then explore the challenges associated with explaining HGNN predictions, providing the foundation for our proposed MR-Explainer.

Definition of Heterogeneous Graphs: A heterogeneous graph is represented as a directed graph $G = (V, E, A, R)$, where V is the set of nodes, E is the set of edges, A is the set of node types, and R is the set of edge types. Each node $v \in V$ and edge $e \in E$ is mapped to a type via functions $\tau(v) : V \to A$ and $\phi(e) : E \to R$, respectively. Since edges in heterogeneous graphs can connect nodes of different types, we introduce the concept of **meta relations**, denoted as $(\tau(s), \phi(e), \tau(t))$, where $\tau(s)$ and $\tau(t)$ are the types of the source and target nodes, and $\phi(e)$ represents the edge type. Meta relations allow for a refined representation of interactions, providing a clearer understanding of the complex relationships in heterogeneous graphs. For instance, in a recommender system, meta relations capture interactions such as "user buys book," denoted as $(\text{"user"}, \text{"buy"}, \text{"book"})$. This definition enables systematic categorization of edges and nodes, aiding in the modeling of diverse interactions.

Definition of Explanations in HGNN: Given an HGNN model $f(\cdot)$, an input graph G, and a prediction y for a node v, the explanatory model aims to identify a minimal subgraph $G_s = (V_s, E_s, A_s, R_s)$ that captures the most relevant substructures from each type of meta relation. This subgraph reflects the critical semantic and structural information influencing the prediction. Grounded in the Graph Information Bottleneck (GIB) theory [12], the model seeks to retain the most relevant information for the prediction while discarding irrelevant details. The goal can be formalized as:

$$\arg\max_{\tau} I(Y, G_s) = H(Y) - H(Y|G_s) \quad \text{s.t.} \quad \min |G_s|, \tag{1}$$

where $I(\cdot)$ denotes mutual information, $H(\cdot)$ denotes entropy, and τ denotes the parameters of the explanatory model. The model outputs a subgraph G_s that identifies the most influential substructures for the HGNN's outcomes.

This definition lays the groundwork for understanding the explanation challenge in HGNNs and introduces the MR-Explainer, designed to capture both structural and semantic information, thereby providing more accurate explanations of HGNN's outcomes.

4 Methodology

To provide accurate explanations for Heterogeneous Graph Neural Networks (HGNNs), we introduce MR-Explainer, an explanatory model that integrates two key components: the heterogeneous information fusion module and the multiplex salience-aware module. These modules collaboratively harness both structural and semantic information across heterogeneous graphs, facilitating context-sensitive explanations that align with the intricate dependencies of the data. The overall framework is illustrated in Figure 2.

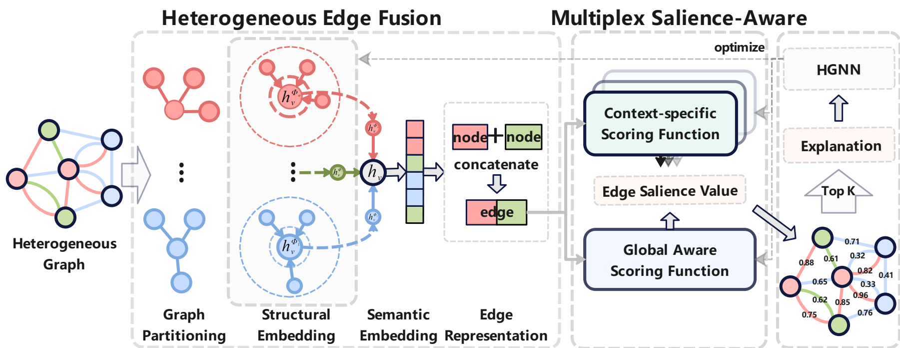


Fig. 2: The framework of MR-Explainer. MR-Explainer employs the heterogeneous information fusion module to integrate structural and semantic information into edge representations. Subsequently, the multiplex salience-aware module computes each edge's global salience relative to the entire graph and context-specific salience within its meta relation, thereby generating explanations tailored to HGNN's outcomes.

4.1 Heterogeneous Information Fusion Module

The heterogeneous information fusion module is designed to effectively capture both structural and semantic information in heterogeneous graphs through a two-step embedding process. First, the heterogeneous graph is partitioned into multiple homogeneous views, each corresponding to a distinct meta relation. This segmentation enables the application of graph neural network (GNN) techniques to aggregate structural information within each view. In the second step, semantic information is integrated across these views, producing unified node representations that capture the full scope of the graph's complexity. The edge representations are derived by concatenating the embeddings of the connected nodes, forming a rich set of features that serve as the foundation for computing edge salience.

The first step involves partitioning the heterogeneous graph G into several homogeneous subgraphs G_i, each corresponding to a meta relation ϕ_i. Formally, this partitioning is represented as:

$$\mathcal{G}_\phi = \{G_i = (V_i, E_i, A_i, R_i) \mid \phi_i \in \phi\}, \tag{2}$$

where each G_i denotes a specific meta relation ϕ_i. The structural embeddings within these views are computed using graph neural network methods. In this paper, we utilize the Graph Convolutional Layer [6] to compute node v's embeddings $h_v^{\phi_i(l+1)}$ at layer $l+1$:

$$h_v^{\phi_i(l+1)} = \sigma\left(\sum_{u \in N(v)_{\phi_i}} \frac{1}{c_{vu}} W_{\phi_i}^{(l)} h_u^{\phi_i(l)}\right), \tag{3}$$

where $N(v)_{\phi_i}$ denotes the neighborhood of v in relation ϕ_i, c_{vu} denotes a normalization factor, and $W_{\phi_i}^{(l)}$ denotes the weight matrix for layer l.

Once the node embeddings are computed, we apply an aggregation strategy to combine the structural embeddings across the different meta relation views. We explore the following aggregation strategies:
Sum Aggregation: This method sums the embeddings from all relations, preserving the full range of relational data:

$$h_v^{\text{sum}} = \sum_i h_v^{\phi_i}. \tag{4}$$

This strategy retains all relational information but may disproportionately emphasize highly active relations.

After constructing the node representations, we derive edge representations by concatenating the embeddings of the source and target nodes:

$$h_m = [h_s, h_t], \tag{5}$$

where h_s and h_t denote the node embeddings for the source and target nodes, respectively. By systematically integrating structural and semantic information, the heterogeneous information fusion module captures the complex interdependencies in heterogeneous graphs, enabling MR-Explainer to produce enriched edge representations that inform the explanation process.

4.2 Multiplex Salience-aware Module

To further enhance the explainability of HGNN, we introduce the multiplex salience-aware module, which assesses edge importance at both the global and meta relation levels. This dual-level salience computation enables a comprehensive understanding of each edge's contribution to the HGNN's outcomes.

The global salience value s_g is computed using a global-aware scoring function $\tau_g(\cdot)$, which evaluates the edge's contribution to the overall HGNN's outcomes:

$$s_g = \tau_g(h_m) = \sigma\left(W_2 \cdot \sigma(W_1 \cdot h_m + b_1) + b_2\right), \tag{6}$$

where W_1 and W_2 are learnable weight matrices, b_1 and b_2 are biases, and σ is the ReLU activation function. This scoring function captures the global significance of each edge within the overall graph structure.

For each meta relation ϕ_i, a context-specific scoring function $\tau_c^i(\cdot)$ is used to evaluate the importance of edges within that specific meta relation. The context-specific salience value s_c is computed as:

$$s_c = \tau_c^i(h_m) = \sigma\left(W_2^{(i)} \cdot \sigma(W_1^{(i)} \cdot h_m + b_1^{(i)}) + b_2^{(i)}\right), \tag{7}$$

where the parameters $W_1^{(i)}$, $W_2^{(i)}$, $b_1^{(i)}$, and $b_2^{(i)}$ are specific to each meta relation ϕ_i. This function captures the unique contextual role of each edge within its corresponding relation. The final contribution weight cw_i of each edge combines both global and context-specific salience values:

$$\mathrm{cw}_i = s_c \cdot s_g. \tag{8}$$

This product ensures that each edge's contribution reflects both its overall importance and its significance within the specific meta relation, yielding a comprehensive measure of its relevance.

The multiplex salience-aware module is jointly optimized with the heterogeneous information fusion module. The overall objective for MR-Explainer follows the Graph Information Bottleneck (GIB), aiming to retain the most relevant information for HGNN's outcomes while minimizing unnecessary details:

$$\mathcal{L} = \arg\max_{\tau} I(Y, G_s) = \arg\min_{\tau} |f(G)(v) - f(G_s)(v)|, \tag{9}$$

where $I(Y, G_s)$ quantifies the mutual information between the predicted outcome and the explanation.

After computing the contribution weights cw_i, the top k edges with the highest weights are selected as the final explanation. This ensures that the explanation remains concise yet informative, offering a clear understanding of the HGNN's outcomes. By assessing edge salience at both the global and meta relation levels, the multiplex salience-aware module enhances the explainability of HGNN. This dual-level salience analysis uncovers the complex interplay between heterogeneous graph, improving the clarity and quality of the explanations.

5 Experiment

5.1 Experimental setups

Datasets. We evaluate the MR-Explainer on five heterogeneous graph datasets: MUTAG, BGS, ACM, AIFB, and AM. MUTAG focuses on bioinformatics, BGS on geological classification, ACM on co-authorship in computer science, and AIFB on the University of Karlsruhe's organizational structure. AM represents consumer behavior in Amazon's co-purchase network. **HGNNs.** In experiments, the Relational Graph Convolutional Network (RGCN) and Relational Graph Attention Network (RGAT) are applied as the HGNNs to be explained. RGCN models multi-relational data, while RGAT enhances the convolution process through attention mechanisms. **Baselines.** We adapt two famous GNN's explanatory models, GNNExplainer and PGExplainer, to heterogeneous graphs, given that existing models are not directly applicable to HGNNs. In addition, a random explanation is built to assign random weights. **Metrics.** We use two metrics to evaluate the explanatory models: Fidelity (FID) [9] and the Area Under the Fidelity-Sparsity Curve (AUSFC). FID measures explanation quality. We utilize the sparsity [16] to quantifies explanation conciseness, and AUSFC captures the trade-offs between fidelity and sparsity. **Implementation Details.** Our implementation is based on PyTorch with random seed 42. We use Xavier initialization for node features, and a two-layer architecture for RGCN and RGAT.

Table 1: Comparative Experiments

		AM		ACM		BGS		AIFB		MUTAG	
		FID↓	AUSFC↓	FID↓	AUSFC↓	FID↓	AUSFC↓	FID↓	AUSFC↓	FID↓	AUSFC↓
RGCN	Random	0.3785	2.2190	0.1617	0.9183	0.0478	0.2570	0.2226	1.3940	0.0278	0.1486
	GNNExplainer	0.3201	1.9943	0.1501	0.8932	0.0592	0.3145	0.2019	1.3119	0.0257	0.1385
	PGExplainer	0.3370	2.1388	0.2042	1.0127	0.0469	0.2514	0.1911	1.2630	0.0271	0.1353
	MR-Explainer	**0.1067**	**0.7398**	**0.0508**	**0.4367**	**0.0311**	**0.1823**	**0.1307**	**0.7317**	**0.0136**	**0.0787**
RGAT	Random	0.0305	0.1945	0.2711	1.9875	0.0096	0.0711	0.1072	0.8667	0.0186	0.1171
	GNNExplainer	0.0269	0.1922	0.2981	2.0637	0.0137	0.0931	0.1202	0.9037	0.0177	0.1135
	PGExplainer	0.0016	0.0476	0.1935	1.6078	0.0124	0.0747	0.1148	0.9019	0.0150	0.1057
	MR-Explainer	**0.0013**	**0.0311**	**0.0515**	**0.4554**	**0.0069**	**0.0383**	**0.0265**	**0.2865**	**0.0128**	**0.0777**

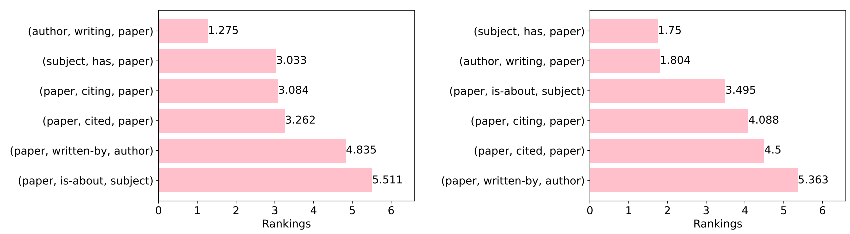
Table 2: Ablation Study for Meta Relation Aware Module

		AM		ACM		BGS		AIFB		MUTAG	
		FID↓	AUSFC↓	FID↓	AUSF↓	FID↓	AUSFC↓	FID↓	AUSFC↓	FID↓	AUSFC↓
RGCN	MR w/o g.a.	0.6063	2.8197	0.2186	1.0605	0.0684	0.3137	0.3415	1.6420	0.0155	0.0840
	MR w/o c.a.	0.3630	2.0033	0.1483	0.9426	0.0335	0.1831	0.2937	1.5034	0.0354	0.1611
	MR-Explainer	**0.1067**	**0.7398**	**0.0508**	**0.4367**	**0.0311**	**0.1823**	**0.1307**	**0.7317**	**0.0136**	**0.0787**
RGAT	MR w/o g.a.	0.0016	0.0312	0.1752	1.2617	0.0054	0.0397	0.3247	1.7448	**0.0117**	**0.0744**
	MR w/o c.a.	0.0028	0.0409	0.4687	2.5366	**0.0000**	**0.0115**	0.3587	1.8654	0.0305	0.1625
	MR-Explainer	**0.0013**	**0.0311**	**0.0515**	**0.4554**	0.0069	0.0383	**0.0265**	**0.2865**	0.0128	0.0777

The dataset is split into training, validation, and test sets (4:4:2). The Adam optimizer is utilized. We sample two-hop subgraphs for computational efficiency, applying a sparsity constraint to retain the top k edges for each meta-relation.

5.2 Comparative Experiment

In this experiment, we compare the explanatory models by evaluating their best FID values at a sparsity level of 0.6, along with their AUSFC across sparsity levels between 0 to 0.6. The detailed result is shown in Table 1. For the RGCN backbone, MR-Explainer achieves the lowest FID (0.1067) and AUSFC (0.7398) on the AM dataset, surpassing PGExplainer (FID: 0.3370, AUSFC: 2.1388) and GNNExplainer (FID: 0.3201, AUSFC: 1.9943). This demonstrates MR-Explainer's superior ability to balance fidelity and sparsity, delivering more accurate and concise explanations. While PGExplainer performs well, it does not match the balance that MR-Explainer achieves, highlighting MR-Explainer's ability to generate explanations that align closely with the original predictions while maintaining brevity. For the RGAT backbone, MR-Explainer again leads with the lowest FID (0.0013) and AUSFC (0.0311) on the AM dataset. PGExplainer follows (FID: 0.0016, AUSFC: 0.0476), showing strong performance but still falling behind MR-Explainer in both fidelity and AUSFC. GNNExplainer, while better than the random baseline, lags behind both PGExplainer and MR-Explainer with an FID of 0.0269 on the AM dataset. These results confirm that MR-Explainer consistently provides the best combination of explanation quality, optimizing both fidelity and sparsity for more robust and precise explanations.



(a) RGCN

(b) RGAT

Fig. 3: Meta-Relation Importance Analysis for ACM Dataset.

5.3 Ablation Study

This study evaluates the impact of the multiplex salience-aware module by testing variants "MR w/o g.a." (without global-aware salience) and "MR w/o c.a." (without context-aware salience). As shown in Table 2, MR-Explainer consistently outperforms both variants in all metrics (FID and AUSFC). For example, on the AM dataset with RGCN, MR-Explainer achieves FID $= 0.1067$ and AUSFC $= 0.7398$, while the variants yield worse results. This demonstrates the importance of both global and context-aware salience in generating accurate, comprehensive explanations. Without global salience, edge significance is poorly captured, and without context-aware salience, key details within meta-relations are missed. MR-Explainer's integration of both salience types enhances its explanatory power.

5.4 Meta-Relation Importance Analysis

We evaluate MR-Explainer on the ACM dataset by analyzing global salience values to assess meta-relation importance in HGNN's outcomes. Figure 3 shows that the meta-relation (author, writing, paper) is the most influential for both RGCN and RGAT, while (paper, written-by, author) ranks lower. This underscores MR-Explainer's ability to differentiate the contributions of meta-relations.

6 Conclusion

In conclusion, this paper presents MR-Explainer, a novel framework for explaining HGNNs. MR-Explainer effectively captures both global and meta relation-specific contributions to HGNN's outcomes through a combination of a heterogeneous information fusion module and a multiplex salience-aware module. The fusion module integrates diverse structural and semantic information, while the salience-aware module computes salience values for edges at both the graph and meta relation levels. This approach ensures that the generated explanations are both comprehensive and precise.

Acknowledgements This research is partially supported by the Natural Science Foundation of China (No.92367202, No.62202279), Excellent Youth Science Fund Project (Overseas) of Shandong (No.2023HWYQ-039), Natural Science Foundation of Shandong Province (No.ZR2022QF018, No.ZR2023LZH006).

References

1. Akkas, S., Azad, A.: Gnnshap: Scalable and accurate gnn explanation using shapley values. In: Proc. ACM Web Conf. 2024. pp. 827–838 (2024)
2. Chen, X., Cai, R., Zhu, Y., Horwood, J., Hao, Z., Li, Z., Hernández-Lobato, J.M., et al.: Feature attribution with necessity and sufficiency via dual-stage perturbation test for causal explanation. In: 41st Int. Conf. Mach. Learn. (2024)
3. Dai, E., Zhao, T., Zhu, H., Xu, J., Guo, Z., Liu, H., Tang, J., Wang, S.: A comprehensive survey on trustworthy graph neural networks: Privacy, robustness, fairness, and explainability. Mach. Intell. Res. pp. 1–51 (2024)
4. Esterhuizen, J.A., Goldsmith, B.R., Linic, S.: Interpretable machine learning for knowledge generation in heterogeneous catalysis. Nature Catal. **5**(3), 175–184 (2022)
5. Hu, Z., Dong, Y., Wang, K., Sun, Y.: Heterogeneous graph transformer. In: Proc. Web Conf. 2020. pp. 2704–2710 (2020)
6. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. In: ICLR 2017. OpenReview.net (2017)
7. Luo, D., Cheng, W., Xu, D., Yu, W., Zong, B., Chen, H., Zhang, X.: Parameterized explainer for graph neural networks. Adv. Neural Info. Process. Syst. **33**, 19620–19631 (2020)
8. Ma, Y., Ye, C., Wu, Z., Wang, X., Cao, Y., Chua, T.S.: Context-aware event forecasting via graph disentanglement. In: Proc. 29th ACM SIGKDD Conf. Knowl. Discov. Data Min. pp. 1643–1652 (2023)
9. Pope, P.E., Kolouri, S., Rostami, M., Martin, C.E., Hoffmann, H.: Explainability methods for graph convolutional neural networks. In: Proc. CVPR 2019. pp. 10772–10781 (2019)
10. Schlichtkrull, M., Kipf, T.N., Bloem, P., Van Den Berg, R., Titov, I., Welling, M.: Modeling relational data with graph convolutional networks. In: ESWC 2018. pp. 593–607. Springer (2018)
11. Vu, M., Thai, M.T.: Pgm-explainer: Probabilistic graphical model explanations for graph neural networks. Adv. Neural Info. Process. Syst. **33**, 12225–12235 (2020)
12. Wu, T., Ren, H., Li, P., Leskovec, J.: Graph information bottleneck. Adv. Neural Info. Process. Syst. **33**, 20437–20448 (2020)
13. Wu, Z., Wang, J., Du, H., Jiang, D., Kang, Y., Li, D., Pan, P., Deng, Y., Cao, D., Hsieh, C.Y., et al.: Chemistry-intuitive explanation of graph neural networks for molecular property prediction with substructure masking. Nature Comm. **14**(1), 2585 (2023)
14. Ye, Z., Huang, R., Wu, Q., Liu, Q.: Same: Uncovering gnn black box with structure-aware shapley-based multipiece explanations. Adv. Neural Info. Process. Syst. **36** (2024)
15. Ying, Z., Bourgeois, D., You, J., Zitnik, M., Leskovec, J.: Gnnexplainer: Generating explanations for graph neural networks. Adv. Neural Info. Process. Syst. **32** (2019)
16. Yuan, H., Yu, H., Gui, S., Ji, S.: Explainability in graph neural networks: A taxonomic survey. IEEE Trans. Pattern Anal. Mach. Intell. (2022)