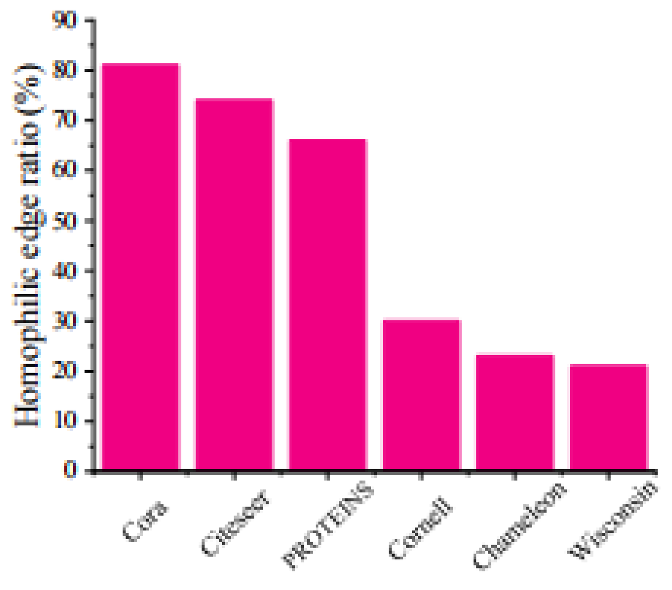
Non-Homophilic Graph Pre-Training and Prompt Learning

Xingtong Yu, Jie Zhang, Renhe Jiang, Yuan Fang

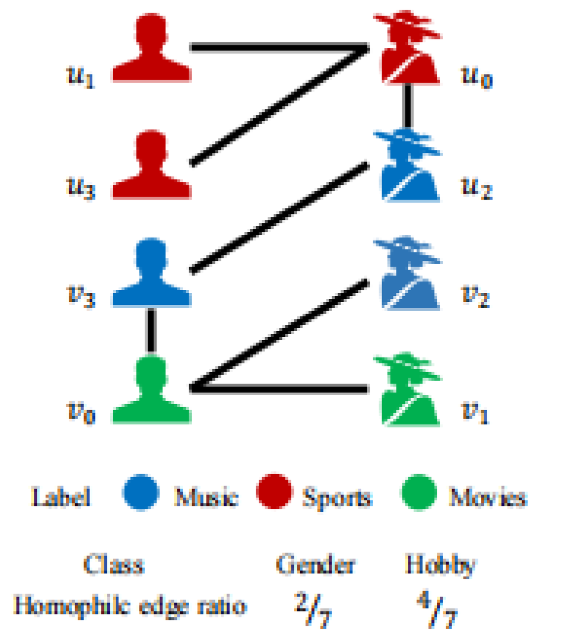
SMU Singapore Management University · NUS National University of Singapore · UTokyo

KDD2025 August 3–7, 2025 Toronto

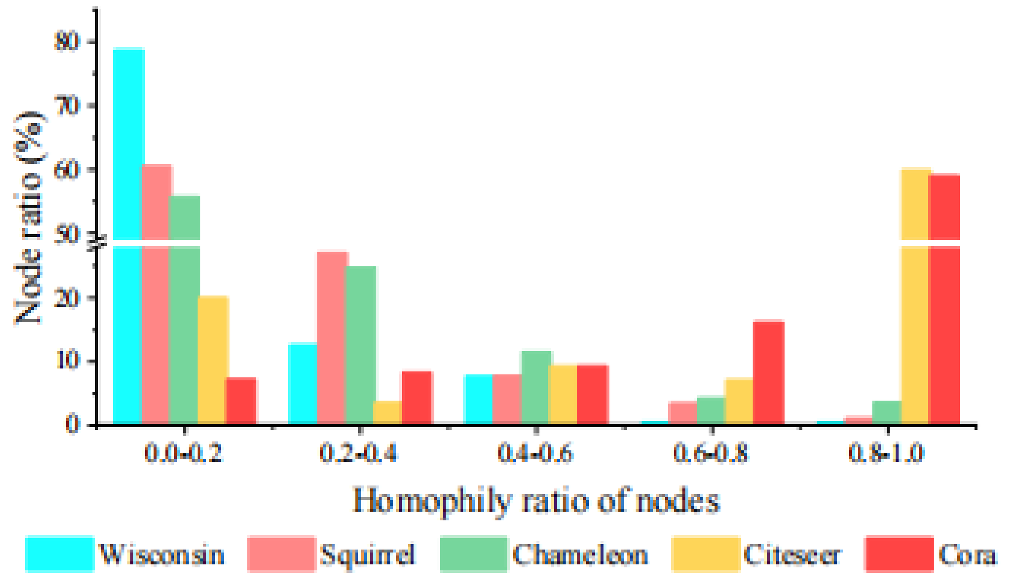
Background



(a) Varying non-homophilic patterns across different graphs



(b) Dependence of homophiliy ratio on the target label



(c) Diverse non-homophilic patterns across nodes in the same graph

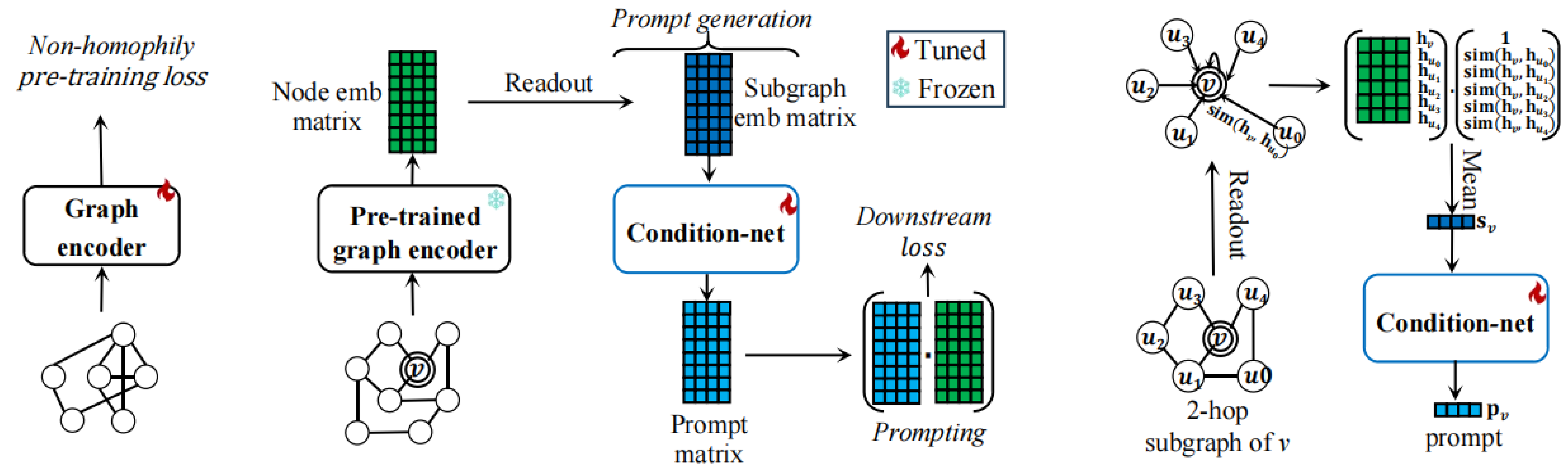
Pre-training and prompt learning have become a popular approach to train Graph Neural Networks (GNNs) without heavy reliance on labeled data.

However, most existing prompt methods do not distinguish between homophilic and heterophilic characteristics in graphs.

Challenges

- How do we pre-train a graph model irrespective of the graph's homophily characteristics?
- How do we capture the fine-grained, node-specific non-homophilic characteristics?

Method: ProNoG



(a) Pre-training **(b) Downstream adaptation with conditional prompting** **(c) Details of prompt generation**

- ProNoG: a novel pre-training and prompt learning framework for non-homophilic graphs.
- In graph pre-training, we pre-train a graph encoder using a non-homophilic pre-training task.
- On downstream tasks, we propose condition-net to generate prompts based on different non-homophilic patterns, allowing for precise adjustments at the node level without updating the weights of the pre-trained model.

Ablation study

Methods	Graph classification: Wisconsin	Squirrel	Chameleon	PROTEINS	ENZYMES
NoPrompt	20.85±6.74	20.18±1.30	22.34±4.15	53.61±8.90	21.85±6.17
SinglePrompt	25.77±6.24	20.68±0.91	27.03±3.98	56.35±10.59	19.38±7.12
NodeCond	25.30±4.62	20.98±1.56	27.24±5.24	**56.61**±10.03	20.70±6.67
ProNoG\sim	22.05±5.86	19.93±0.42	20.20±1.11	52.30±10.94	16.70±1.28
ProNoG	**31.54**±5.30	**20.92**±1.37	**28.50**±5.30	56.11±10.19	**22.55**±6.70

Methods	Node classification: Wisconsin	Squirrel	Chameleon	PROTEINS	ENZYMES
NoPrompt	25.41±3.13	20.60±1.30	22.71±3.54	47.22±11.05	66.59±19.28
SinglePrompt	32.76±5.21	20.85±1.32	22.78±3.35	30.33±19.59	65.32±21.67
NodeCond	35.56±4.65	21.26±3.95	21.13±2.23	36.01±19.70	68.54±19.31
ProNoG\sim	30.65±4.05	20.05±0.59	20.96±4.21	33.73±17.82	36.02±20.64
ProNoG	**44.72**±11.93	**24.59**±3.41	**30.67**±3.73	**48.95**±10.85	**72.94**±20.23

ProNoG consistently outperforms these variants in all but one instance, in which its performance is still competitive. This highlights the necessity of reading out subgraphs with similarity weighting in order to capture the characteristics of each node, and the advantage of using conditional prompt learning to adapt to each node.

Results & Experiments

Methods	Few-shot node classification: Chameleon	Cornell	PROTEINS	ENZYMES	Few-shot graph classification: Chameleon	Cornell	PROTEINS	ENZYMES
GCN	25.11 ± 4.19	21.81 ± 4.71	43.32 ± 9.35	48.08 ± 4.71	17.21 ± 4.80	26.36 ± 4.35	51.66 ± 10.87	19.30 ± 6.36
GAT	24.82 ± 4.35	23.03 ± 13.19	31.79 ± 20.11	35.32 ± 18.72	25.71 ± 3.32	22.66 ± 12.46	51.33 ± 11.02	20.24 ± 6.39
H2GCN	25.89 ± 4.96	32.77 ± 14.88	29.60 ± 6.99	37.27 ± 8.73	<u>26.76</u> ± 3.98	23.11 ± 11.78	53.81 ± 8.85	19.40 ± 5.57
FAGCN	22.71 ± 3.74	28.67 ± 17.64	32.63 ± 9.94	35.87 ± 13.47	25.93 ± 4.03	25.71 ± 13.12	55.45 ± 11.57	19.95 ± 5.94
DGI	19.33 ± 4.57	32.54 ± 15.66	45.22 ± 11.09	48.05 ± 14.83	24.29 ± 4.33	18.60 ± 12.79	50.32 ± 13.47	21.57 ± 5.37
GraphCL	27.16 ± 4.31	24.69 ± 14.06	46.15 ± 10.94	48.88 ± 15.98	26.45 ± 4.30	20.03 ± 10.05	54.81 ± 11.44	19.93 ± 5.65
DSSL	<u>27.92</u> ± 3.93	20.36 ± 5.38	40.42 ± 10.08	<u>66.59</u> ± 19.28	26.19 ± 3.72	18.38 ± 10.63	52.73 ± 10.98	**23.14** ± 6.71
GraphACL	26.72 ± 4.67	<u>33.17</u> ± 16.06	42.16 ± 13.50	47.57 ± 14.36	26.28 ± 3.93	<u>26.50</u> ± 17.18	**56.11** ± 13.95	20.28 ± 5.60
GPPT	24.53 ± 2.55	25.09 ± 2.92	35.15 ± 11.40	35.37 ± 9.37	26.51 ± 4.67	24.06 ± 13.71	53.61 ± 8.90	21.85 ± 6.17
GraphPrompt	25.36 ± 3.99	31.00 ± 13.88	<u>47.22</u> ± 11.05	53.54 ± 15.46	26.03 ± 4.17	25.31 ± 7.65	54.55 ± 12.61	21.85 ± 5.15
GraphPrompt	25.73 ± 4.50	31.65 ± 14.48	46.08 ± 9.96	57.68 ± 13.12				
ProNoG	**30.67** ± 3.73	**37.90** ± 9.31	**48.95** ± 10.85	**72.94** ± 20.23	**28.50** ± 5.30	**27.17** ± 9.58	<u>56.11</u> ± 10.19	<u>22.55</u> ± 6.70

Few-shot node classification — Few-shot graph classification

ProNoG surpasses all baseline methods across all settings in one shot scenarios, outperforming the best competitor by up to 21.49% on node classification and 6.50% on graph classification.

ProNoG significantly outperforms all baselines in low-shot scenarios with very limited labeled data (e.g., $k \le 5$),