GraphPrompt: Unifying Pre-Training and Downstream Tasks for Graph Neural Networks

Zemin Liu[1*], Xingtong Yu[2*], Yuan Fang[3†], Xinming Zhang [2†]

[1] National University of Singapore, Singapore
[2] University of Science and Technology of China, China
[3] Singapore Management University, Singapore

In Proceeding of THE WEB CONFERENCE, APRIL 30 - MAY 4, 2023

* Co-first author
† Corresponding author

Outline

1 .Motivation

2 .Challenges

3 .Proposed Model: GraphPrompt

4 .Experiment

5 .Conclusions

Motivation

Problem1:

- task-specific labeled data is often difficult or costly to obtain

Problem2:

- pre-training step aims to preserve various intrinsic graph properties
- fine-tuning step aims to reduce the downstream task loss

GNNs' performance heavily depends on labeled data[1,2]

Scarce of labeled data

Pre-Training+Finetuning [3,4]

Gap between pre-train and downstream tasks[5]

Pre-Training+Prompt

[1] Will Hamilton et.al. 2017. Inductive representation learning on large graphs. NIPS.
[2] Thomas N Kipf and Max Welling. 2017. Semi-supervised classification with graph convolutional networks. ICLR.
[3] Weihua Hu et.al. 2020. Strategies for Pre-training Graph Neural Networks. ICLR.
[4] Ziniu Hu et.al. 2020. GPT-GNN: Generative pre-training of graph neural networks. KDD.
[5] Pengfei Liu et.al. 2021. Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. ACM Computing Survey.

Outline

1 .Motivation

2 .Challenges

3 .Proposed Model: GraphPrompt

4 .Experiment

5 .Conclusions

Challenges

Challenges

- Different downstream tasks often have different objectives[6]
- Distinction between various downstream tasks

C1: How to unify pre-training with various downstream tasks on graph?
C2: How to design prompts on graphs?[7]

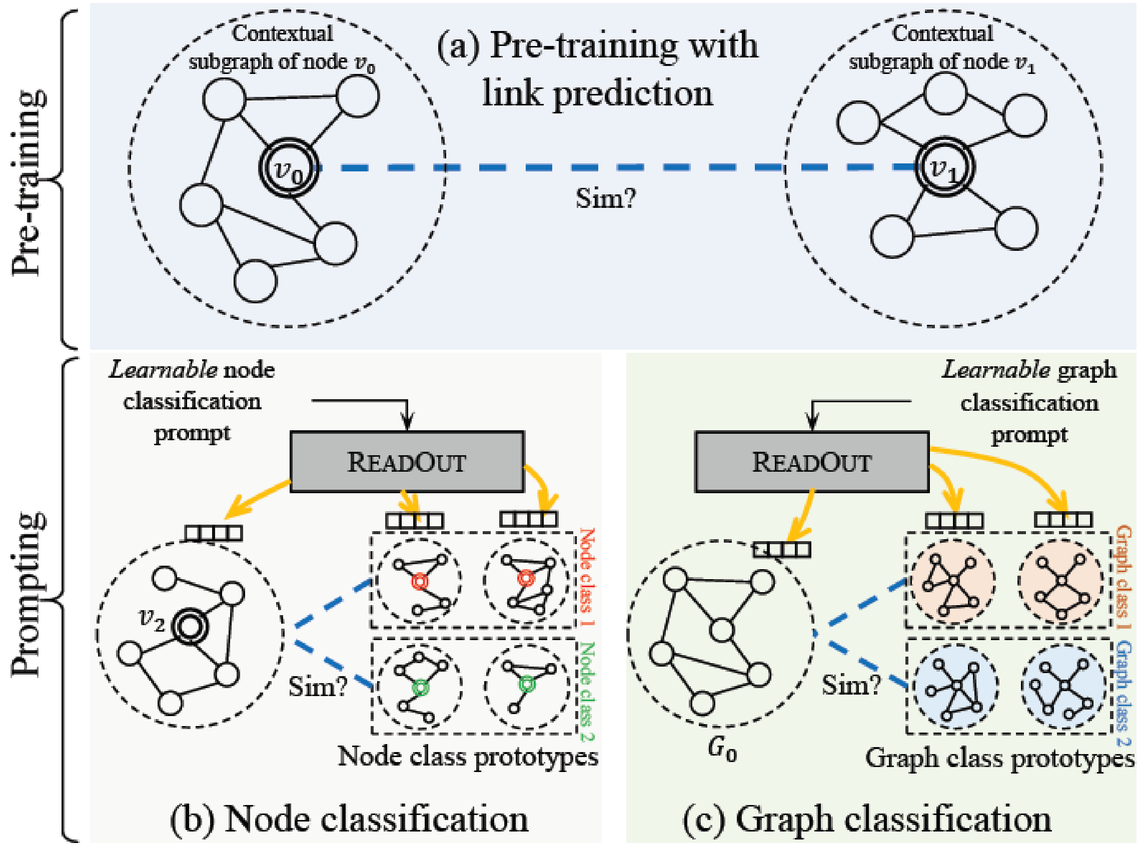


Figure 1: Illustration of the motivation. (a) Pre-training on graphs. (b/c) Downstream node/graph classification.

[6] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. NeurIPS.
[7] Mingchen Sun, Kaixiong Zhou, Xin He, Ying Wang, and Xin Wang. 2022. GPPT: Graph Pre-training and Prompt Tuning to Generalize Graph Neural Networks. SIGKDD

Outline

1 .Motivation
2 .Challenges
3 .Proposed Model: GraphPrompt
4 .Experiment
5 .Conclusions

Proposed Method: GraphPrompt

Unified task template

Link Prediction

$$\mathrm{sim}(\mathbf{s}_v, \mathbf{s}_a) > \mathrm{sim}(\mathbf{s}_v, \mathbf{s}_b)$$

Node Classification(NC)

$$\tilde{\mathbf{s}}_c = \frac{1}{k}\sum_{(v_i,\ell_i)\in D, \ell_i=c} \mathbf{s}_{v_i}$$

$$\ell_j = \arg\max_{c\in C} \mathrm{sim}(\mathbf{s}_{v_j}, \tilde{\mathbf{s}}_c)$$

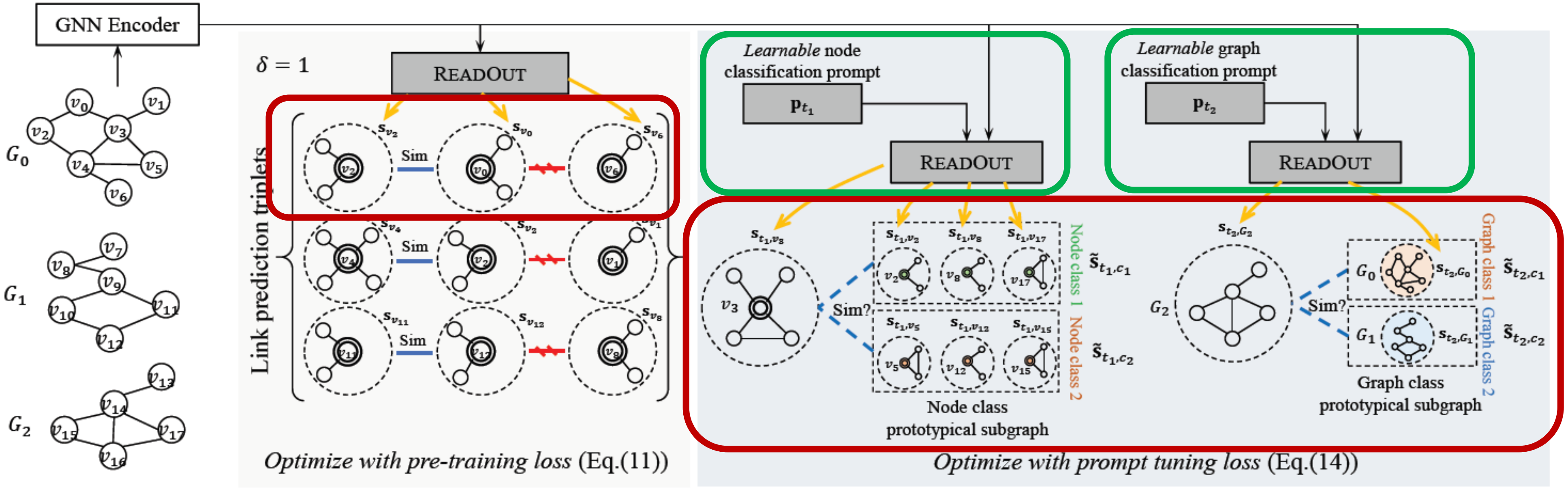
Graph Classification(GC)

$$\tilde{\mathbf{s}}_c = \frac{1}{k}\sum_{(G_i,L_i)\in \mathcal{D}, L_i=c} \mathbf{s}_{G_i}$$

$$L_j = \arg\max_{c\in C} \mathrm{sim}(\mathbf{s}_{G_j}, \tilde{\mathbf{s}}_c)$$

mean embedding of (sub)graphs

class label



(a) Toy graphs (b) Pre-training (c) Prompting for node classification (left) or graph classification (right)

Figure 2: Overall framework of GraphPrompt.

A Notation for NC and GC

$$y = \arg\max_{c\in Y} \mathrm{sim}(\mathbf{s}_x, \tilde{\mathbf{s}}_c)$$

$$\mathbf{s}_x = \textsc{ReadOut}(\{\mathbf{h}_v : v \in V(S_x)\})$$

Pre-Training Objective

$$\mathcal{L}_{\text{pre}}(\Theta) = -\sum_{(v,a,b)\in\mathcal{T}_{\text{pre}}} \ln \frac{\exp(\mathrm{sim}(\mathbf{s}_v, \mathbf{s}_a)/\tau)}{\sum_{u\in\{a,b\}}\exp(\mathrm{sim}(\mathbf{s}_v, \mathbf{s}_u)/\tau)}$$

Prompt Design

$$\mathbf{s}_{t,x} = \textsc{ReadOut}(\{\mathbf{p}_t \odot \mathbf{h}_v : v \in V(S_x)\})$$

Outline

1 .Motivation

2 .Challenges

3 .Proposed Model: GraphPrompt

4 .Experiment

5 .Conclusions

Experiment

Node Classification and Graph Classification

Table 2: Accuracy evaluation on node classification.

All tabular results are in percent, with best **bolded** and runner-up underlined.

Methods	Flickr 50-shot	PROTEINS 1-shot	ENZYMES 1-shot
GCN	9.22 ± 9.49	59.60 ± 12.44	61.49 ± 12.87
GraphSAGE	13.52 ± 11.28	59.12 ± 12.14	61.81 ± 13.19
GAT	16.02 ± 12.72	58.14 ± 12.05	60.77 ± 13.21
GIN	10.18 ± 5.41	60.53 ± 12.19	63.81 ± 11.28
DGI	17.71 ± 1.09	54.92 ± 18.46	63.33 ± 18.13
GraphCL	18.37 ± 1.72	52.00 ± 15.83	58.73 ± 16.47
GPPT	18.95 ± 1.92	50.83 ± 16.56	53.79 ± 17.46
GraphPrompt	**20.21 ± 11.52**	**63.03 ± 12.14**	**67.04 ± 11.48**

Table 3: Accuracy evaluation on graph classification.

Methods	PROTEINS 5-shot	COX2 5-shot	ENZYMES 5-shot	BZR 5-shot
GCN	54.87 ± 11.20	51.37 ± 11.06	20.37 ± 5.24	56.16 ± 11.07
GraphSAGE	52.99 ± 10.57	52.87 ± 11.46	18.31 ± 6.22	57.23 ± 10.95
GAT	48.78 ± 18.46	51.20 ± 27.93	15.90 ± 4.13	53.19 ± 20.61
GIN	58.17 ± 8.58	51.89 ± 8.71	20.34 ± 5.01	57.45 ± 10.54
InfoGraph	54.12 ± 8.20	54.04 ± 9.45	20.90 ± 3.32	57.57 ± 9.93
GraphCL	56.38 ± 7.24	55.40 ± 12.04	28.11 ± 4.00	59.22 ± 7.42
GraphPrompt	**64.42 ± 4.37**	**59.21 ± 6.82**	**31.45 ± 4.32**	**61.63 ± 7.68**

- GraphPrompt outperforms all baselines for both node classification task and graph classification task, which implies
  - GraphPrompt is able to narrow the gap between pre-training task and downstream tasks.
  - GraphPrompt could effectively derive the downstream tasks to exploit the pre-trained model in task-specific manner.

Experiment

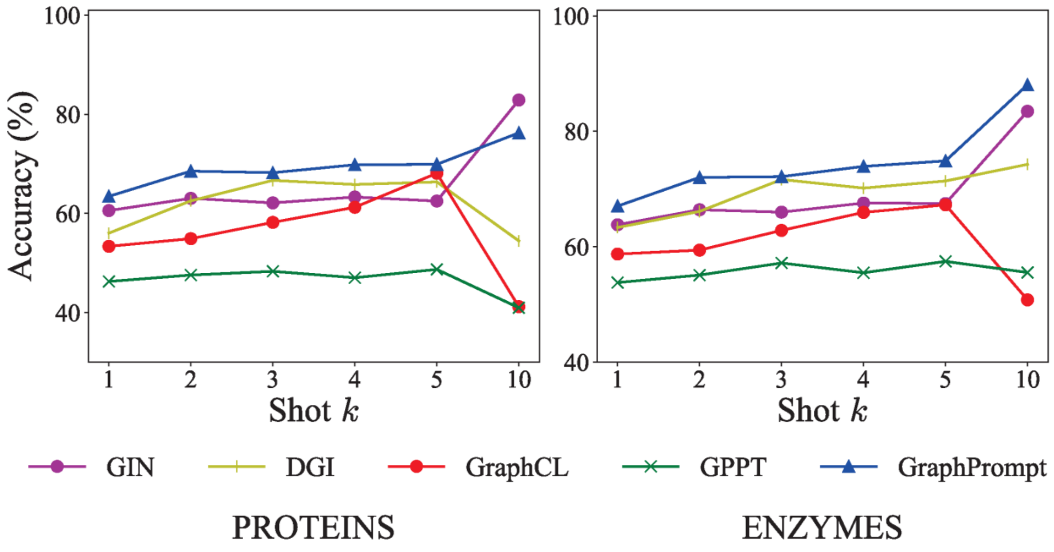


Figure 3: Impact of shots on few-shot node classification.

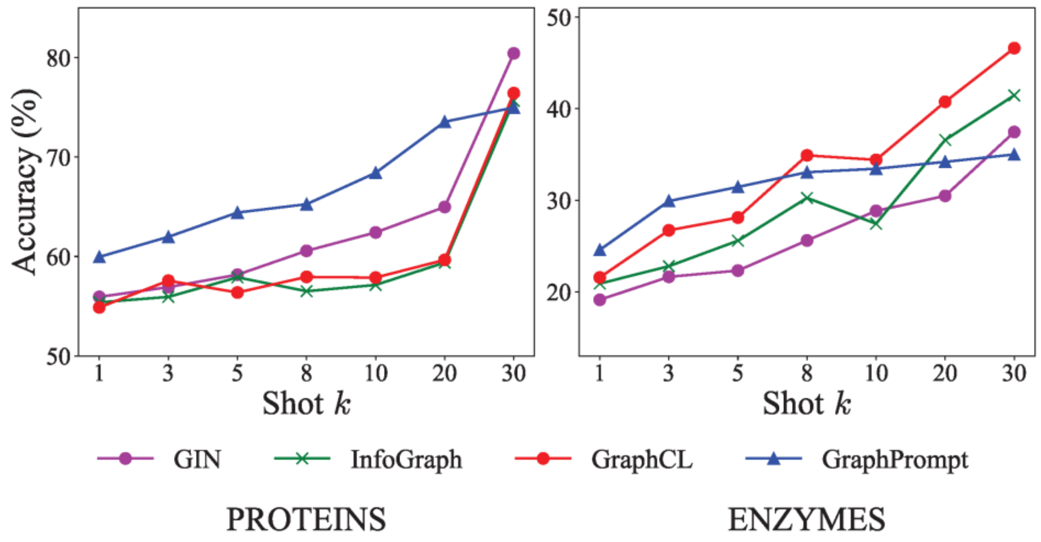


Figure 4: Impact of shots on few-shot graph classification.

- GraphPrompt consistently outperforms the baselines especially with lower shots
- For node classification task, 10 shot is sufficient for semi-supervised learning since graph is small
- For graph classification task, GraphPrompt can be surpassed by some baselines when given more shots

Outline

1 .Motivation

2 .Challenges

3 .Proposed Model: GraphPrompt

4 .Experiment

5 .Conclusions

Conclusions

- **Problem: Pretraining-Prompting**
  - Unify pre-training task and downstream tasks
  - Attain task-specific optima
- **Proposed-Model: GraphPrompt**
  - Unify upstream and downstream tasks via subgraph similarity
  - Using prompt vector to change the feature weights of each dimension of the node embedding to guide subgraph readout
- **Experiment**
  - GraphPrompt outperforms all baselines for both node classification task and graph classification task

Thanks!

Paper, data & code available at https://xingtongyu.netlify.app/

Zemin Liu* ,Xingtong Yu*, Yuan Fang†, Xinming Zhang†

GraphPrompt: Unifying Pre-Training and Downstream Tasks for Graph Neural Networks

n Proceeding of THE WEB CONFERENCE, APRIL 30 - MAY 4, 2023