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

Claudio Moroni(✉)[1,2], Claudio Borile[2], Carolina Mattsson[2], Michele Starnini[2,3], André Panisson[2]

[1]Dedagroup, Torino, Italy, claudio.moroni@dedagroup.it

[2]CENTAI Institute, Torino, Italy, {name.surname}@centai.eu

[3]Department of Engineering, Universitat Pompeu Fabra, Barcelona, 08018, Spain

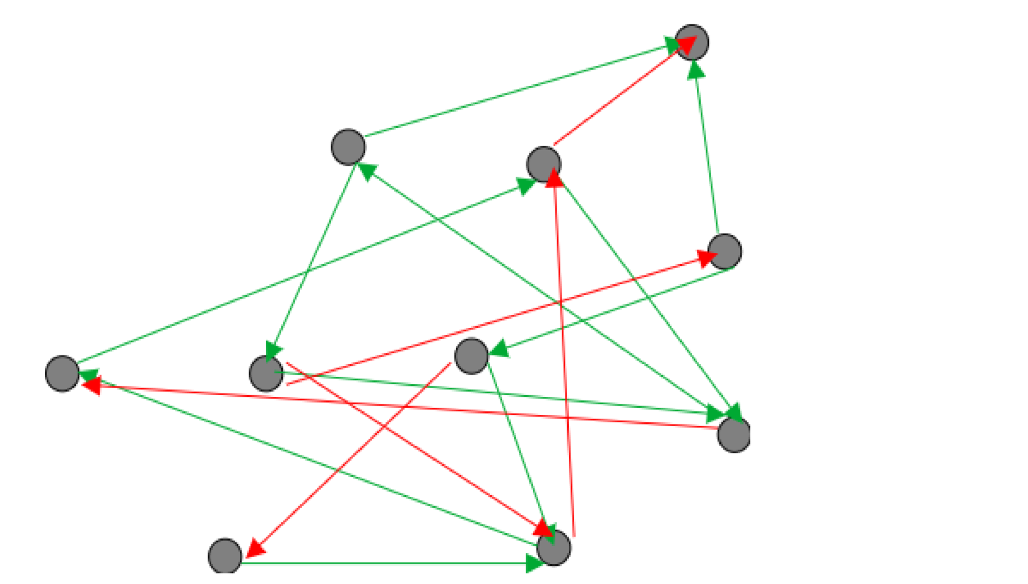
CENTAI

deda.ai

(NAIVE) DIRECTED LINK PREDICTION

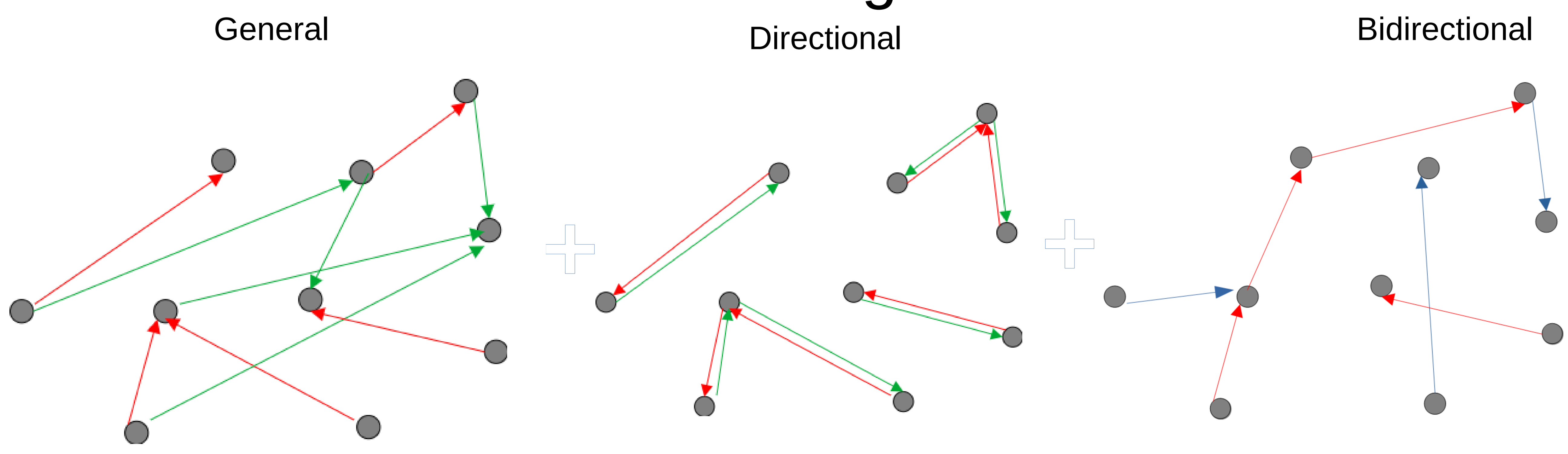
In the context of Graph Machine Learning, Link Prediction (LP) is a foundational task. Graph Neural Networks (GNN) perform LP by decoding pairs of node embeddings into a probability that an edge between them exists. We find that LP's directed variant (DLP) has been poorly explored. Infact, evaluating DLP models require a nuanced setup made of three test sets, respectively measuring the model's ability to predict edge's existence, direction and bidirectionality. Nevertheless, we find that training general DLP models using the standard contrastive setup as in Undirected Link Prediction (ULP) does not work. A few prior works have, at best, trained specialized models for each test set, leaving the question open on how get good performance across the three test sets out of a general DLP-capable GNN model.

Training



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

Testing



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	$\sigma(\vec{z}_v[0] - \lambda \ln(\lvert\lvert \vec{z}_u[1:] - \vec{z}_v[1:]\rvert\rvert_2^2))$	89.2	92.4	63.4	61.5	69.1	66.5
MLP-GAE	MLP($\vec{z}_v \lvert\lvert \vec{z}_u$)	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

NEW TRAINING STRATEGIES

We devise three training strategies that allow a general DLP-capable GNN model to perform well across the three sub tasks simultaneously. The first strategy, Multi-Class NDLP (MC-NDLP) recognizes that the underwhelming performances reported above are due to an imbalance between uni- and bi-directional edges in the naive training set, and makes up for it by mapping NDLP to a four-class classification task where the loss contributions of uni- and bi-directional edges are rebalanced together with that of positives and negatives. The second and third tasks instead split the training set in three sets whose composition mimics that of the three test sets, and apply multi-task learning techniques, respectively MGDA and Scalarization.

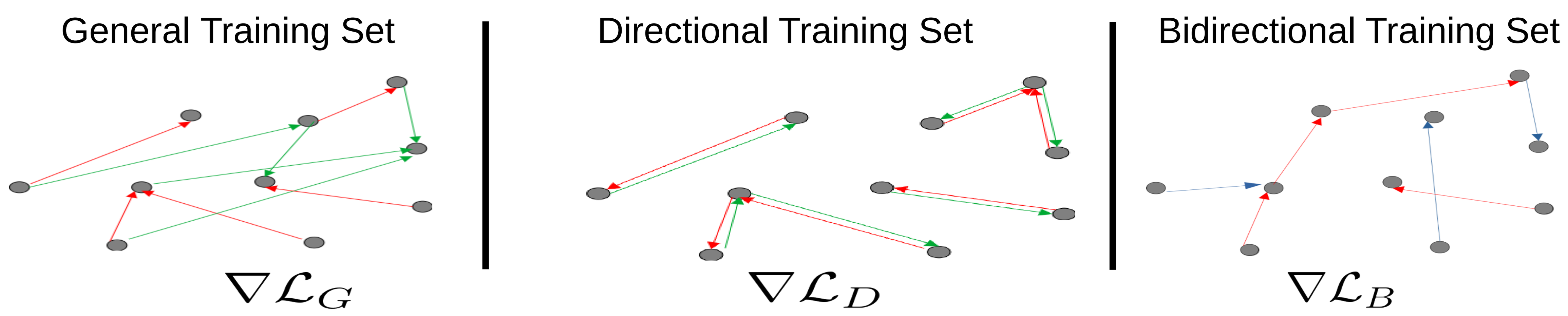
Multi Class

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

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

MGDA

Scalarization



$$\vec{g}_{\text{MGDA}} = \min_{\lvert\lvert \sum_i \alpha_i \nabla\mathcal{L}_i \rvert\rvert_2^2} \sum_i \alpha_i \nabla\mathcal{L}_i \quad \text{s.t} \quad \sum_i \alpha_i = 1$$

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

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

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

model	strategy	GENERAL ROC-AUC	GENERAL AUPRC	DIRECTIONAL ROC-AUC	DIRECTIONAL AUPRC	BIDIRECTIONAL ROC-AUC	BIDIRECTIONAL AUPRC
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**
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

- Salha, G., Limnios, S., Hennequin, R., Tran, V.A., Vazirgiannis, M.: Gravityinspired graph autoencoders for directed link prediction. In: Proceedings of the 28th ACM International Conference on Information and Knowledge Management. pp. 589–598 (2019)
- Zhang, X., He, Y., Brugnone, N., Perlmutter, M., Hirn, M.: Magnet: A neural network for directed graphs. In: Advances in Neural Information Processing Systems (2021)
- Kollias, G., Kalantzis, V., Id'e, T., Lozano, A.C., Abe, N.: Directed graph autoencoders. In: AAAI Conference on Artificial Intelligence (2022)
- Désidéri, J.A.: Multiple-gradient descent algorithm (mgda) for multiobjective optimization. Comptes Rendus Mathematique 350(5), 313–318 (2012)