Hierarchical Graph Representation Learning with Differentiable Pooling

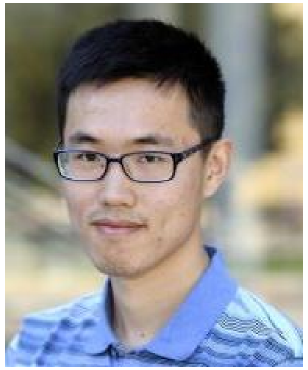
Authors: Rex Ying, Jiaxuan You, Christopher Morris, Xiang Ren, William L. Hamilton, Jure Leskovec

Speaker: Ziyuan Ye

Thursday, August 18, 2022

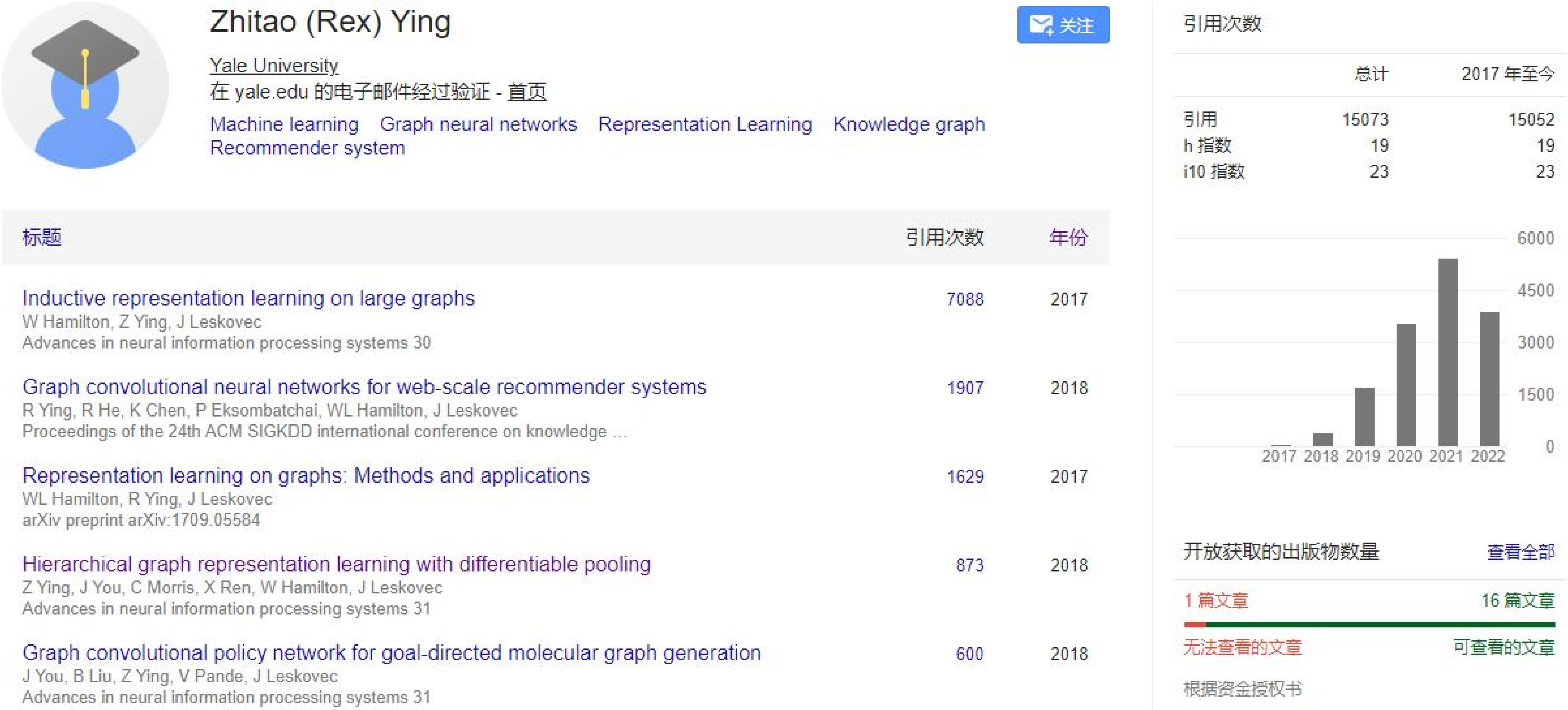
Ying, Z., You, J., Morris, C., Ren, X., Hamilton, W., & Leskovec, J. (2018). Hierarchical graph representation learning with differentiable pooling. Advances in neural information processing systems (NeurIPS), 31.

About the authors



Representative Work:

- ✓ GraphSAGE
- ✓ GNNExplainer
- ✓ DIFFPOOL



Zhitao (Rex) Ying

关注

Yale University

在 yale.edu 的电子邮件经过验证 - 首页

Machine learning   Graph neural networks   Representation Learning   Knowledge graph   Recommender system

标题	引用次数	年份
Inductive representation learning on large graphs W Hamilton, Z Ying, J Leskovec Advances in neural information processing systems 30	7088	2017
Graph convolutional neural networks for web-scale recommender systems R Ying, R He, K Chen, P Eksombatchai, WL Hamilton, J Leskovec Proceedings of the 24th ACM SIGKDD international conference on knowledge ...	1907	2018
Representation learning on graphs: Methods and applications WL Hamilton, R Ying, J Leskovec arXiv preprint arXiv:1709.05584	1629	2017
Hierarchical graph representation learning with differentiable pooling Z Ying, J You, C Morris, X Ren, W Hamilton, J Leskovec Advances in neural information processing systems 31	873	2018
Graph convolutional policy network for goal-directed molecular graph generation J You, B Liu, Z Ying, V Pande, J Leskovec Advances in neural information processing systems 31	600	2018

引用次数

	总计	2017 年至今
引用	15073	15052
h 指数	19	19
i10 指数	23	23

开放获取的出版物数量 查看全部

1 篇文章 16 篇文章

无法查看的文章 可查看的文章

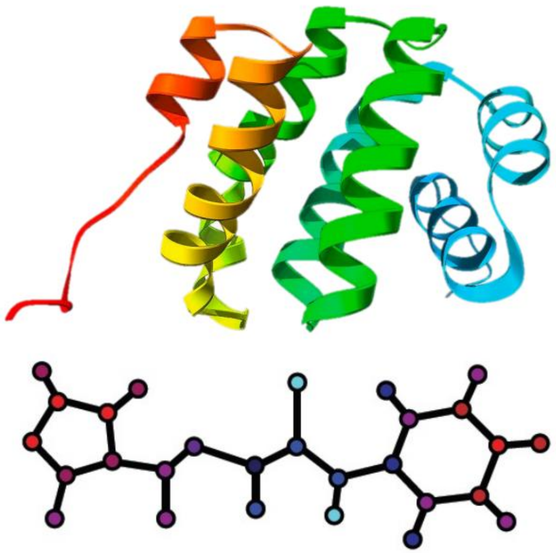
根据资金授权书

Content

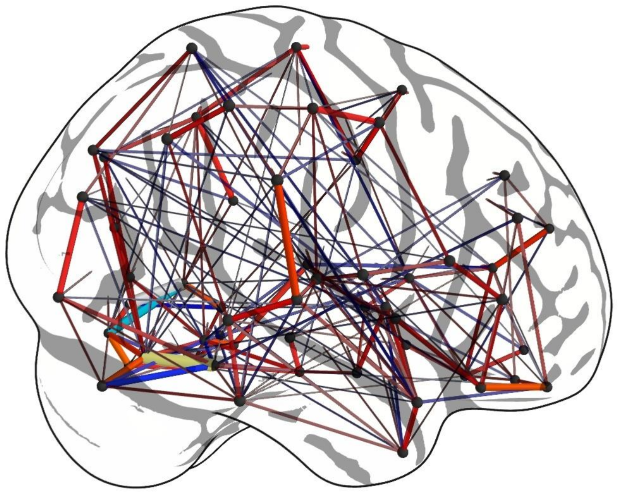
1. Background
2. Research content
3. Experimental results
4. Take-home message

Content

Graphs = System of Relations and Interactions



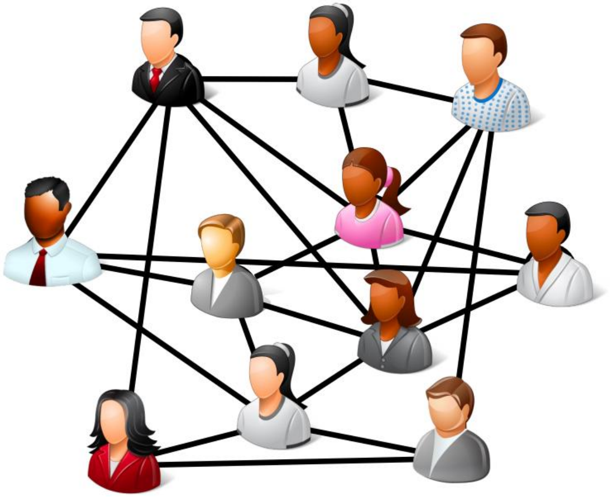
Molecules



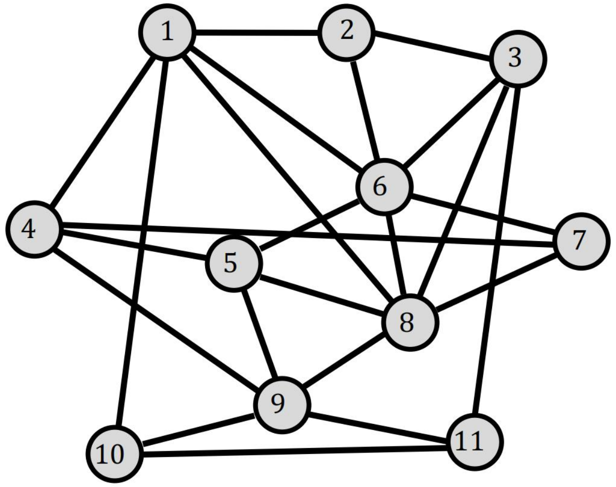
Brain connectivity

Social Network

Graph Classification via Graph Neural Networks (GNNs)

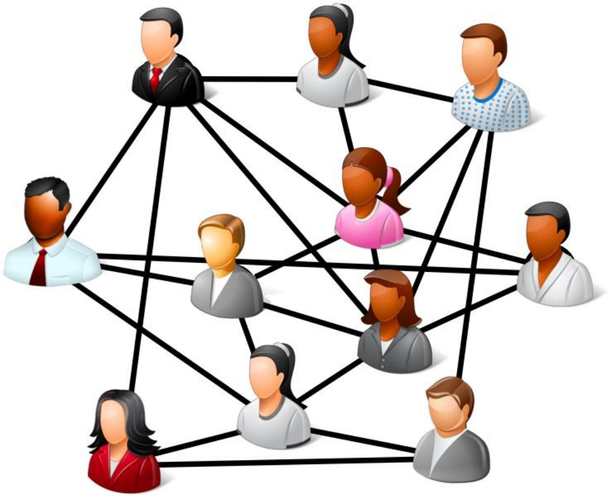


Social network

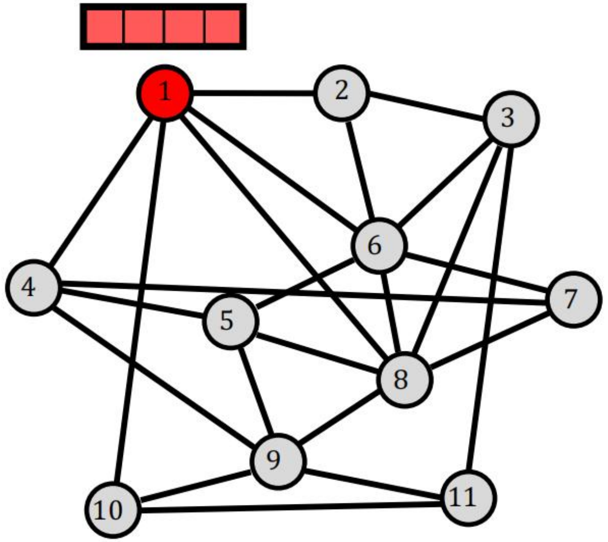


Graph

Graph Classification via GNNs

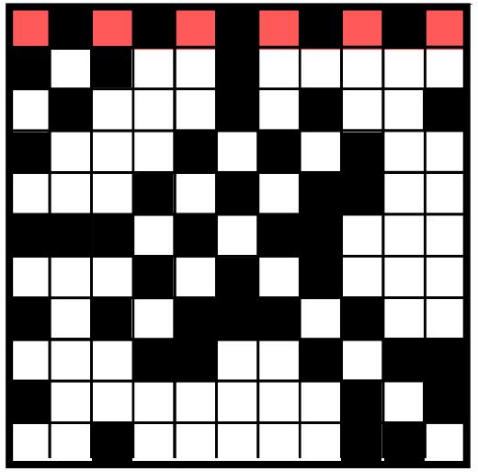


Social network



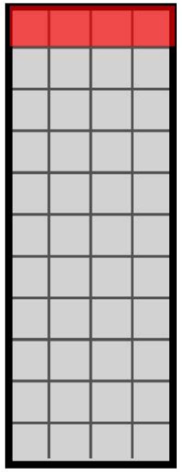
Graph

Adjacency matrix $n \times n$



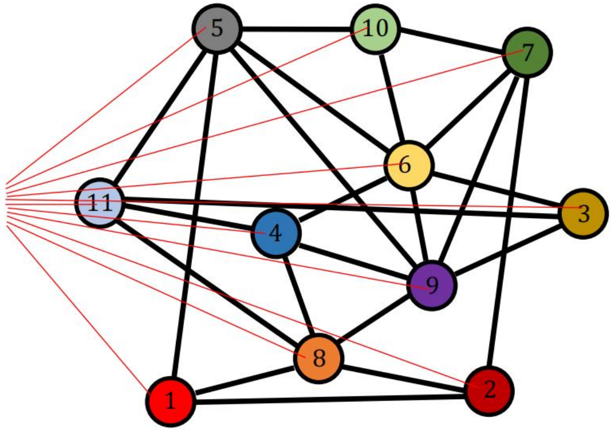
A

Feature matrix $n \times d$

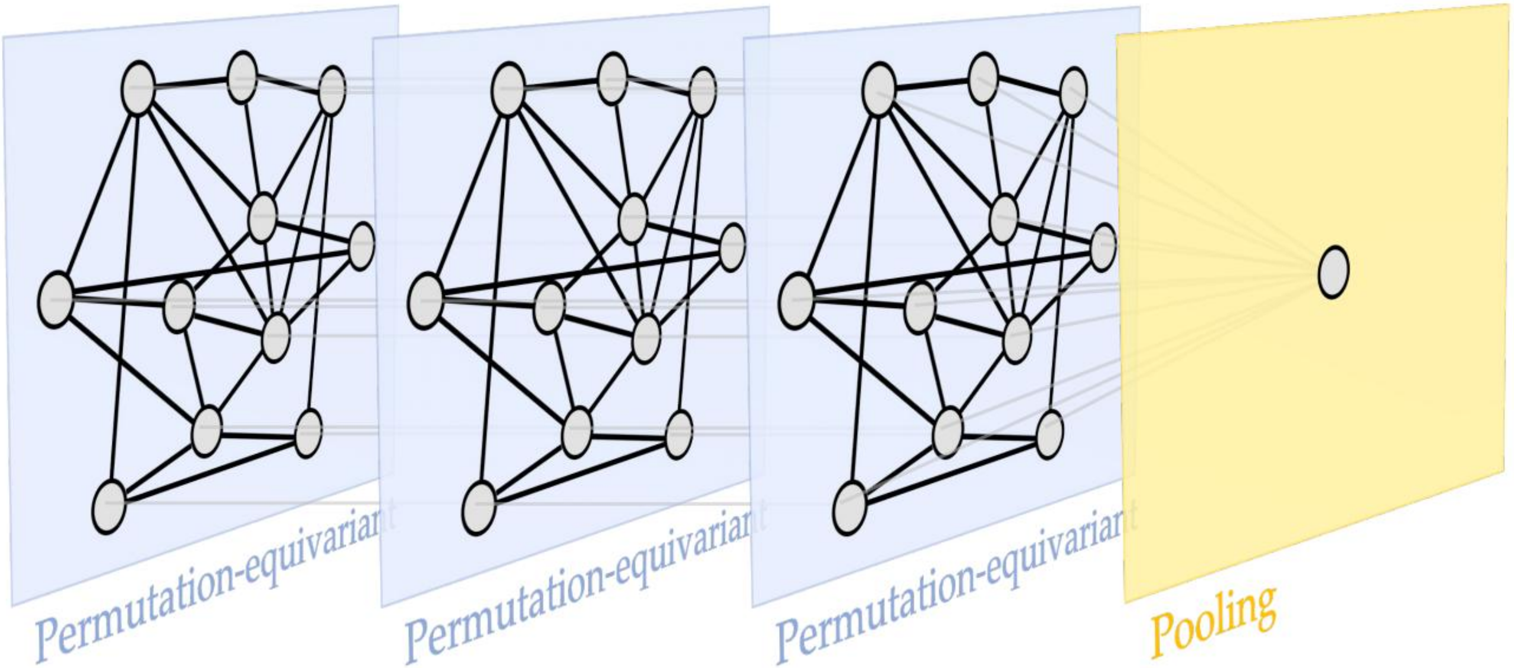


X

Graph function $f(\boldsymbol{X}, \boldsymbol{A})$

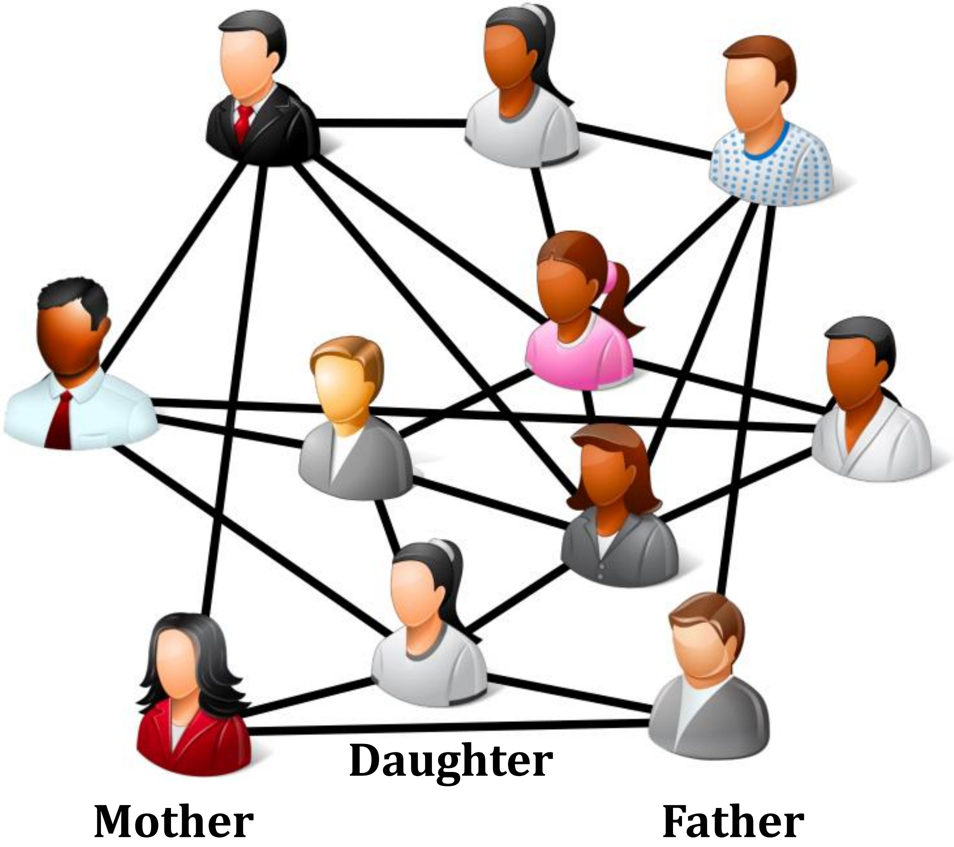


Graph pooling in GNNs

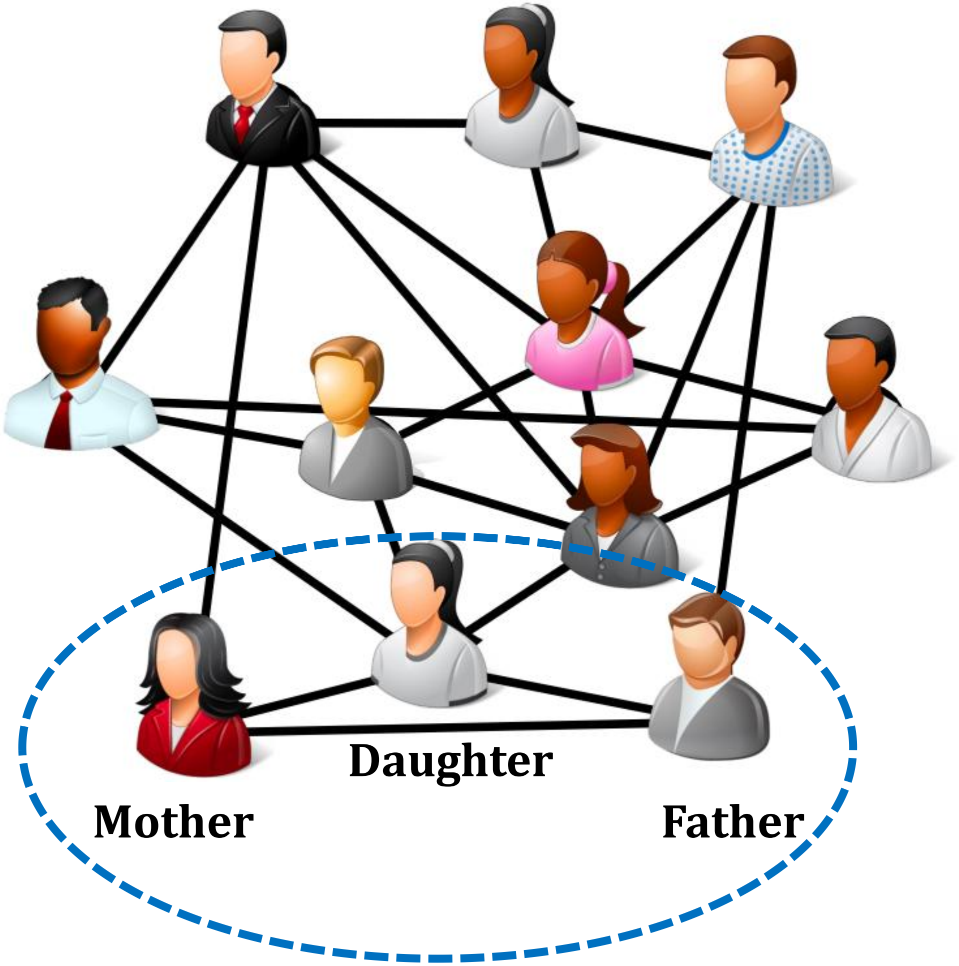


Ignoring any hierarchical structure that might be present in the graph

Graph pooling in GNNs



Graph pooling in GNNs



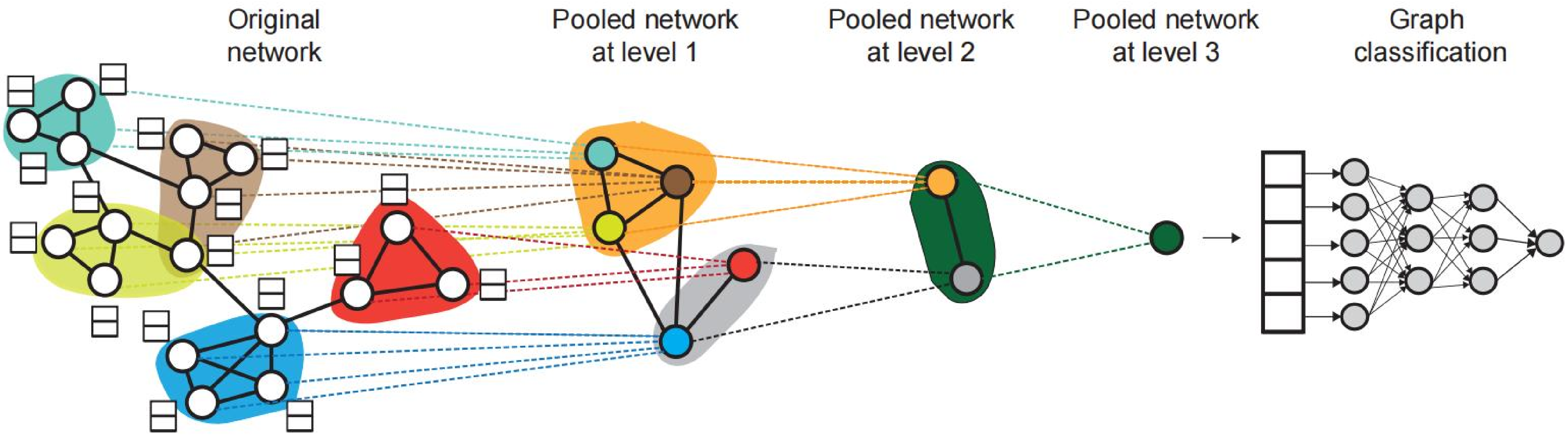
Content

1. Background
2. Research content
3. Experimental results
4. Take-home message

Notation

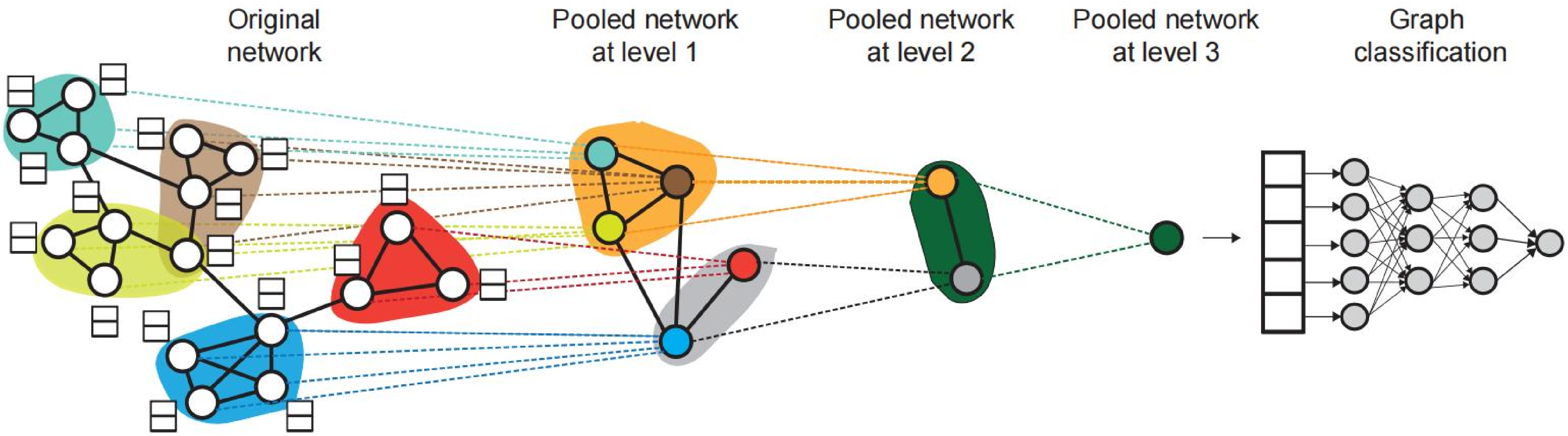
Adjacency matrix	$A \in \{0,1\}^{n \times n}$
Feature matrix	$F \in \mathbb{R}^{n \times d}$
Graph	G as (A, F)
Node embeddings computed after k steps of GNN	$H^{(k)} \in \mathbb{R}^{n \times d}$
Node embeddings of coarsened graph	$Z^{(k)} \in \mathbb{R}^{m \times d}$, $m < n$
Dataset	$\mathcal{D} = \{(G_1, y_1), (G_2, y_2), ...\}$, $y_i \in \mathcal{Y}$
Learnable cluster assignment matrix	$S^{(l)} \in \mathbb{R}^{n_l \times n_{l+1}}$
Goal of graph classification	$f : \mathcal{G} \to \mathcal{Y}$

DIFFPOOL



Graph coarsening and classification via DIFFPOOL

DIFFPOOL

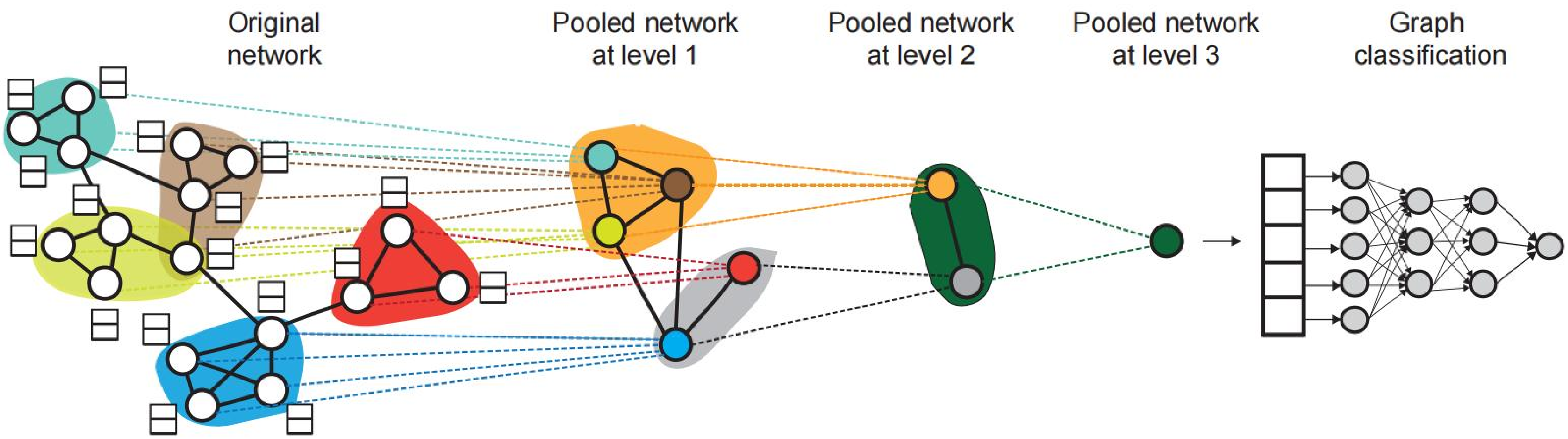


Graph coarsening and classification via DIFFPOOL

Learning cluster assignment matrix

$$S^{(l)} = \text{softmax}\left(\text{GNN}_{l,\text{pool}}(A^{(l)}, X^{(l)})\right)$$

DIFFPOOL



Graph coarsening and classification via DIFFPOOL

Learning cluster assignment matrix

$$S^{(l)} = \text{softmax}\left(\text{GNN}_{l,\text{pool}}(A^{(l)}, X^{(l)})\right)$$

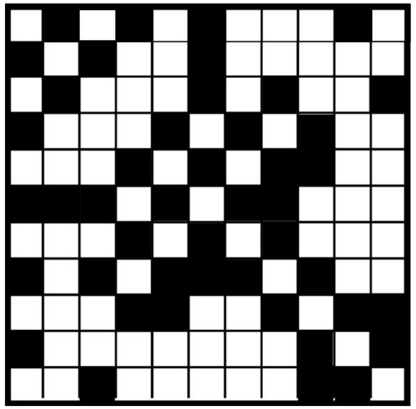
Updating embeddings and adjacency matrix

$$Z^{(l)} = \text{GNN}_{l,\text{embed}}(A^{(l)}, X^{(l)})$$

$$X^{(l+1)} = {S^{(l)}}^T Z^{(l)} \in \mathbb{R}^{n_{l+1} \times d},$$

$$A^{(l+1)} = {S^{(l)}}^T A^{(l)} S^{(l)} \in \mathbb{R}^{n_{l+1} \times n_{l+1}}$$

DIFFPOOL: Regularization terms

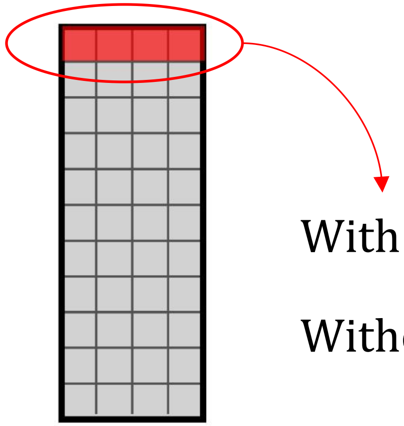


Adjacency matrix $n \times n$

Entropy Regularization

$$L_{\mathrm{E}} = \frac{1}{n} \sum_{i=1}^{n} H(S_i)$$

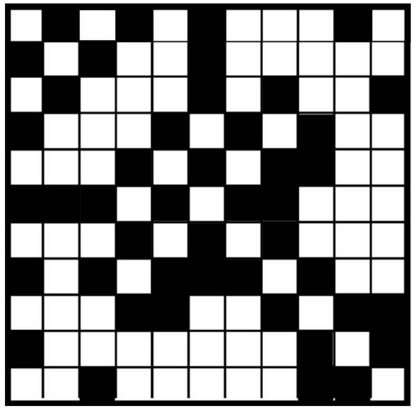
Assignment matrix $n \times m$



With: [0.95, 0.01, 0.02, 0.02]

Without: [0.33, 0.34, 0.20, 0.13]

DIFFPOOL: Regularization terms

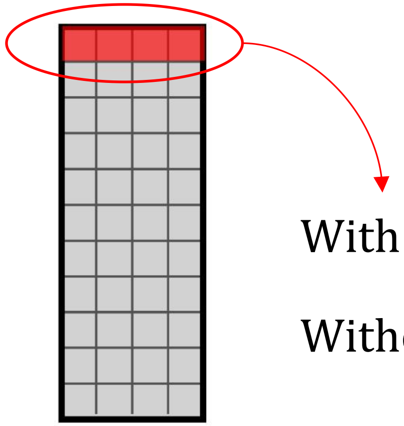


Adjacency matrix $n \times n$

Entropy Regularization

$$L_{\mathrm{E}} = \frac{1}{n} \sum_{i=1}^{n} H(S_i)$$

Assignment matrix $n \times m$

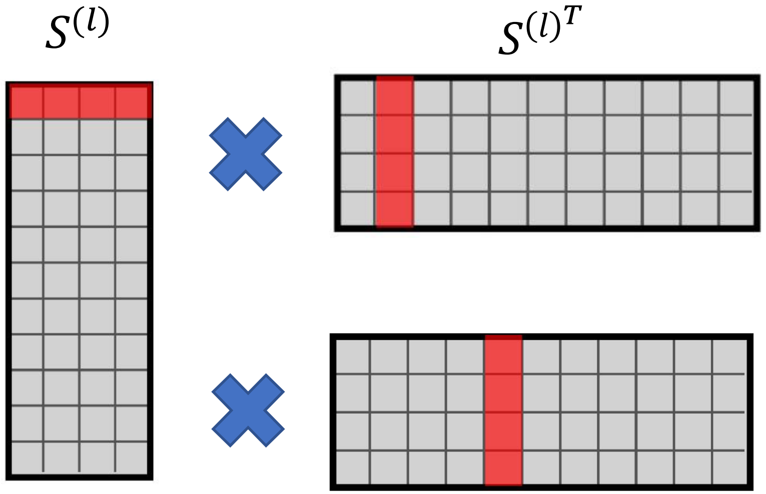


With: [0.95, 0.01, 0.02, 0.02]

Without: [0.33, 0.34, 0.20, 0.13]

Auxiliary Link Prediction Objective

$$L_{\mathrm{LP}} = ||A^{(l)}, S^{(l)} S^{(l)^T} ||_F$$



The higher the better

The lower the better

Content

1. Background
2. Research content
3. Experimental results
4. Take-home message

Experimental results

	Method	Data Set					
		ENZYMES	D&D	REDDIT-MULTI-12K	COLLAB	PROTEINS	Gain
Kernel	GRAPHLET	41.03	74.85	21.73	64.66	72.91	
	SHORTEST-PATH	42.32	78.86	36.93	59.10	76.43	
	1-WL	53.43	74.02	39.03	78.61	73.76	
	WL-OA	60.13	79.04	44.38	80.74	75.26	
GNN	PATCHYSAN	–	76.27	41.32	72.60	75.00	4.17
	GRAPHSAGE	54.25	75.42	42.24	68.25	70.48	–
	ECC	53.50	74.10	41.73	67.79	72.65	0.11
	SET2SET	60.15	78.12	43.49	71.75	74.29	3.32
	SORTPOOL	57.12	79.37	41.82	73.76	75.54	3.39
	DIFFPOOL-DET	58.33	75.47	46.18	**82.13**	75.62	5.42
	DIFFPOOL-NOLP	61.95	79.98	46.65	75.58	76.22	5.95
	DIFFPOOL	**62.53**	**80.64**	**47.08**	75.48	**76.25**	**6.27**

Content

Take-home Message

- **Motivation:**
  - ➢ Original GNN architectures are unable to infer and aggregate the information in a hierarchical way.
- **Challenge:**
  - ➢ Graphs contain no natural notion of spatial locality, i.e., one cannot simply pool together all nodes in a "m × m patch" on a graph.
- **Main contributions:**
  - ➢ Propose DIFFPOOL that can generate hierarchical representations of graphs;
  - ➢ DIFFPOOL can be combined with various graph neural network architectures in an end-to-end fashion;

Thanks for your attention!