Multi-Class and Multi-Task Strategies for Neural Directed Link Prediction

Claudio Moroni

Executive Summary

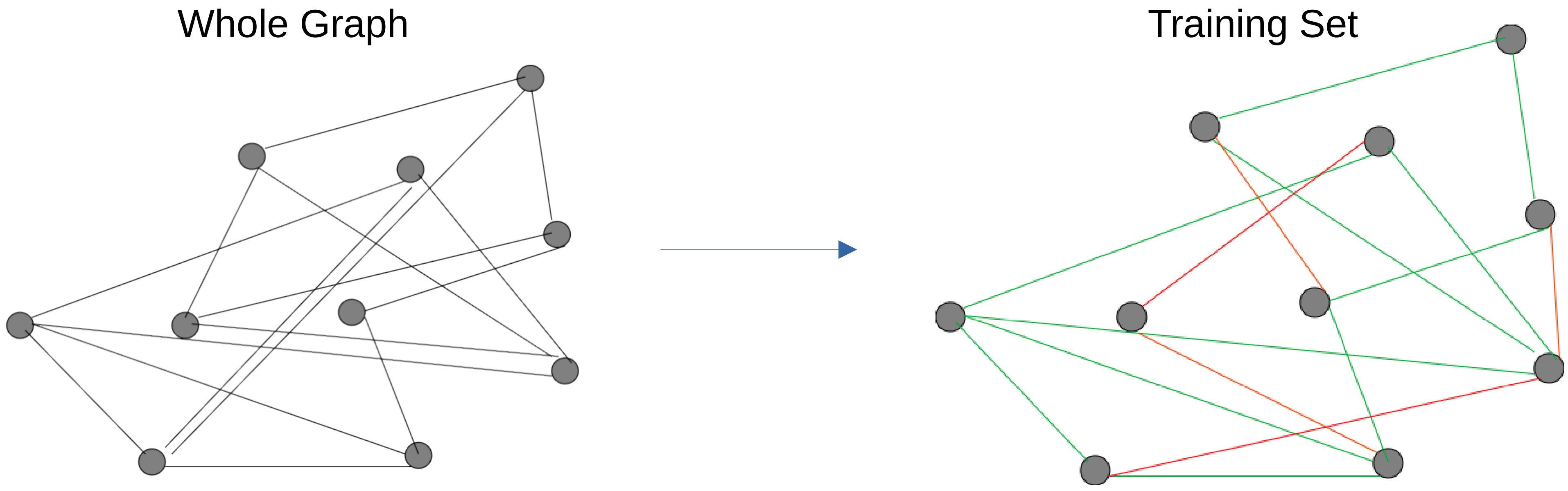
1) Naively porting ULP testing techniques to DLP **is not sound → need for a “three-way” testing setup.**
2) Naively porting ULP training techniques to DLP **leads to subotpimal performances → gap**
3) We fill this gap by **developing three DLP training strategies** perform better in the three-way testing setup

Outline

1) Undirected Link Prediction
2) Naive Directed Link Prediction
3) New Training Strategies & Results

Undirected Link Prediction

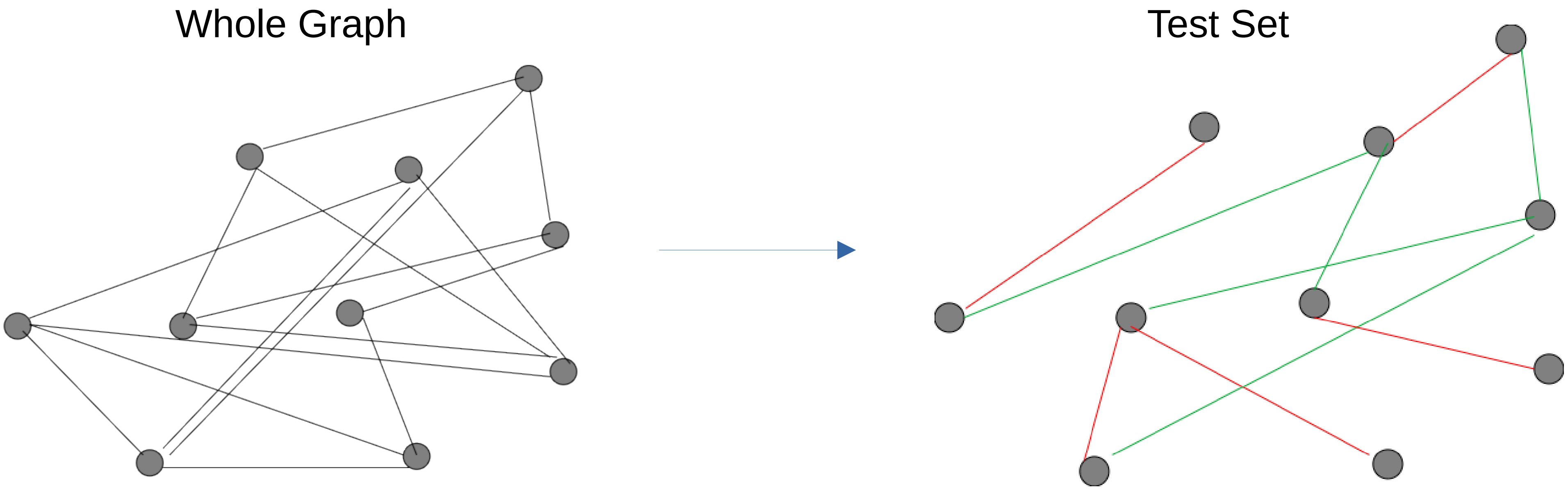
Training



$$\mathcal{L}(\Theta) = \sum_{e \in E_{s+n}} w_p y_e \ln(\hat{p}_\Theta(e)) + (1 - y_e) \ln(1 - \hat{p}_\Theta(e))$$

Undirected Link Prediction

Testing

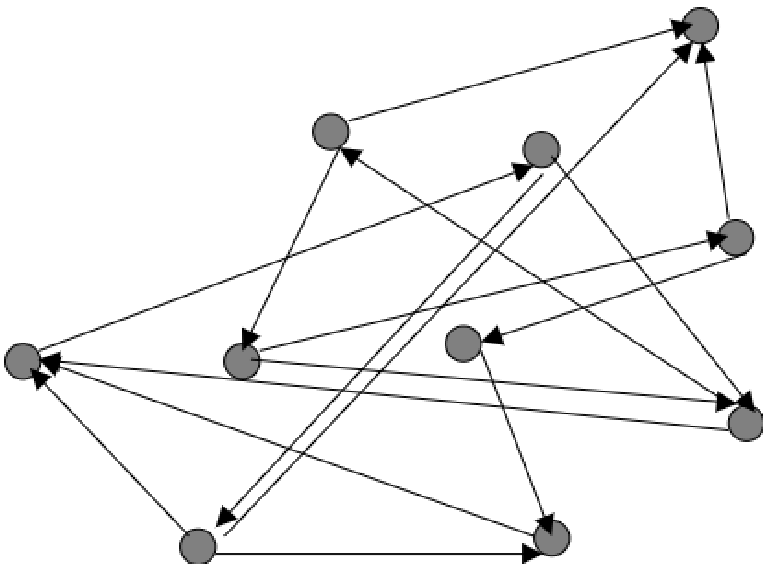


AUC, AP, Hits@K, MRR

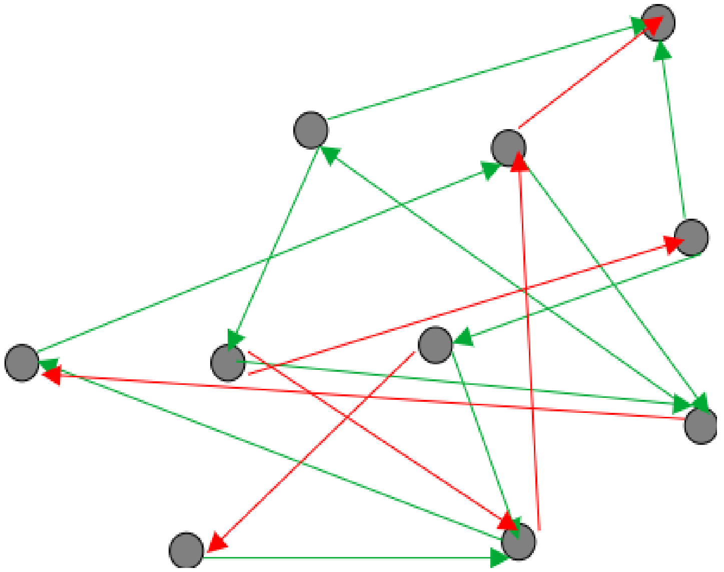
Directed Link Prediction

Naive Training

Whole Graph



Training Set

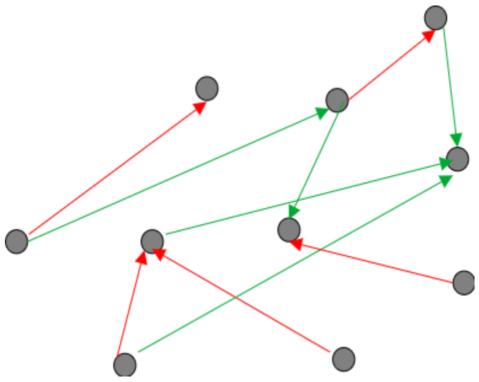


$$\mathcal{L}(\Theta) = \sum_{e \in E_{s+n}} w_p y_e \ln(\hat{p}_\Theta(e)) + (1 - y_e) \ln(1 - \hat{p}_\Theta(e))$$

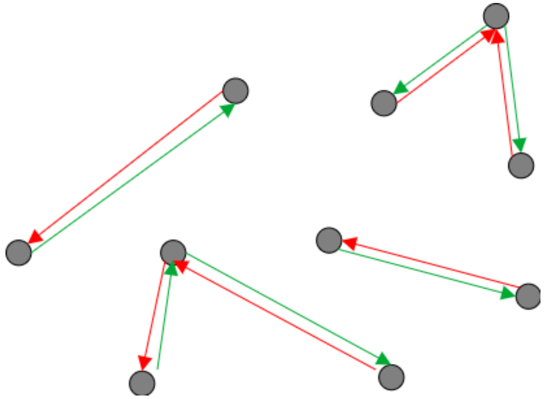
Directed Link Prediction

Testing

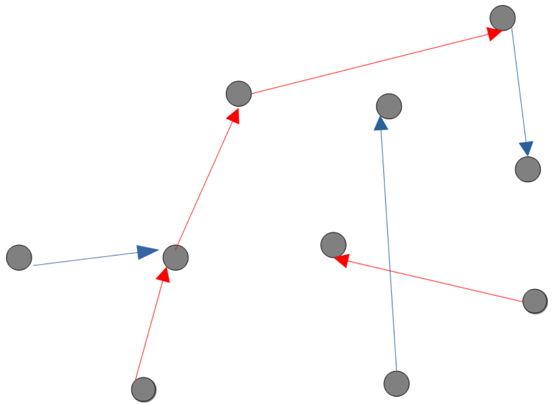
General Test Set



Directional Test Set



Bidirectional Test Set



AUC, AP, Hits@K, MRR

Directed Link Prediction

Results with Naive Training

model	decoder	GENERAL ROC-AUC	GENERAL AUPRC	DIRECTIONAL ROC-AUC	DIRECTIONAL AUPRC	BIDIRECTIONAL ROC-AUC	BIDIRECTIONAL AUPRC
GAE	$\vec{z}_i \cdot \vec{z}_j$	84.6	88.6	50.0	50.0	62.4	64.0
GR-GAE	$\sigma(\vec{z}_v[0] - \lambda \ln(\|\|\vec{z}_u[1:] - \vec{z}_v[1:]\|\|_2^2))$	89.2	92.4	63.4	61.5	69.1	66.5
ST-GAE	$\sigma(\vec{z}_v[:\frac{L}{2}] \cdot \vec{z}_u[\frac{L}{2}:])$	87.8	90.1	60.8	64.5	74.6	74.1
DiGAE	$\sigma(\vec{z}_v^S \cdot \vec{z}_v^T)$	80.4	85.3	57.5	63.0	70.4	68.6
MLP-GAE	$\mathrm{MLP}(\overrightarrow{z}_v \|\| \vec{z}_u)$	77.1	78.2	90.7	90.7	69.9	69.7
MAGNET	MLP-like	75.2	77.8	90.4	89.8	71.9	70.4
dMPLP	MLP-like	86.1	88.0	75.7	76.8	81.1	82.2

Directed Link Prediction

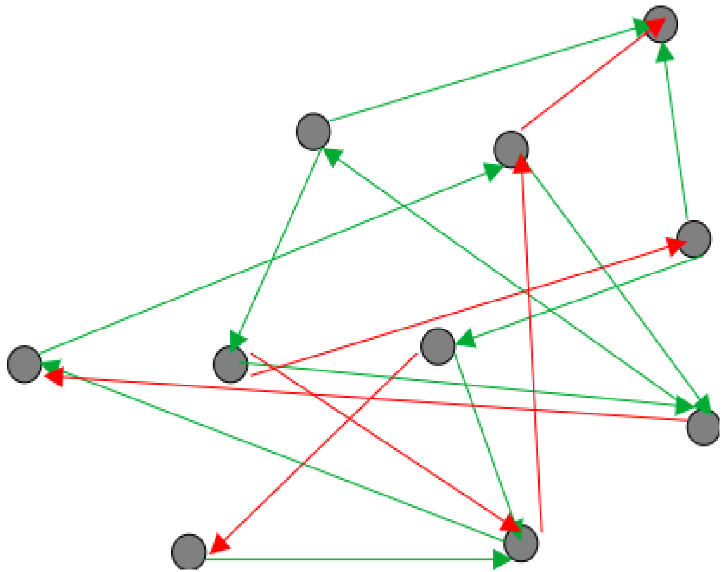
Existing Solutions

- Naive Training Kollias et al. 2022
- Train one model per task Salha et al. 2019
- Ad-hoc architectures Zhang et al. 2021

We wish to develop techniques that allow any NDLP-capable architecture to achieve good performance on all three tasks simultaneously

New Training Strategies

Multi Class



We need to **balance the loss** w.r.t. **positives vs negatives** AND **unidirectionals vs bidirectionals** SIMULTANEOUSLY!

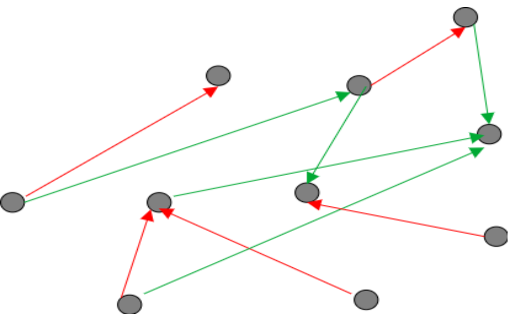
$$
\begin{aligned}
[\hat{p}_{uv}^{nb}, \hat{p}_{uv}^{nu}, \hat{p}_{uv}^{pu}, \hat{p}_{uv}^{pb}] = [(1-\hat{p}_{uv})(1-\hat{p}_{vu}),\\
(1-\hat{p}_{uv})\hat{p}_{vu},\\
\hat{p}_{uv}(1-\hat{p}_{vu}),\\
\hat{p}_{uv}\hat{p}_{vu}]
\end{aligned}
$$

$$\mathcal{L}_{MC\text{-}NDLP}(\Theta) = -\sum_{c \in C} \sum_{uv \in T} w_{y_{uv}} \mathbb{I}(y_{uv} = c) \log(\hat{p}_{uv}^{y_{uv}}), \quad w_{y_{uv}} = \frac{n_x}{n_{y_{uv}}}$$

New Training Strategies

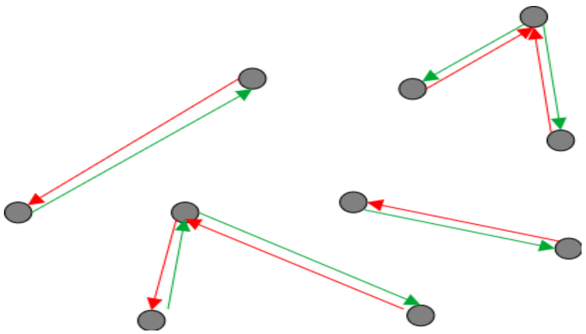
Multi Task - MGDA

General Training Set



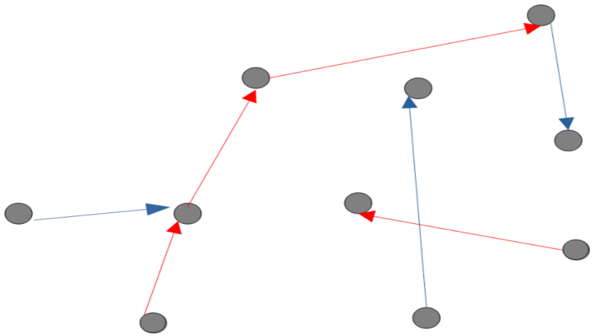
$\nabla \mathcal{L}_G$

Directional Training Set



$\nabla \mathcal{L}_D$

Bidirectional Training Set



$\nabla \mathcal{L}_B$

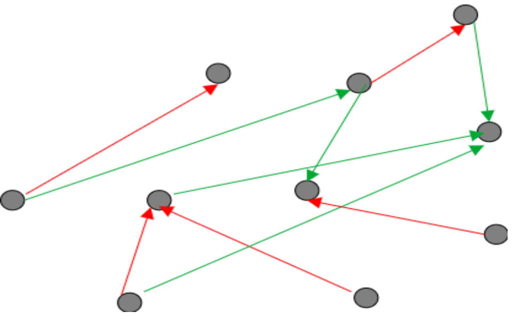
$$\vec{g}_{\mathrm{MGDA}} = \min_{||\sum_i \alpha_i \nabla \mathcal{L}_i||_2^2} \sum_i \alpha_i \nabla \mathcal{L}_i \quad \text{s.t} \quad \sum_i \alpha_i = 1$$

$$\theta' = \theta - \eta \vec{g}_{\mathrm{MGDA}}$$

New Training Strategies

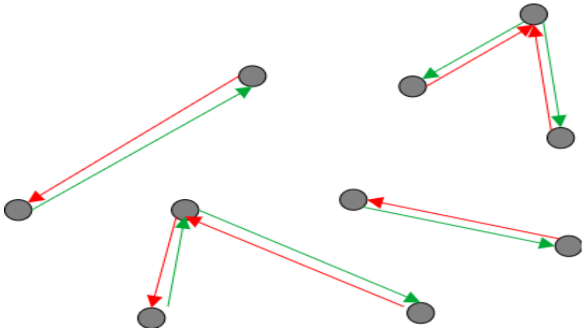
Multi Task - Scalarization

General Training Set



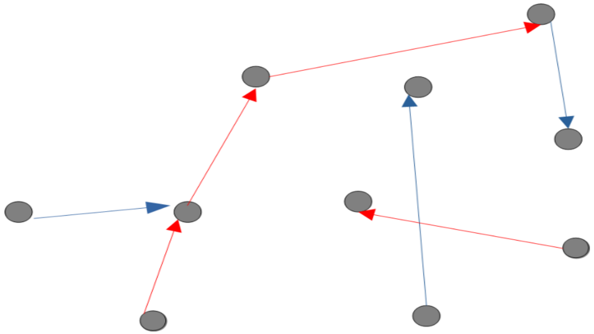
$\nabla \mathcal{L}_G$

Directional Training Set



$\nabla \mathcal{L}_D$

Bidirectional Training Set



$\nabla \mathcal{L}_B$

$$\vec{g}_{\mathrm{S}} = \sum_i \alpha_i \nabla \mathcal{L}_i$$

$$\theta' = \theta - \eta \vec{g}_{\mathrm{S}}$$

New Training Strategies

Results - Cora

		GENERAL		DIRECTIONAL		BIDIRECTIONAL	
model	strategy	ROC-AUC	AUPRC	ROC-AUC	AUPRC	ROC-AUC	AUPRC
GAE	BASELINE	84.6 ± 0.4	88.6 ± 0.3	50.0 ± 0.0	50.0 ± 0.0	62.4 ± 3.0	64.0 ± 3.1
GR-GAE	BASELINE	**89.2 ± 0.4**	**92.4 ± 0.2**	63.4 ± 2.5	61.5 ± 2.7	69.1 ± 3.1	66.5 ± 3.3
	MO-NDLP	84.5 ± 1.1	86.3 ± 1.1	80.6 ± 0.7	80.2 ± 0.9	79.6 ± 4.3	84.6 ± 3.5
	MC-NDLP	88.6 ± 0.4	90.0 ± 0.4	82.1 ± 0.5	**81.8 ± 0.7**	77.3 ± 2.2	76.3 ± 1.7
	S-NDLP	87.8 ± 0.6	89.5 ± 0.5	**82.3 ± 0.5**	81.6 ± 0.4	**89.6 ± 1.6**	**92.4 ± 1.1**
ST-GAE	BASELINE	**87.8 ± 0.7**	**90.1 ± 0.5**	60.8 ± 0.5	64.5 ± 0.6	74.6 ± 1.8	74.1 ± 2.2
	MO-NDLP	86.3 ± 0.5	86.2 ± 0.4	**79.3 ± 1.0**	80.0 ± 0.9	79.3 ± 0.5	79.5 ± 1.9
	MC-NDLP	80.7 ± 2.0	80.1 ± 2.1	79.0 ± 2.3	**81.6 ± 1.9**	70.3 ± 3.0	68.1 ± 2.1
	S-NDLP	84.5 ± 0.4	84.9 ± 0.7	75.8 ± 1.0	78.4 ± 0.9	**81.1 ± 0.9**	**80.4 ± 1.6**
DiGAE	BASELINE	**80.4 ± 1.1**	**85.3 ± 0.8**	57.5 ± 1.3	63.0 ± 1.4	70.4 ± 2.2	68.6 ± 1.2
	MO-NDLP	70.2 ± 3.8	72.6 ± 3.6	73.6 ± 5.4	76.0 ± 4.2	67.3 ± 4.6	69.6 ± 4.1
	MC-NDLP	75.4 ± 0.9	77.4 ± 1.0	**84.3 ± 0.6**	**85.4 ± 0.8**	68.9 ± 1.5	69.3 ± 1.1
	S-NDLP	72.5 ± 4.0	77.4 ± 4.4	61.6 ± 1.3	69.2 ± 1.4	**72.1 ± 5.6**	**74.4 ± 5.7**
MLP-GAE	BASELINE	**77.1 ± 0.9**	**78.2 ± 0.6**	90.7 ± 0.6	90.7 ± 0.6	69.9 ± 3.2	69.7 ± 3.7
	MO-NDLP	76.0 ± 0.8	76.4 ± 0.7	93.4 ± 0.6	93.5 ± 0.6	**80.7 ± 1.6**	**79.2 ± 2.4**
	MC-NDLP	74.5 ± 0.7	75.6 ± 0.7	**94.3 ± 0.6**	**94.4 ± 0.5**	71.7 ± 2.4	65.7 ± 1.8
	S-NDLP	74.7 ± 1.0	74.9 ± 0.9	90.5 ± 0.7	90.0 ± 0.9	72.0 ± 2.6	70.5 ± 2.9
MAGNET	BASELINE	**75.2 ± 1.4**	**77.8 ± 1.0**	90.4 ± 0.9	89.8 ± 0.8	**71.9 ± 2.3**	**70.4 ± 2.8**
	MO-NDLP	74.4 ± 1.4	77.4 ± 1.1	91.3 ± 1.0	90.9 ± 1.0	70.6 ± 2.7	68.6 ± 2.7
	MC-NDLP	74.4 ± 1.0	77.4 ± 1.0	**92.1 ± 0.7**	**91.6 ± 0.7**	71.8 ± 2.6	70.0 ± 2.6
	S-NDLP	74.6 ± 1.3	77.5 ± 1.1	91.0 ± 1.0	90.4 ± 1.0	71.8 ± 2.8	70.2 ± 2.9
dMPLP	BASELINE	**86.1 ± 0.5**	**88.0 ± 0.9**	75.7 ± 2.2	76.8 ± 1.6	81.1 ± 3.6	82.2 ± 5.3
	MO-NDLP	83.5 ± 0.6	85.1 ± 0.6	**89.1 ± 1.7**	**89.0 ± 2.1**	**85.8 ± 3.3**	**89.3 ± 2.5**
	MC-NDLP	81.4 ± 1.7	82.0 ± 1.5	83.7 ± 4.2	83.7 ± 3.6	70.0 ± 4.3	71.5 ± 3.9
	S-NDLP	85.6 ± 1.0	86.9 ± 1.0	84.8 ± 2.7	86.3 ± 2.3	83.6 ± 4.7	87.0 ± 4.3

Thanks!