TouchUp-G: Improving Feature Representation through Graph-Centric Finetuning

Jing Zhu, Xiang Song, Vassilis Ioannidis, Danai Koutra, Christos Faloutsos



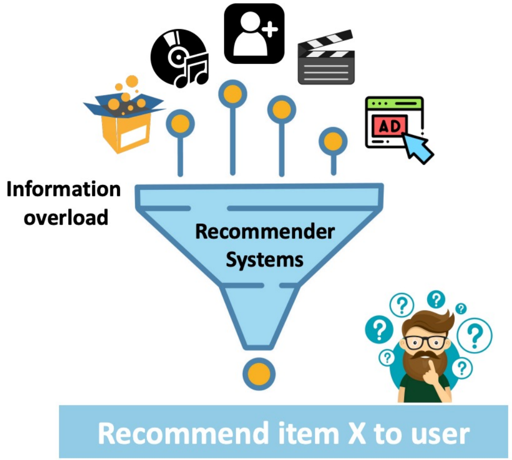


KnowledgeNLP – Aug.6 2023

Recommendation System

Recommendation has been widely applied in online services:
E-commerce, Content Sharing, Social Networking ...

Age of Information Explosion



Items can be: Products, News, Movies, Videos, Friends, etc.

Image credit: Jiliang Tang

Recommendation System as LP

One typical approach for solving this is to cast as a Link Prediction (LP) task for feature-rich graphs.

- Co-purchasing: Given item A, which item tend to be bought together with item A
- Features: Text descriptions, images of candidate items

Product Recommendation

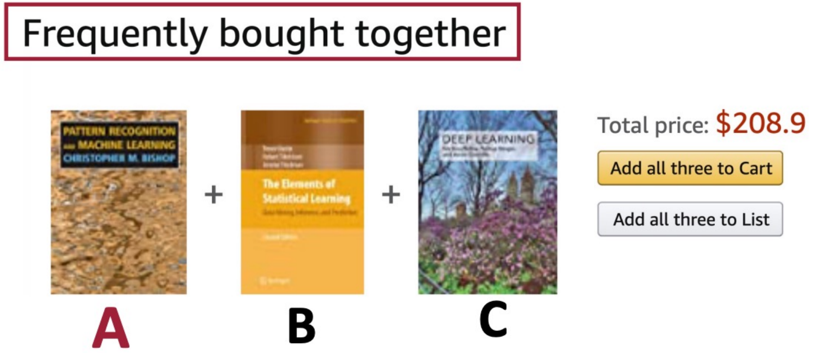
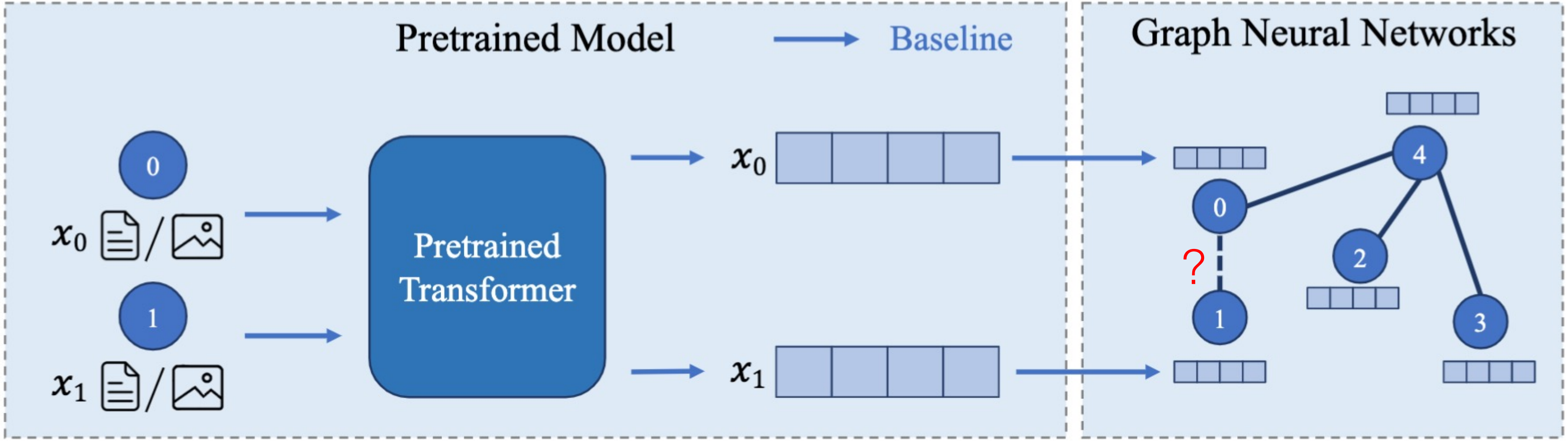


Image credit: Jiliang Tang

Pretrained Models for GNNs

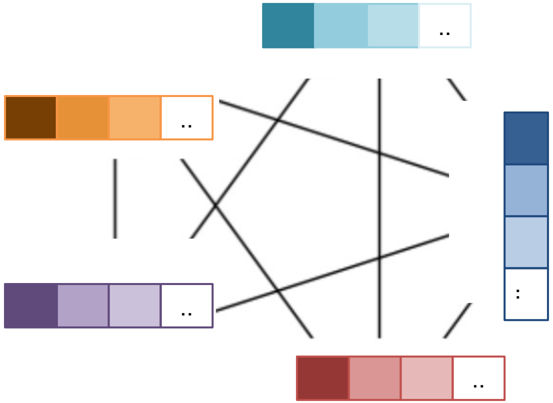
For feature-rich graphs: Pretrained models such as BERT is typically adopted to generate feature embeddings for each node.



Jing Zhu et. al

Why pretrained features may fail

Pretrained Model
(BERT, ViT etc.)

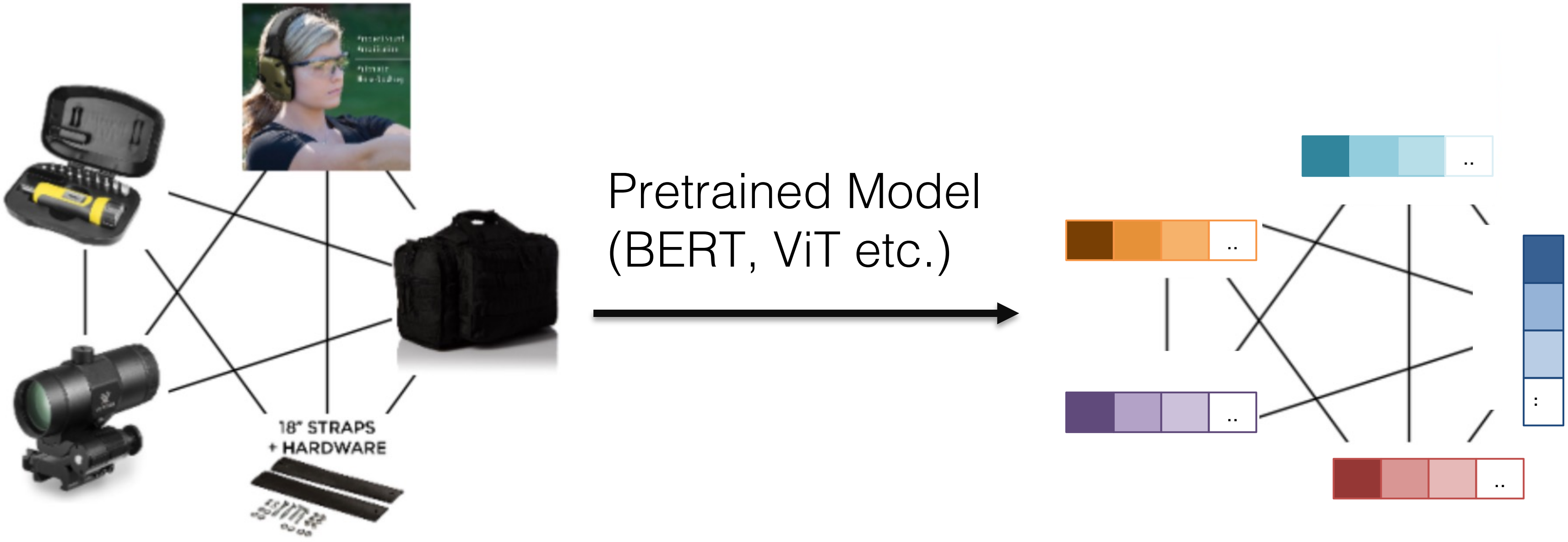


A subgraph of Amazon Co-purchasing graph (Amazon-CP). Products have completely different visual features, but they are often bought together.

NO!

Jing Zhu et. al

Why pretrained features may fail

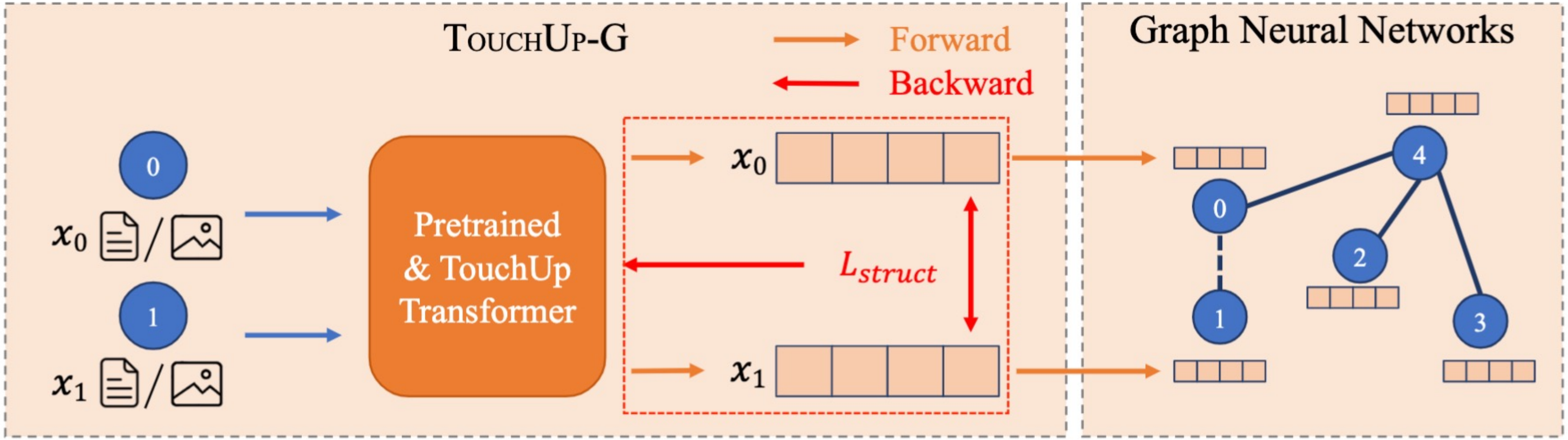


The graph agnostic feature embeddings prevents fully utilize the potential correlations between graph structures and node features.

Research Question

How can we improve node features obtained from Pretrained Models (PMs) for downstream graph tasks such as link prediction?

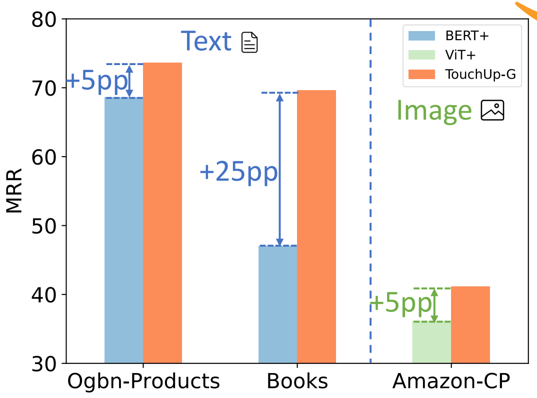
Graph-Centric Finetuning on the PMs!

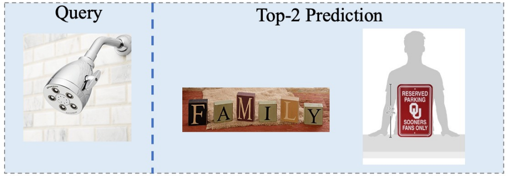


Jing Zhu et. al

TouchUp-G

- General: Can be applied to a variety of graph tasks.
- Multi-modal: Can be applied any pretrained models from any modality, e.g., texts, images etc.
- Principled: Propose a novel metric: feature homophily to measure the correlation between node features and graph structure.
- Effective: Outperforms baselines on 4 real datasets, with up to 2× performance improvement across various tasks, metrics and modalities.





ViT+

TouchUp-G Wins!

Jing Zhu et. al

Previous Works

- Most work focuses on finetuning language models (LMs) improving node representation for node classification on text-rich graphs.

- TouchUp-G can be applied to any pretrained models from any modality, e.g., texts, images, and a variety of graph tasks, including node classification and link prediction.

Property \ Method	GLEM [55]	GIANT [5]	BERT+ [9]	TouchUp-G
General	✓		✓	✓
Multi-modal				✓
Principled		✓		✓
Effective	✓	✓	?	✓

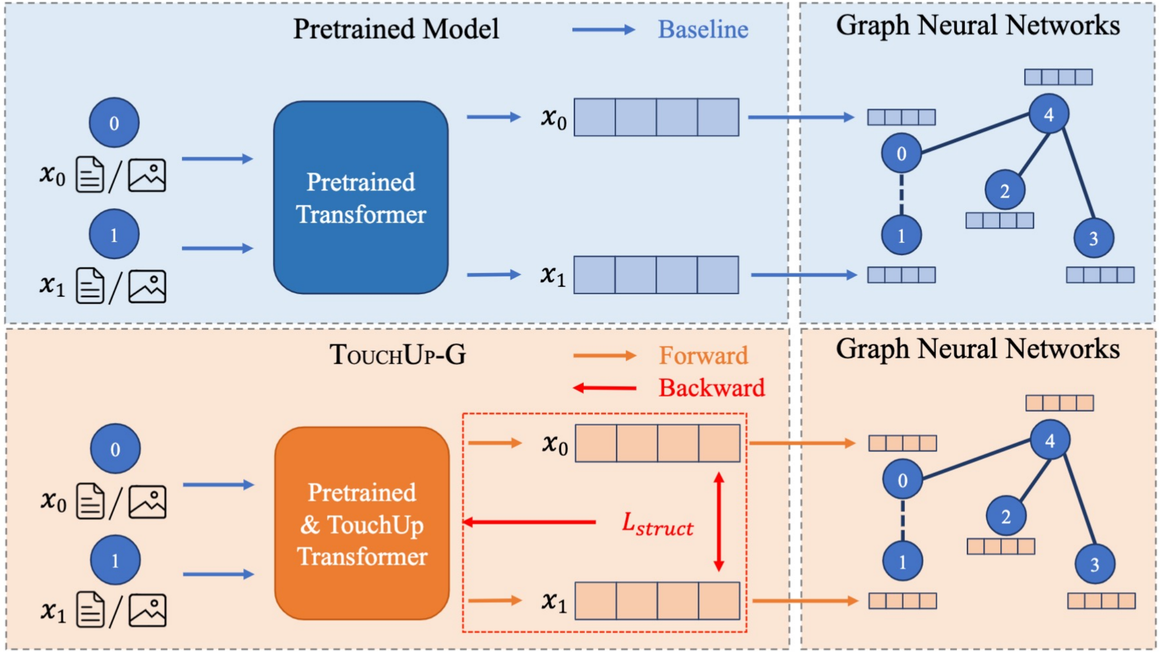
[Chien+ ’22, *Zhao+ ’23*]

TouchUp-G Pipeline

Key Intuition:

Make the node features obtained from the PMs graph-aware (L_{struct}).

Measure the awareness by feature homophily score.

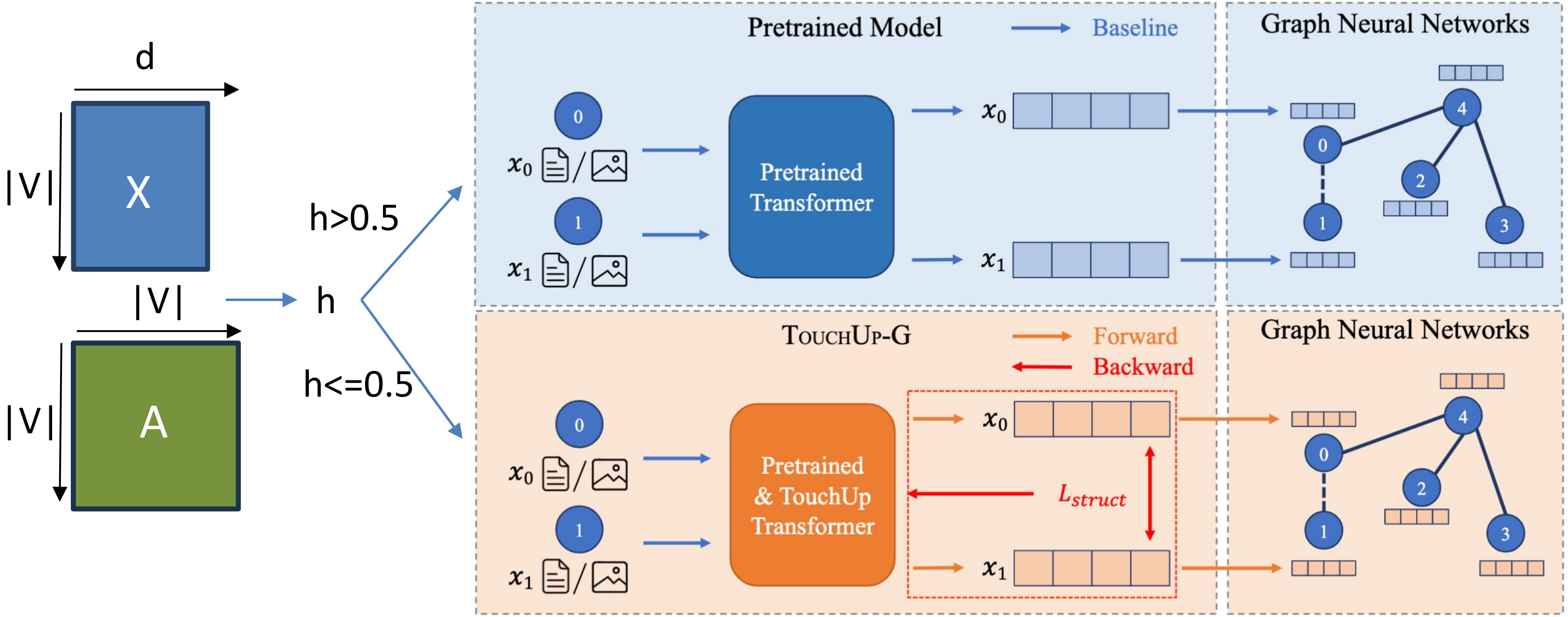


Jing Zhu et. al

TouchUp-G Pipeline

① Feature Homophily h

② Graph-Centric Finetuning

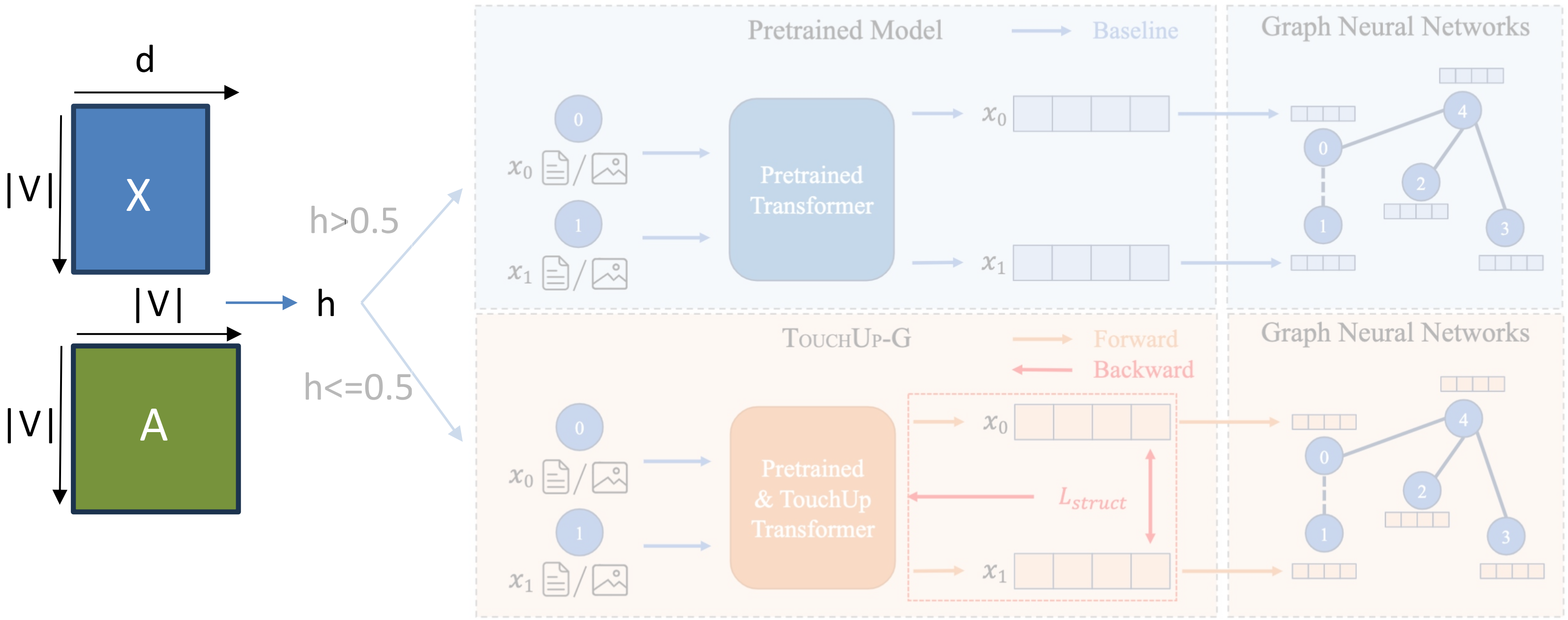


Jing Zhu et. al

TouchUp-G Pipeline

① Feature Homophily h

② Graph-Centric Finetuning



Jing Zhu et. al

Feature Homophily h

Def: The tendency of nodes with similar features to be connected to each other.

Feature Homophily is an implicit assumption shared by most GNNs. [Yang+ '21]

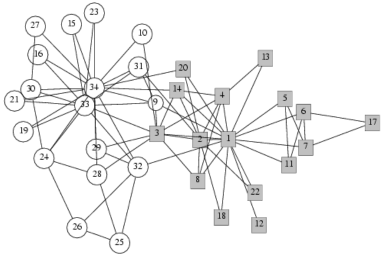
Goal: Quantify the correlation between features and structures for any graph with features and decide if TouchUp-G is needed.

Homophily

"Birds of a feather, flock together"

Majority of linked nodes are similar

- Social Networks (wrt. political beliefs, age)
- Citation Networks (wrt. research area)



Zachary's Karate Club

Feature Homophily h

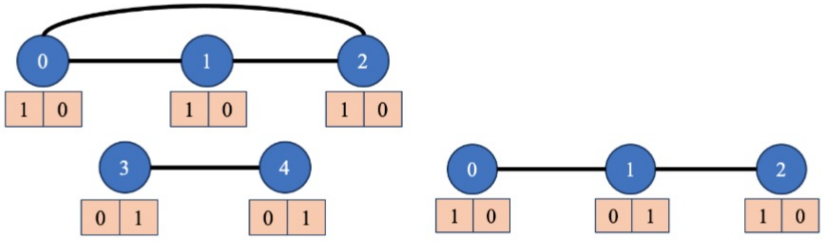
DETAILS

A vectorized extension of scalar assortativity.

$$h = \frac{\sum_{ij\in E}(x_i - \bar{x}) \cdot (x_j - \bar{x})}{\sqrt{\sum_{ij\in E}(x_i - \bar{x}) \cdot (x_i - \bar{x})} \cdot \sqrt{\sum_{ij\in E}(x_j - \bar{x}) \cdot (x_j - \bar{x})}}$$

$$\bar{x} = \frac{\sum_{ij\in E}(x_i + x_j)}{2|E|}$$

The Pearson correlation between the set of all head node features x_i and the set of all tail node features x_j. Bounded in [-1, 1].



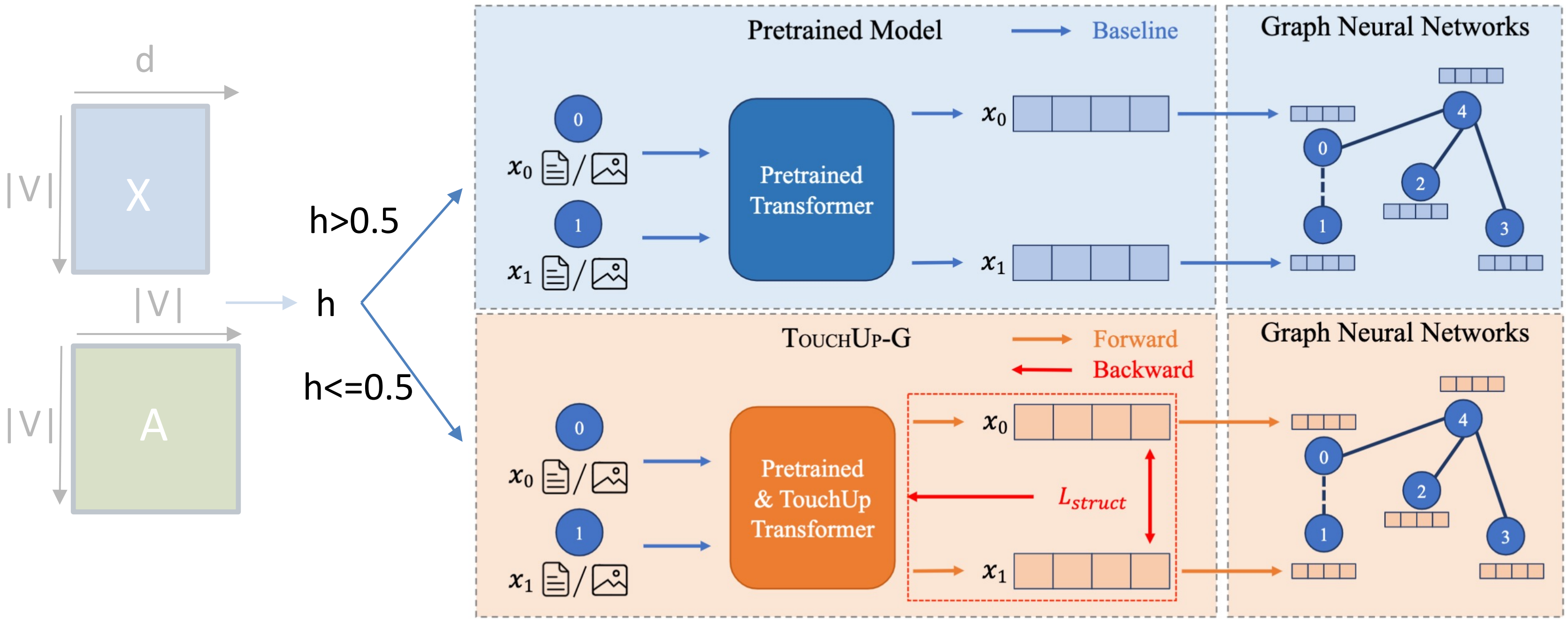
(a) **h = 1** (b) **h = -1**

Jing Zhu et. al

TouchUp-G Pipeline

①Feature Homophily h

② Graph-Centric Finetuning



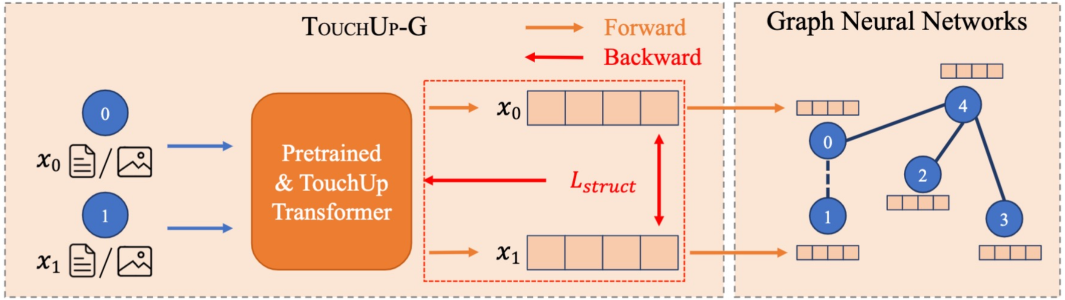
Jing Zhu et. al

Graph-centric Finetuning

DETAILS

$$L_{\text{struct}} = -\frac{1}{|E|} \sum_{\substack{(u,v) \in E \\ (u,m) \notin E}} \left(\log(x'_u \cdot x'_v) + \log(1 - x'_u \cdot x'_m) \right)$$

- For each training edge $(u, v) \in E$, randomly sample a negative node m
- $x'_u = \max(T(x_u), 0)$, $x'_v = \max(T(x_v), 0)$, $x'_m = \max(T(x_m), 0)$ are the pretrained feature embeddings obtained from a pretrained model T
- Finetune T using binary cross entropy loss



Jing Zhu et. al

Experiments

- ✧ RQ1 - Effective: How accurate is TouchUp-G?
- ✧ RQ2 - Multi-modal: Can TouchUp-G handle other modalities (like, images), in addition to text?
- ✧ RQ3 - General: Can TouchUp-G handle other downstream tasks (like node classification), besides link prediction?
- ✧ RQ4 - Principled: How well is the feature representation learnt by TouchUp-G, according to feature homophily score, compared with features obtained directly from pretrained models?

Experiment Details

Name	Nodes	Edges	Description	Node Features	Pretrained Model	Downstream Task
Ogb-Products [18]	2,449,029	61,859,140	Purchasing Network	Text	BERT [9]	LP & NC
Books [46, 47]	1,098,672	33,619,434	Recommendation Network	Text	BERT [9]	LP
Amazon-CP [36]	379,770	4,102,444	Purchasing Network	Image	ViT [11]	LP
Ogb-Arxiv [18]	169,343	1,166,243	Citation Network	Text	SciBERT [1]	NC

- Amazon-CP : We extract the copurchasing information from the metadata in Amazon-Review. Each product's high-resolution image are used as feature.
- Books: Extract node feature is each book's description and the links capture if a reader who recommends one book will recommend the other book.
- For textual features, BERT/SciBERT is used as pretrained models.
- For visual features, Vision Transformer (ViT) is used as pretrained models.

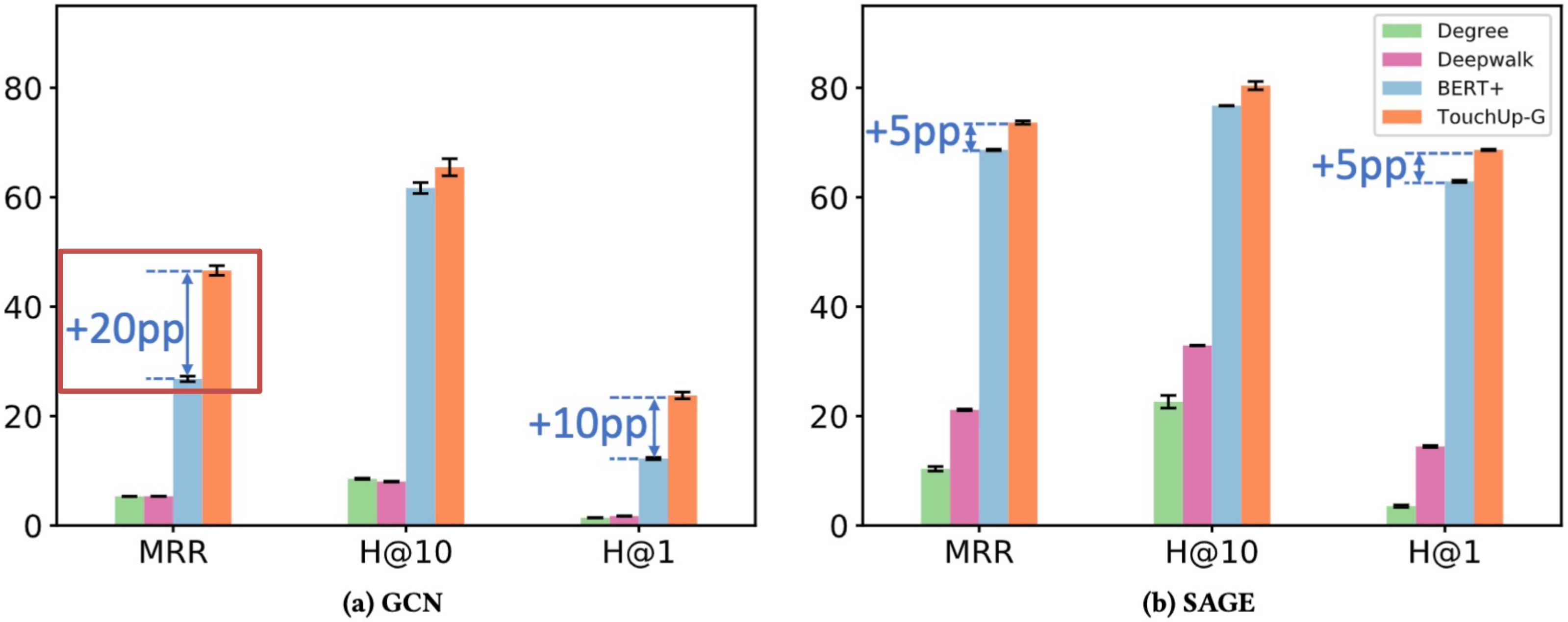
Experiment Details

Name	Nodes	Edges	Description	Node Features	Pretrained Model	Downstream Task
Ogb-Products [18]	2,449,029	61,859,140	Purchasing Network	Text	BERT [9]	LP & NC
Books [46, 47]	1,098,672	33,619,434	Recommendation Network	Text	BERT [9]	LP
Amazon-CP [36]	379,770	4,102,444	Purchasing Network	Image	ViT [11]	LP
Ogb-Arxiv [18]	169,343	1,166,243	Citation Network	Text	SciBERT [1]	NC

- Baselines:
  - Structure embeddings: Degree, Deepwalk
  - Feature embeddings: BERT+, ViT+, Ogb+, SciBERT+, DeBERTa+
- Evaluation:
  - LP – report MRR, Hits@10, and Hits@1
  - NC – report Acc
- GNN Backbones: SAGE, GCN, GATv2

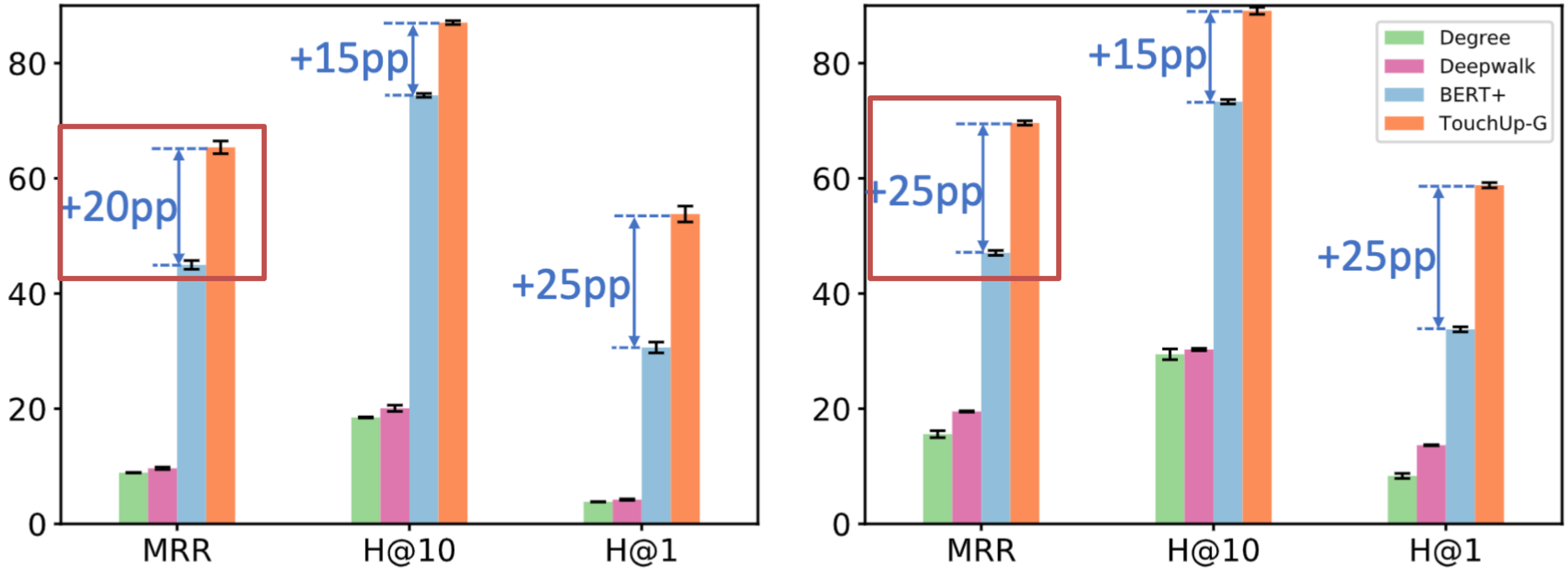
RQ1 Effective (LP)

Ogb-Products



RQ1 Effective (LP)

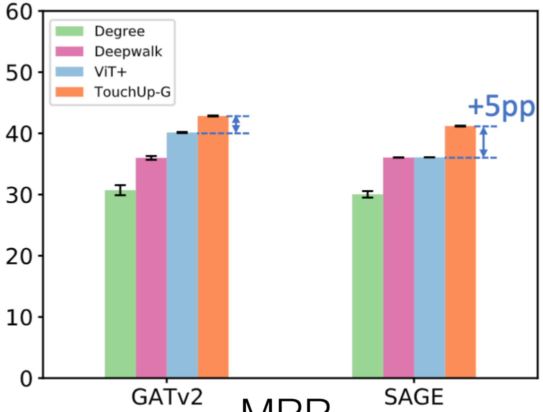
Books



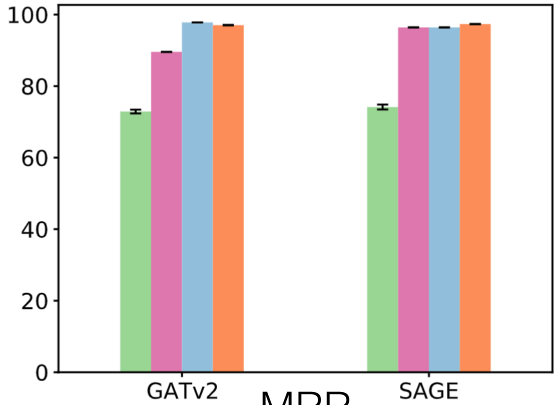
(a) GCN (b) SAGE

Jing Zhu et. al

RQ2 Multimodal(LP)

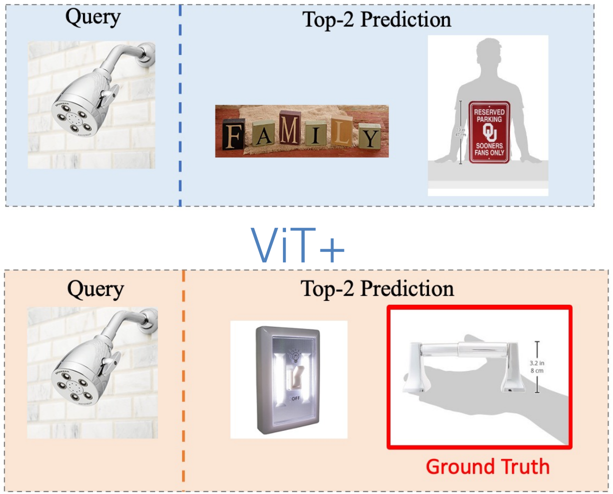


MRR



MRR

Amazon-CP



TouchUp-G Wins!

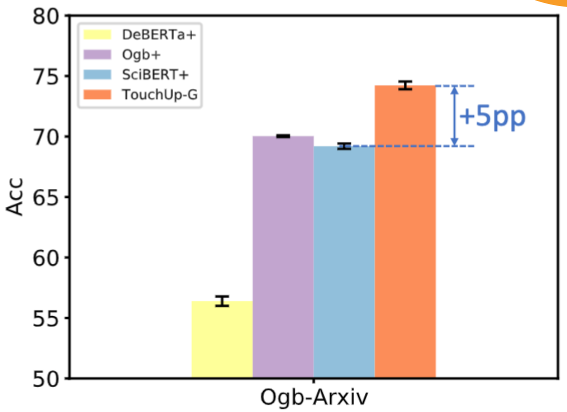
TouchUp-G works better than ViT+ both quantitatively and qualitatively.
TouchUp-G correctly predicts the ground truth in top-2 predictions while ViT+ fails.

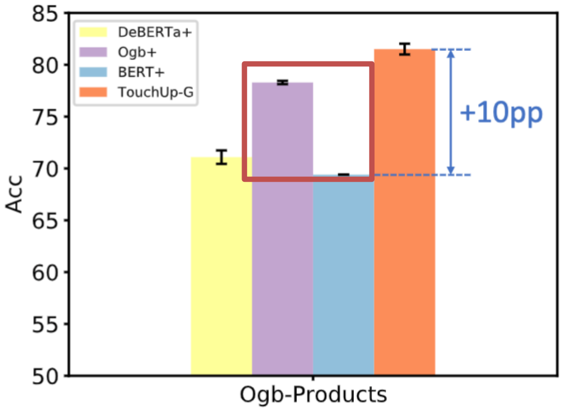
RQ3 General (NC)

TouchUp-G works better than BERT+, SciBERT+.

For Ogb-Products, ogb+ mainly uses bag of words representation and works much better than BERT+, DeBERTa+.

This indicates contextualized embeddings from PMs hurts downstream graph task performance, if the contextualization is irrelevant wrt. the downstream task.





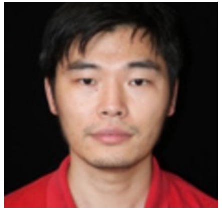
Jing Zhu et. al

RQ4 Principled: Feature Homophily Score

Dataset	BERT+ [9]	SciBERT+ [1]	ViT+ [11]	TouchUp-G
Ogb-Products	0.223	-	-	**0.762** (3.4x)
Books	0.137	-	-	**0.579** (4.2x)
Amazon-CP	-	-	0.173	**0.622** (3.6x)
Ogb-Arxiv	-	0.194	-	**0.408** (2.1x)

All datasets exhibits a more than 2× increase in feature homophily score.

Thanks! Questions?


We thank the following for the helpful feedback on this project.