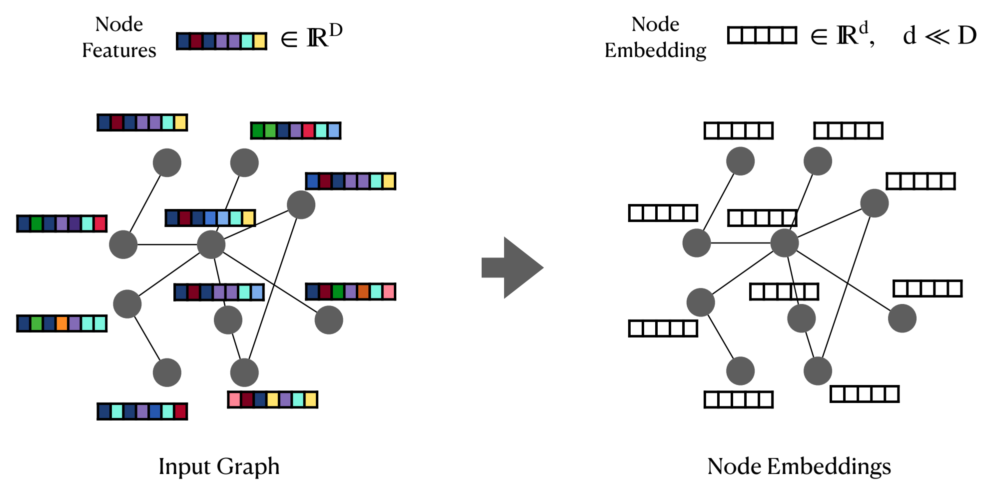
Università degli Studi di Padova
Dipartimento di Ingegneria dell'Informazione

A Meta-Learning Approach for Graph Representation Learning in Multi-Task Settings

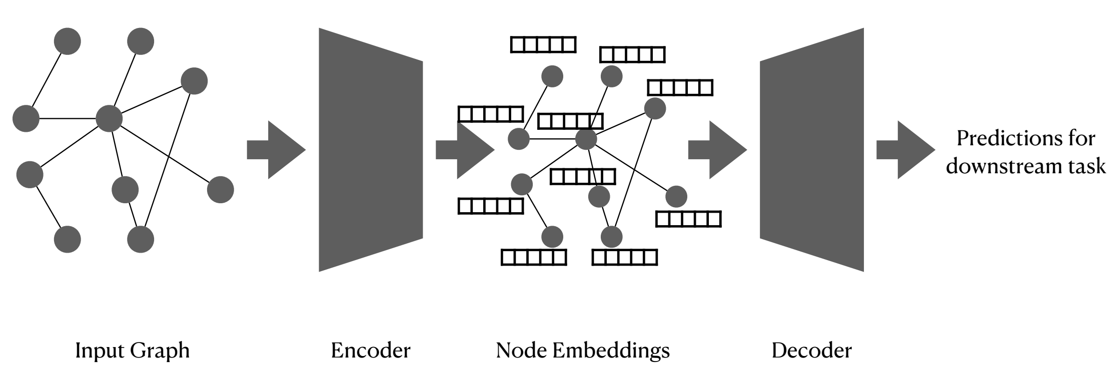
Davide Buffelli, Fabio Vandin
{davide.buffelli, fabio.vandin}@unipd.it

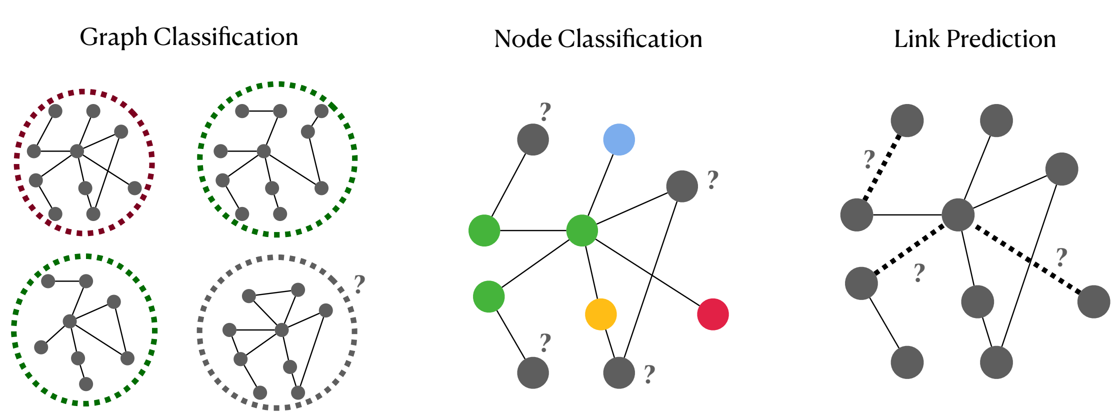
Motivation

Graph Representation Learning

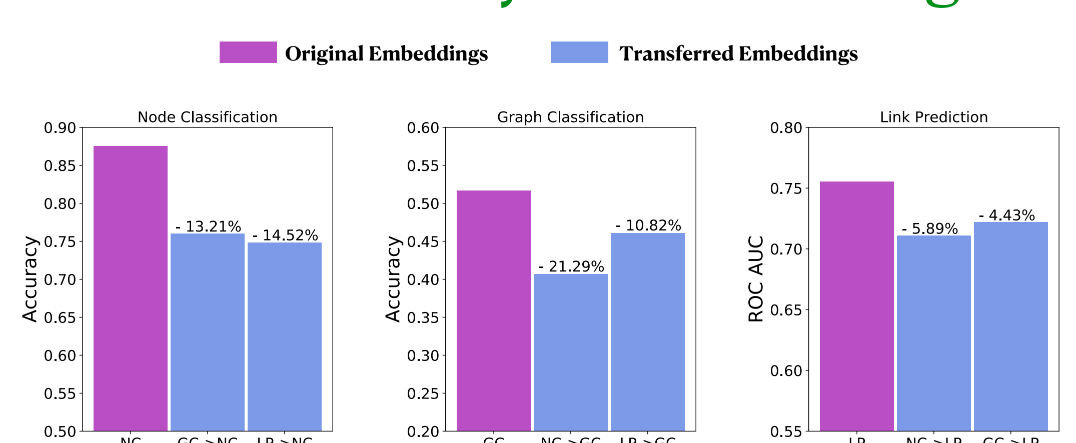


Framework & Applications





Problem: Transferability of Embeddings



Can we generate node embeddings that generalize across tasks?

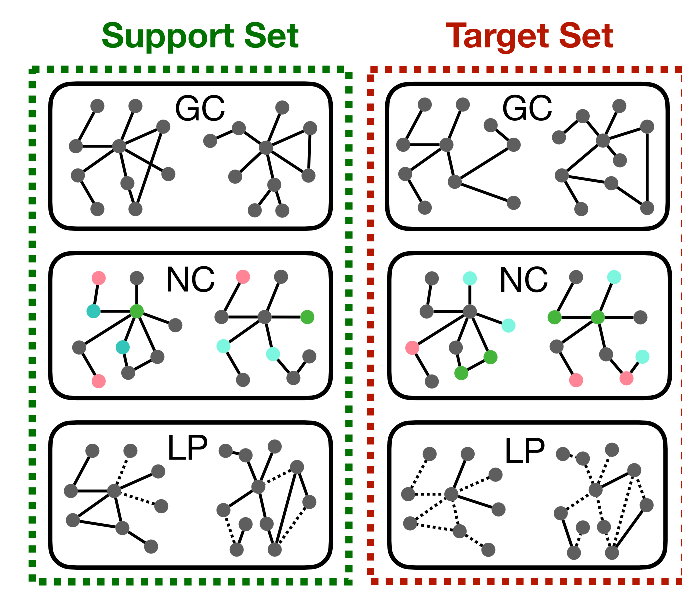
Our Method

SAME

Single-Task Adaptation for Multi-Task Embeddings

Multi-Task Episodes

- 1 task-specific support set per task
- 1 task-specific target set per task
- Each task can have a different loss



Support set ⟷ **Training set**
Target set ⟷ **Validation set**

The outer loop's objective becomes to maximize the performance on a validation set, given a training set, hence pushing towards generalization.

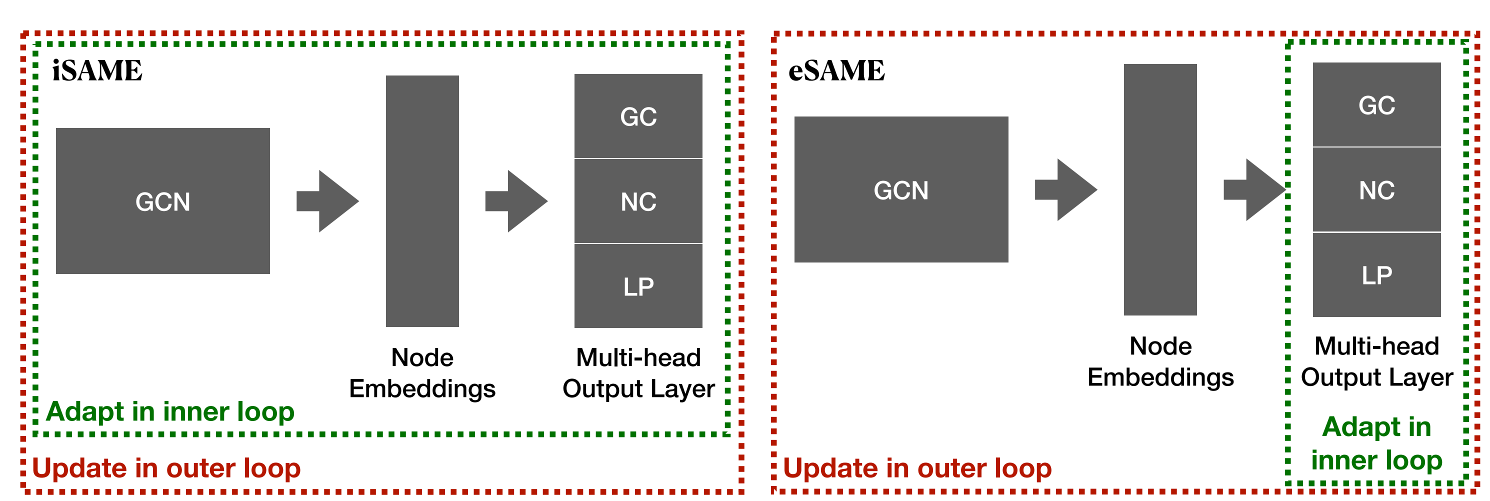
Meta-Learning Procedure

- Separate adaptation for each task
- Unique outer loop update
- Don't learn to perform all tasks **concurrently**, learn to quickly adapt to each of them **singularly**!

Algorithm 1: Proposed Meta-Learning Procedure

Input : Model f_θ; Episodes $\mathcal{E} = \{\mathcal{E}_1, .., \mathcal{E}_n\}$
init(θ)
for $\mathcal{E}_i$ *in* $\mathcal{E}$ **do**
  o_loss $\leftarrow 0$
  for t *in (GC, NC, LP)* **do**
    $\theta'^{(t)} \leftarrow \theta$
    $\theta'^{(t)} \leftarrow \mathtt{ADAPT}(f_\theta, \mathcal{S}^{(t)}_{\mathcal{E}_i}, \mathcal{L}^{(t)}_{\mathcal{E}_i})$
    o_loss $\leftarrow$ o_loss $+ \mathtt{TEST}(f_{\theta'^{(t)}}, \mathcal{T}^{(t)}_{\mathcal{E}_i}, \mathcal{L}^{(t)}_{\mathcal{E}_i})$
  end
  $\theta \leftarrow \mathtt{UPDATE}(\theta, \texttt{o\_loss}, \theta'^{(GC)}, \theta'^{(NC)}, \theta'^{(LP)})$
end

Two variants:
- *Implicit*- SAME (iSAME)
- *Explicit*-SAME (eSAME)



Experimental Results

Setting

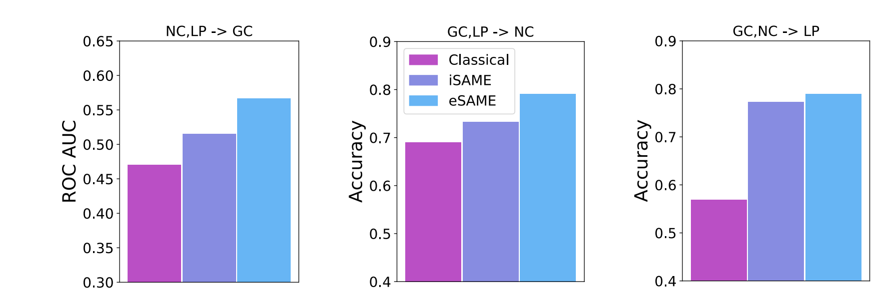
- Same architecture for all training strategies
- Datasets from TUDortmund with node labels, graph labels, and node attributes
- Linear model on top of the embeddings from SAME
- Tests in single-task and multi-task scenarios

Q1-Q2: Do iSAME and eSAME lead to high quality node embeddings in the single-task & multi-task settings?

Task	Model	ENZYMES	PROTEINS	DHFR	COX2
NC	Cl	87.5 ± 1.9	72.3 ± 4.4	97.3 ± 0.2	96.4 ± 0.3
	iSAME	87.3 ± 0.8	81.8 ± 1.6	96.6 ± 0.3	96.1 ± 0.4
	eSAME	87.8 ± 0.7	82.4 ± 1.6	96.8 ± 0.2	96.5 ± 0.6
GC	Cl	51.6 ± 4.2	73.3 ± 3.6	71.5 ± 2.3	76.7 ± 4.7
	iSAME	50.8 ± 2.9	73.5 ± 1.2	73.2 ± 3.2	76.3 ± 4.6
	eSAME	52.1 ± 5.0	72.6 ± 1.6	71.6 ± 2.4	75.6 ± 4.1
LP	Cl	75.5 ± 3.0	85.6 ± 0.8	98.8 ± 0.7	98.3 ± 0.8
	iSAME	81.7 ± 1.7	84.0 ± 1.1	99.2 ± 0.4	99.1 ± 0.5
	eSAME	80.1 ± 3.4	84.1 ± 0.9	99.2 ± 0.3	99.2 ± 0.7

Task GC	Task NC	Task LP	ENZYMES GC	ENZYMES NC	ENZYMES LP	PROTEINS GC	PROTEINS NC	PROTEINS LP	DHFR GC	DHFR NC	DHFR LP	COX2 GC	COX2 NC	COX2 LP
Classical End-to-End Training														
✓			51.6			73.3			71.5			76.7		
	✓			87.5			72.3			97.3			96.4	
		✓			75.5			85.6			98.8			98.3
Fine-Tuning														
✓	✓		48.3	85.3		73.6	72.0		66.4	92.4		80.0	92.3	
✓		✓	49.3		71.6	69.6		80.7	65.3		58.9	80.2		50.9
	✓	✓		87.7	73.9		80.4	81.5		80.7	56.6		87.4	52.3
iSAME (ours)														
✓	✓		50.1	86.1		73.1	76.6		71.6	94.8		75.2	95.4	
✓		✓	50.7		83.1	73.4		85.2	71.6		99.2	77.5		98.9
	✓	✓		86.3	83.4		79.4	87.7		96.5	99.3		95.5	99.0
✓	✓	✓	50.0	86.5	82.3	71.4	76.6	87.3	71.2	95.5	99.5	75.4	95.2	99.2
eSAME (ours)														
✓	✓		51.7	86.1		71.5	79.2		70.1	95.7		75.6	95.5	
✓		✓	51.9		80.1	71.7		85.4	70.1		99.1	77.5		98.8
	✓	✓		86.7	82.2		80.7	86.3		96.6	99.4		95.6	99.1
✓	✓	✓	51.5	86.3	81.1	71.3	79.6	86.8	70.2	95.3	99.5	77.7	95.7	98.8

Q3: Do iSAME and eSAME extract information that is not captured by classically trained multi-task models?



Q4: Can the node embeddings learned by iSAME and eSAME be used to perform multiple tasks with comparable or better performance than classical multi-task models?

GC	NC	LP	Model	ENZYMES	PROTEINS	DHFR	COX2
✓	✓		Cl	-0.1 ± 0.5	4.0 ± 1.0	-0.3 ± 0.2	0.5 ± 0.1
			FT	-4.5 ± 1.2	0.1 ± 0.5	-7.4 ± 1.4	0.1 ± 0.4
			iSAME	-2.3 ± 0.9	2.7 ± 1.5	-1.2 ± 0.4	-1.6 ± 0.2
			eSAME	-0.8 ± 0.8	3.2 ± 1.4	-1.8 ± 0.3	-1.2 ± 0.3
✓		✓	Cl	-25.3 ± 3.2	-5.3 ± 1.2	-28.3 ± 4.3	-21.4 ± 3.4
			FT	-5.1 ± 1.9	-5.4 ± 1.5	-24.5 ± 3.7	-22.6 ± 3.8
			iSAME	4.1 ± 0.5	-0.2 ± 0.9	0.2 ± 3.2	0.2 ± 0.5
			eSAME	3.2 ± 0.4	-1.2 ± 1.1	-0.7 ± 3.4	-0.8 ± 0.7
	✓	✓	Cl	7.2 ± 2.7	6.8 ± 0.9	-29.1 ± 7.7	-28.2 ± 4.5
			FT	-1.0 ± 0.3	3.1 ± 1.2	-28.9 ± 6.4	-28.3 ± 4.2
			iSAME	4.4 ± 1.1	6.1 ± 1.0	-0.1 ± 6.2	-0.6 ± 2.5
			eSAME	3.9 ± 1.3	6.1 ± 1.1	0.1 ± 6.4	-0.6 ± 2.6
✓	✓	✓	Cl	1.6 ± 1.3	2.9 ± 0.3	-18.9 ± 2.3	-16.9 ± 3.1
			iSAME	1.5 ± 1.0	2.2 ± 0.2	-0.5 ± 1.4	-0.9 ± 1.3
			eSAME	1.8 ± 0.9	2.8 ± 0.2	-1.0 ± 1.7	-0.4 ± 1.2