



One Prompt Fits All: Universal Graph Adaptation for Pretrained Models

Yongqi Huang¹, Jitao Zhao¹, Dongxiao He^{1*}, Xiaobao Wang¹, Yawen Li²,
Yuxiao Huang³, Di Jin¹, Zhiyong Feng¹

¹Tianjin University, ²Beijing University of Posts and Telecommunications,

³George Washington University

¹{yqhuang, zjtao, hedongxiao, wangxiaobao, jindi, zyfeng}@tju.edu.cn

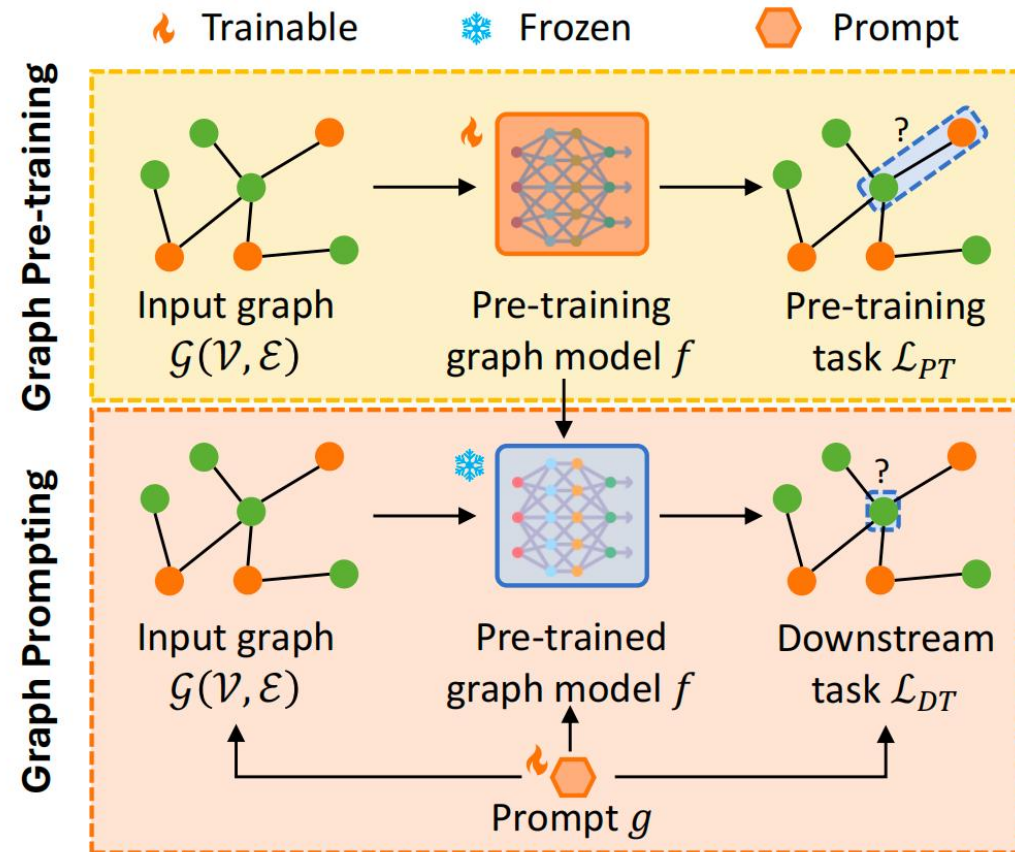
²warmly0716@126.com, ³yuxiaohuang@gwu.edu

Content



- **Background: Graph Prompt Learning**
- **GPL Limitaions**
- **Motivation Experiment**
- **Theoretical Analysis and Discussions**
- **Methods / Experiments**
- **Conclusion**

Background: Graph Prompt Learning



Overview of Graph Prompt Learning^[1].

Graph Prompt Learning (GPL), which aims to design diverse graph prompt strategies, has emerged as a promising and effective alternative paradigm that bridges between graph pretraining and downstream scenarios, which overcomes the limitations of label dependency and the misalignment between upstream pretraining and downstream tasks.

Background: Graph Prompt Learning

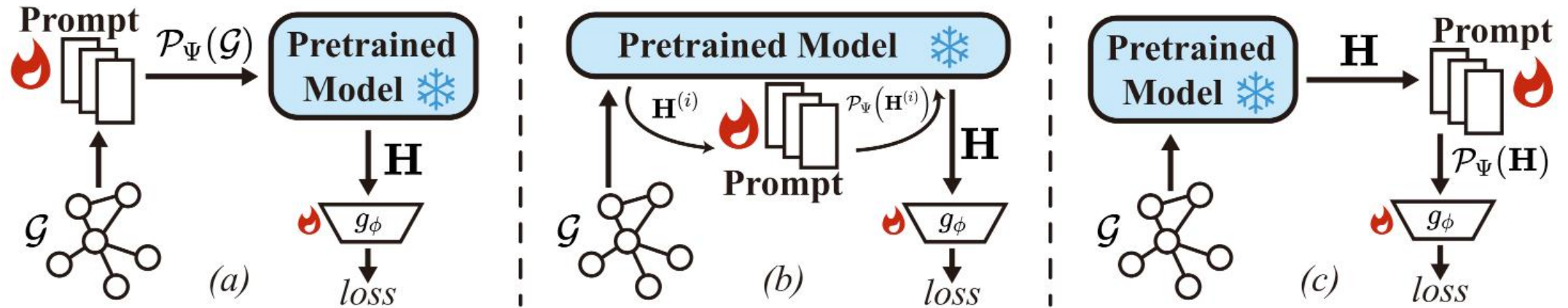


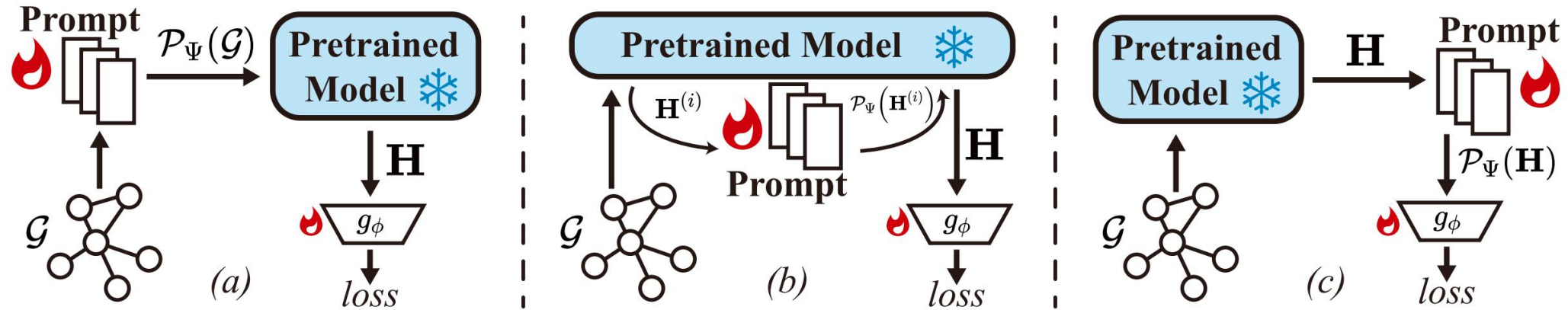
Figure 1: Three different graph prompting mechanisms: input-level prompt (left), layer-wise prompt (middle), and representation-level prompt (right).

Graph Prompt Learning (GPL) has emerged as a promising paradigm that bridges graph pretraining models and downstream scenarios, mitigating label dependency and the misalignment between upstream pretraining and downstream tasks. However, their effectiveness and underlying principles remain unclear.

Limitations

Through our analysis of existing methods, we identify two major issues:

(1). Lack of consensus on underlying mechanisms



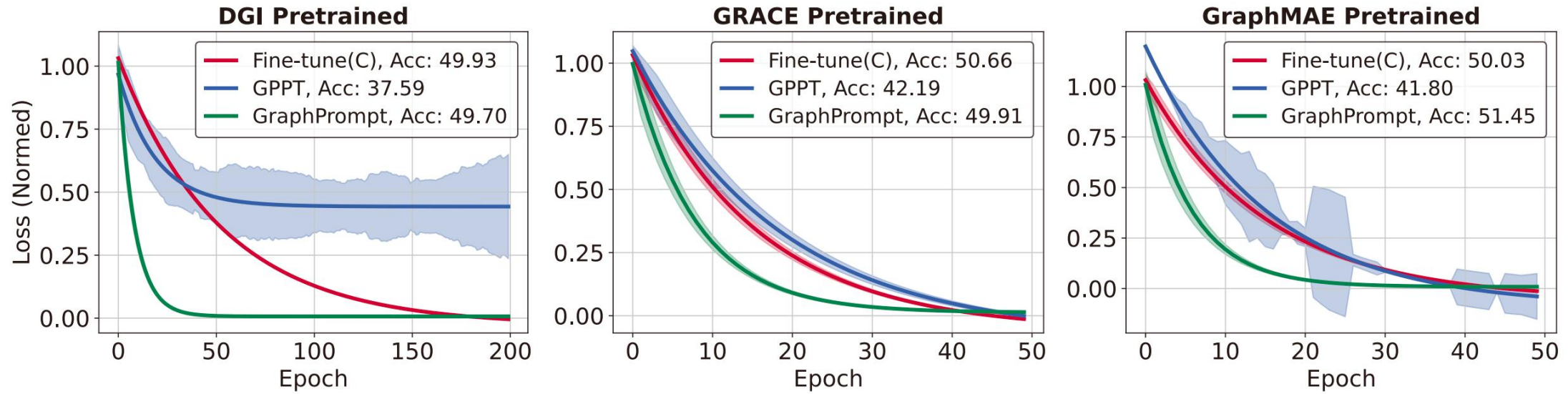
There is no consensus on how prompts interact with pretrained models, as different strategies intervene at varying spaces within the model, i.e., input-level, layer-wise, and representation-level prompts.

(2). Limited scenario adaptability

Most methods fail to generalize across diverse downstream scenarios, especially under data distribution shifts (e.g., homophilic-to-heterophilic graphs).

From the prompting mechanism to the downstream scenario, existing graph prompt learning methods exhibit an adaptation gap.

Motivation Experiment



Existing representation-level prompt GPLs fail to consistently adapt well to different pretrained models. Moreover, they show no significant performance improvement compared to linear probe (only fine-tune the classifier), which achieves good and stable results. **This motivates us to explore the relationship between different types of prompts and linear probe.**



Theoretical Analysis and Discussions

Theorem 4.1 (Parameter Objective Equivalence) *Given a linear prompt function $T(\mathbf{h}) = \mathbf{W}_T \mathbf{h} + \mathbf{b}_T$ and classifier $C(\mathbf{h}) = \mathbf{W}_C^\top \mathbf{h}$, the following properties hold:*

1. **Function Space Equivalence:** *There exists a linear classifier $C'(\mathbf{h}) = \mathbf{W}_{C'}^\top \mathbf{h} + \mathbf{b}_{C'}$ such that $(C \circ T)(\mathbf{h}) = C'(\mathbf{h})$ for all \mathbf{h} ;*
2. **Optimization Objective Equivalence:** *The optimization problems $\min_{\mathbf{W}_T, \mathbf{b}_T, \mathbf{W}_C} L(C \circ T(\mathbf{h}), y)$ and $\min_{\mathbf{W}_{C'}, \mathbf{b}_{C'}} L(C'(\mathbf{h}), y)$ are equivalent in parameter space and gradient update paths.*

The function space equivalence is guaranteed by Proposition [4.1](#), and the optimization equivalence is demonstrated in Proposition [4.2](#).

Proposition 4.1 (Function Space Equivalence) *For any linear transformation $T(\mathbf{h}) = \mathbf{W}_T \mathbf{h} + \mathbf{b}_T$ and classifier $C(\mathbf{h}) = \mathbf{W}_C^\top \mathbf{h}$, there exists an equivalent classifier $C'(\mathbf{h}) = \mathbf{W}_{C'}^\top \mathbf{h} + \mathbf{b}_{C'}$ such that $(C \circ T)(\mathbf{h}) = C'(\mathbf{h})$.*

Proposition 4.2 (Optimization Objective Equivalence) *For $(C \circ T)(\mathbf{h})$ and $C'(\mathbf{h})$, we consider the same loss function L , the optimization problems $\min_{\mathbf{W}_T, \mathbf{W}_C, \mathbf{b}_T} L((C \circ T)(\mathbf{h}), y)$ and $\min_{\mathbf{W}_{C'}, \mathbf{b}_{C'}} L(C'(\mathbf{h}), y)$ are equivalent in the parameter space.*

Method

① Prompt Initialization.

we propose an edge prompt strategy that uses kNN to generate a topological prompt with tunable edge weights

$$(\tilde{\mathbf{A}}_{\text{init}})_{ij} = \begin{cases} \mathbf{S}_{ij}, & \text{if } \mathbf{S}_{ij} \in \text{top-k } \{\mathbf{S}_{i\cdot}\}, \\ 0, & \text{otherwise.} \end{cases}, \quad \mathbf{S}_{ij} = \frac{\mathbf{x}_i \mathbf{x}_j^\top}{\|\mathbf{x}_i\|_2 \|\mathbf{x}_j\|_2},$$

② Parameterization.

we use a learnable scalar weight w_{ij} , which forms the set of prompt parameters $\Psi = \{w_{ij}\}$. To enable the model to select the most relevant prompt edges and ensure non-negative weights, we apply a gating mechanism using a scaled and shifted ELU activation function.

$$\tilde{\mathbf{A}}_{ij} = \text{ELU}(w_{ij} \cdot \alpha - \alpha) + 1,$$

③ Bootstrapped Prompt Integration.

we use bootstrap to balance between original and prompt topology, preserving severe overfitting and model collapse. $\hat{\mathbf{A}}^{(t)} = \tau \hat{\mathbf{A}}^{(t-1)} + (1 - \tau) \tilde{\mathbf{A}},$

④ Optimization Objective.

$$\min_{\phi, \Psi} \frac{1}{|\mathcal{V}_L|} \sum_{v_i \in \mathcal{V}_L} \ell_D(g_\phi(f_\theta(p_\Psi(\mathbf{A}, \mathbf{X}))_i), y_i),$$

Experiments

Pretrain	Methods	Cora	CiteSeer	PubMed	Cornell	Texas	Wisconsin	Chameleon	Actor	Squirrel
DGI	Fine-tune	50.22 \pm 9.28	42.58 \pm 8.87	53.90 \pm 8.30	35.23 \pm 8.84	37.50 \pm 13.57	33.91 \pm 10.56	24.42 \pm 3.19	21.36 \pm 3.28	22.27 \pm 4.10
	Linear-probe	49.77 \pm 9.74	43.16 \pm 7.60	55.76 \pm 9.43	34.56 \pm 8.60	36.21 \pm 13.77	28.71 \pm 9.38	23.64 \pm 2.17	21.33 \pm 2.62	22.82 \pm 4.10
	GPPT	37.59 \pm 7.38	36.01 \pm 6.33	51.56 \pm 6.64	29.01 \pm 8.32	31.26 \pm 8.51	28.56 \pm 6.50	22.15 \pm 2.50	19.81 \pm 1.63	20.71 \pm 1.24
	GraphPrompt	49.70 \pm 10.27	43.98 \pm 7.61	46.32 \pm 7.80	22.29 \pm 6.44	27.62 \pm 11.08	22.62 \pm 8.14	23.59 \pm 2.54	19.84 \pm 2.79	22.85 \pm 3.28
	All-in-one	32.10 \pm 6.50	28.77 \pm 3.12	35.87 \pm 7.53	26.67 \pm 12.42	31.53 \pm 13.14	24.82 \pm 8.77	22.41 \pm 3.58	19.93 \pm 5.23	21.61 \pm 5.87
	GPF	51.68 \pm 9.52	43.11 \pm 5.76	53.09 \pm 9.66	26.76 \pm 8.87	34.04 \pm 15.54	26.59 \pm 8.94	23.29 \pm 3.67	20.31 \pm 4.17	21.66 \pm 3.28
	GPF+	48.66 \pm 6.80	44.89 \pm 6.61	52.58 \pm 9.79	25.23 \pm 8.76	28.55 \pm 13.49	22.82 \pm 8.89	22.98 \pm 3.66	20.81 \pm 3.08	21.56 \pm 3.68
	EdgePrompt	42.05 \pm 6.36	38.54 \pm 6.37	47.67 \pm 4.73	28.00 \pm 8.51	31.32 \pm 15.82	32.64 \pm 11.87	23.17 \pm 3.78	21.36 \pm 2.76	21.99 \pm 2.50
	EdgePrompt+	41.74 \pm 6.73	36.10 \pm 6.15	46.73 \pm 5.53	28.37 \pm 7.94	33.75 \pm 13.57	33.38 \pm 11.81	22.95 \pm 3.78	20.16 \pm 2.65	21.74 \pm 2.10
	UniPrompt	49.95 \pm 10.48	45.57 \pm 8.63	57.17 \pm 7.11	51.13 \pm 13.26	48.21 \pm 15.95	58.75 \pm 13.41	23.75 \pm 4.02	25.38 \pm 4.86	24.20 \pm 2.35
GRACE	Fine-tune	48.59 \pm 9.20	46.16 \pm 6.30	57.97 \pm 7.55	34.18 \pm 10.18	31.52 \pm 13.08	32.23 \pm 8.96	26.22 \pm 2.73	20.81 \pm 2.86	21.16 \pm 2.57
	Linear-probe	46.22 \pm 7.92	46.10 \pm 6.32	57.87 \pm 7.60	34.92 \pm 9.74	34.84 \pm 15.65	31.66 \pm 8.18	24.27 \pm 3.84	20.53 \pm 3.11	20.81 \pm 1.82
	GPPT	42.19 \pm 6.42	37.42 \pm 9.10	47.62 \pm 7.86	27.88 \pm 8.09	32.97 \pm 13.84	26.53 \pm 8.72	25.46 \pm 5.43	19.20 \pm 4.16	21.56 \pm 2.30
	GraphPrompt	49.91 \pm 9.60	35.64 \pm 8.35	53.63 \pm 9.01	23.20 \pm 5.83	30.19 \pm 13.63	23.07 \pm 6.73	28.28 \pm 4.38	19.15 \pm 3.39	22.48 \pm 2.66
	All-in-one	34.53 \pm 5.86	24.06 \pm 6.18	34.51 \pm 7.45	22.17 \pm 5.40	27.37 \pm 13.79	36.17 \pm 6.32	19.46 \pm 0.29	19.04 \pm 4.30	22.03 \pm 2.46
	GPF	48.41 \pm 8.17	36.78 \pm 4.96	50.59 \pm 7.18	28.21 \pm 8.25	29.98 \pm 14.44	27.58 \pm 5.74	25.25 \pm 4.33	20.20 \pm 2.65	20.80 \pm 3.05
	GPF+	47.06 \pm 8.14	44.46 \pm 6.76	51.38 \pm 7.19	28.91 \pm 8.85	31.49 \pm 14.92	27.49 \pm 8.38	26.03 \pm 4.37	20.13 \pm 2.90	21.41 \pm 2.96
	EdgePrompt	41.95 \pm 8.19	36.65 \pm 6.07	48.20 \pm 10.08	31.85 \pm 6.19	29.27 \pm 11.99	38.62 \pm 8.25	23.23 \pm 3.25	20.78 \pm 2.67	21.76 \pm 1.66
	EdgePrompt+	45.32 \pm 9.03	35.80 \pm 6.37	50.01 \pm 11.96	32.13 \pm 7.42	31.95 \pm 6.51	38.68 \pm 7.78	23.79 \pm 3.31	20.63 \pm 2.95	20.97 \pm 1.06
	UniPrompt	44.73 \pm 10.78	47.53 \pm 10.13	57.88 \pm 4.80	52.80 \pm 11.08	45.38 \pm 19.87	50.98 \pm 15.38	26.67 \pm 2.51	26.23 \pm 4.46	23.98 \pm 2.53
GraphMAE	Fine-tune	45.92 \pm 9.67	36.47 \pm 8.35	54.29 \pm 9.52	35.82 \pm 11.30	37.07 \pm 14.08	33.54 \pm 10.16	22.08 \pm 3.19	20.85 \pm 1.68	21.32 \pm 2.65
	Linear-probe	50.13 \pm 12.06	48.08 \pm 6.96	58.61 \pm 8.34	32.27 \pm 11.28	38.32 \pm 13.61	28.40 \pm 8.67	23.02 \pm 2.08	20.56 \pm 2.91	21.05 \pm 1.87
	GPPT	41.80 \pm 8.72	31.96 \pm 5.26	49.10 \pm 8.06	26.74 \pm 7.86	35.16 \pm 15.12	25.86 \pm 8.65	21.87 \pm 3.25	19.36 \pm 3.72	20.59 \pm 1.80
	GraphPrompt	51.45 \pm 9.63	37.07 \pm 6.19	50.87 \pm 6.84	23.82 \pm 7.50	26.04 \pm 11.72	26.78 \pm 9.77	22.05 \pm 2.61	17.82 \pm 2.84	20.71 \pm 4.21
	All-in-one	28.96 \pm 4.87	31.72 \pm 2.78	39.99 \pm 6.21	22.33 \pm 6.43	29.71 \pm 20.15	29.83 \pm 13.99	20.13 \pm 1.81	21.08 \pm 2.17	20.39 \pm 0.93
	GPF	46.74 \pm 8.50	40.07 \pm 8.34	55.38 \pm 7.53	27.21 \pm 7.70	28.98 \pm 14.02	25.65 \pm 8.15	22.30 \pm 2.58	20.20 \pm 3.78	20.19 \pm 0.80
	GPF+	43.30 \pm 11.40	40.15 \pm 6.79	52.92 \pm 7.95	26.38 \pm 8.48	34.83 \pm 16.64	26.79 \pm 9.14	22.35 \pm 3.60	20.44 \pm 3.64	20.26 \pm 0.57
	EdgePrompt	39.16 \pm 9.95	35.03 \pm 6.90	49.79 \pm 7.47	25.26 \pm 7.20	35.02 \pm 16.61	26.02 \pm 8.60	22.27 \pm 3.90	19.93 \pm 3.19	20.16 \pm 1.09
	EdgePrompt+	40.11 \pm 10.12	37.13 \pm 6.93	50.77 \pm 7.91	26.15 \pm 7.77	34.21 \pm 15.55	25.84 \pm 9.35	22.47 \pm 3.82	20.20 \pm 3.00	20.73 \pm 1.10
	UniPrompt	47.05 \pm 9.17	49.29 \pm 11.20	57.47 \pm 6.86	51.28 \pm 12.45	49.83 \pm 17.85	61.38 \pm 13.58	24.29 \pm 3.64	23.35 \pm 3.57	22.08 \pm 2.03

▲Table 1: In-domain 1-shot node classification over different pretrained models.

Methods	Cora	Citeseer	PubMed	Cornell	Squirrel	Chameleon
GCN*	28.57 \pm 5.07	31.27 \pm 4.53	40.55 \pm 5.65	31.81 \pm 4.71	20.00 \pm 0.29	24.17 \pm 5.21
GAT*	28.40 \pm 6.25	30.76 \pm 5.40	39.99 \pm 4.96	28.03 \pm 13.19	21.55 \pm 2.30	23.93 \pm 4.11
DGI*	29.30 \pm 5.82	30.03 \pm 4.88	41.85 \pm 7.78	31.54 \pm 15.66	21.15 \pm 1.68	21.73 \pm 5.47
GraphCL*	34.94 \pm 6.49	30.58 \pm 4.58	40.37 \pm 7.81	27.15 \pm 12.64	21.42 \pm 2.23	22.49 \pm 3.02
GPPT*	17.52 \pm 5.52	21.45 \pm 3.45	36.56 \pm 5.31	25.09 \pm 2.92	20.09 \pm 0.91	24.53 \pm 2.55
GPF*	37.84 \pm 11.07	37.61 \pm 8.87	46.36 \pm 7.48	34.54 \pm 7.73	21.92 \pm 3.50	25.90 \pm 8.51
GCOPE*	34.23 \pm 8.16	39.05 \pm 8.82	44.85 \pm 6.72	34.02 \pm 11.94	22.46 \pm 1.96	24.61 \pm 3.99
MDGPT*	39.54 \pm 9.02	39.24 \pm 8.95	45.39 \pm 11.01	33.58 \pm 10.38	22.35 \pm 3.77	23.68 \pm 1.56
MDGFM*	44.83 \pm 7.41	42.18 \pm 6.41	46.84 \pm 7.31	40.77 \pm 5.96	24.30 \pm 3.26	28.36 \pm 3.65
UniPrompt(Ours)	45.37 \pm 9.08	43.25 \pm 9.61	55.01 \pm 3.36	51.58 \pm 9.91	25.29 \pm 3.86	25.14 \pm 5.65

▲Table 2: Cross-domain 1-shot node classification.

Methods	Cora(5)	CiteSeer(5)	Pubmed(5)	Cornell(3)	Squirrel(3)	Chameleon(5)
GCN*	60.15 \pm 5.33	45.54 \pm 4.71	57.82 \pm 8.26	39.53 \pm 13.57	21.61 \pm 4.22	22.09 \pm 0.99
GAT*	59.79 \pm 3.89	50.48 \pm 2.94	57.55 \pm 9.37	34.53 \pm 13.01	20.11 \pm 3.11	20.83 \pm 1.52
DGI*	56.76 \pm 11.29	42.67 \pm 8.98	54.04 \pm 11.59	43.22 \pm 5.84	20.23 \pm 1.12	27.68 \pm 5.21
GraphCL*	61.59 \pm 5.71	47.05 \pm 6.85	58.50 \pm 7.38	32.77 \pm 6.23	21.18 \pm 0.96	27.45 \pm 2.58
GPPT*	43.67 \pm 7.11	47.31 \pm 6.93	40.47 \pm 10.17	34.69 \pm 8.54	22.14 \pm 1.53	28.25 \pm 1.39
GPF*	51.21 \pm 11.44	56.90 \pm 8.84	58.76 \pm 7.70	38.17 \pm 8.15	21.62 \pm 3.10	28.09 \pm 4.93
GCOPE*	54.63 \pm 3.98	53.18 \pm 4.47	57.74 \pm 2.73	48.21 \pm 11.97	21.37 \pm 4.20	25.50 \pm 1.23
MDGPT*	59.64 \pm 5.73	52.71 \pm 5.71	58.65 \pm 7.54	35.18 \pm 8.90	21.42 \pm 4.16	26.18 \pm 5.18
MDGFM*	64.56 \pm 7.29	61.24 \pm 4.82	63.50 \pm 5.81	49.56 \pm 6.92	23.00 \pm 4.39	30.54 \pm 2.87
UniPrompt(Ours)	65.64 \pm 3.53	59.37 \pm 2.78	65.09 \pm 2.51	52.09 \pm 5.37	26.70 \pm 3.78	31.38 \pm 2.67

▲Table 3: Cross-domain 3/5-shot node classification.

Conclusion

- ✓ We identify two key issues in existing GPLs: lack of consensus on underlying mechanisms, and limited scenario adaptability. We propose that **graph prompt learning should focus on unleashing the capability of pretrained models, and the classifier adapts to downstream scenarios.**
- ✓ We propose UniPrompt, **a novel universal GPL method that adapts any pretrained models.** This method leverages a learnable prompt graph while preserving the original structure to unleash the capability of pretrained models.
- ✓ We conduct **extensive experiments on homophilic and heterophilic datasets, evaluating in-domain and cross-domain performance under few-shot settings.** Experimental results demonstrate that our method consistently outperforms state-of-the-art GPL baselines.

Thank you for listening!

Mails: yqhuang@tju.edu.cn

Code: <https://github.com/hedongxiao-tju/UniPrompt>



My Wechat!