INDIGO: GNN-Based Inductive Knowledge Graph Completion Using Pair-Wise Encoding

Shuwen Liu[1], **Bernardo Cuenca Grau**[1], **Ian Horrocks**[1], **and Egor V. Kostylev**[2]
[1]Department of Computer Science, University of Oxford, UK
{shuwen.liu, bernardo.cuenca.grau, ian.horrocks}@cs.ox.ac.uk
[2]Department of Informatics, University of Oslo
egork@ifi.uio.no

A Proof of Proposition 1

Proposition 1. *Let r be a rule of form $\forall \mathbf{x}.\, B_1, \ldots, B_n \rightarrow H$, let C be a set of $|\mathbf{x}|$ constants, and let Σ be the set of all assignments of constants from C to $\mathbf{x}$. A constant-agnostic completion function f captures r if and only if $f(\mathcal{K}_r^\sigma, t_r^\sigma)$ is true for each $\sigma \in \Sigma$.*

Proof. The "only if" direction of the claim holds by definition, so we concentrate on the "if" direction. Let f be a constant-agnostic completion function such that $f(\mathcal{K}_r^\sigma, t_r^\sigma)$ is true for each $\sigma \in \Sigma$. To show that f captures r, consider an arbitrary assignment σ' of constants to $\mathbf{x}$ (which is not necessarily in Σ). Fix an arbitrary renaming ϱ of constants in the range of σ' by constants in C, which exists since $|C| = |\mathbf{x}|$. On the one hand, we have

$$f(\mathcal{K}_r^{\sigma'}, t_r^{\sigma'}) = f(\varrho(\mathcal{K}_r^{\sigma'}), \varrho(t_r^{\sigma'})),$$

since f is constant-agnostic. On the other hand,

$$f(\varrho(\mathcal{K}_r^{\sigma'}), \varrho(t_r^{\sigma'})) = f(\mathcal{K}_r^{\sigma}, t_r^{\sigma})$$

by construction, where σ is the substitution from Σ that is the composition of σ' and ϱ. Thus, the claim follows by the assumption that $f(\mathcal{K}_r^\sigma, t_r^\sigma)$ is true. $\square$

B Construction of INDIGO-BM

Our INDIGO-BM benchmark was constructed in the following way. We first collected from Freebase all the type triples for constants contained in FB15K-237, and merged them with the triples in FB15K-237. Second, we randomly sampled $1,000$ triples from FB15K-237, and set all the constants contained in these triples as *unseen constants*. Third, for the triples containing no unseen constants, we split them into a training set $\mathcal{T}$ and a validation set $\mathcal{V}$ with a ratio of 9:1. Finally, the triples containing at least one unseen constant were split into an incomplete graph $\mathcal{K}$ and a set $\Lambda_{\mathsf{test}}^{+}$ of positive test triples with a ratio of 9:1. To simulate a KG evolution scenario, we took $\mathcal{K}_{\mathsf{test}} = \mathcal{K} \cup \mathcal{T}$.

C Definition of F1-score and AUC

Besides other metrics, we use *F1-score* (i.e., *balanced F-score*) $2 \cdot prec \cdot rec/(prec + rec)$, which is based on precision $prec$ and recall rec for given numbers of true and false positives and negatives. Moreover, we also use *the area under the precision-recall curve* (*AUC*), which is defined as follows. A precision-recall curve is a graphical plot with coordinates x and y, where coordinate x corresponds to the recall, coordinate y corresponds to the precision, and each point in the plot represents the precision and recall values for a specific threshold, which are computed using confidence-based

35th Conference on Neural Information Processing Systems (NeurIPS 2021).

Table 4: Additional classification-based metric results on the benchmarks in %; R, G, H, and I stand for R-GCN, GraIL, the system of Hamaguchi et al., and INDIGO, respectively

Bench-			Precision				Recall				F1-Score			
mark			R	G	H	I	R	G	H	I	R	G	H	I
GraIL-BM	FB15K-237	v1	51.1	41.5	-	**92.2**	47.3	**92.2**	-	75.0	49.1	57.1	-	**82.7**
		v2	51.5	62.7	-	**95.7**	44.4	**95.8**	-	82.4	47.6	75.8	-	**88.5**
		v3	54.0	63.9	-	**95.1**	66.8	**97.0**	-	82.2	59.7	77.0	-	**88.2**
		v4	51.4	63.7	-	**93.8**	77.9	**92.6**	-	81.0	61.9	75.5	-	**86.9**
	NELL-995	v1	29.3	**96.4**	-	78.2	93.8	98.2	-	**99.0**	44.6	**97.3**	-	87.3
		v2	51.8	38.7	-	**94.6**	57.6	**95.6**	-	72.4	54.5	55.1	-	**82.0**
		v3	52.1	49.3	-	**95.6**	56.9	**98.7**	-	83.3	54.4	65.8	-	**89.0**
		v4	52.7	61.1	-	**85.8**	71.3	49.7	-	**84.5**	60.6	54.8	-	**85.1**
	WN18RR	v1	50.2	79.0	-	**79.4**	62.2	**98.0**	-	96.7	55.6	**87.5**	-	87.2
		v2	52.9	62.6	-	**82.0**	50.1	**99.6**	-	91.7	51.5	76.9	-	**86.3**
		v3	51.2	54.8	-	**85.7**	**94.9**	94.1	-	82.3	66.5	69.3	-	**84.0**
		v4	48.5	74.0	-	82.1	49.8	**98.3**	-	90.8	49.1	84.5	-	**86.2**
Hamaguchi-BM		h-1K	45.0	36.9	**82.4**	75.7	63.1	49.1	**85.6**	74.9	52.5	42.1	**84.0**	75.2
		h-3K	45.8	37.9	77.1	**82.6**	64.5	50.3	**82.7**	78.1	53.6	43.2	79.8	**80.3**
		h-5K	47.5	38.4	77.7	**82.2**	65.5	50.1	83.4	**85.2**	55.0	43.5	80.5	**83.7**
		t-1K	49.8	45.3	79.0	**81.3**	64.8	54.2	73.8	**82.5**	56.4	49.4	76.3	**82.0**
		t-3K	44.8	44.4	76.0	**86.1**	64.7	53.8	73.9	**86.0**	53.0	48.6	74.9	**85.9**
		t-5K	43.7	44.4	73.3	**86.0**	63.3	54.2	75.7	**88.6**	51.7	48.8	74.5	**87.3**
		b-1K	29.7	46.5	**83.3**	81.9	40.5	54.8	87.7	**95.2**	34.3	50.3	85.4	**88.0**
		b-3K	33.5	46.2	75.3	**83.6**	46.0	54.5	87.3	**91.8**	38.7	50.0	80.9	**87.5**
		b-5K	34.0	43.2	71.1	**86.9**	46.0	53.3	86.9	**91.6**	39.1	47.7	78.3	**89.2**
INDIGO-BM			66.8	77.5	-	**98.1**	**95.0**	93.9	-	90.5	78.4	84.9	-	**94.1**

predictions—that is, for a given threshold θ, the plot has a point $(r(\theta), p(\theta))$, where $r(\theta)$ and $p(\theta)$ are the recall and the precision for θ. Then, relying on the fact that functions r and p are monotonic, AUC is $\int_{x=0}^{1} p(r^{-1}(x))dx$.

D Additional Evaluation Results

Table 4 and Table 5 show the results on the benchmarks for additional classification-based metrics (precision, recall, and F1) and ranking-based metrics (e-Hits@1, e-Hits@10, e-MRR, r-Hits@1, r-Hits@10, and r-MRR), respectively. As we can see, INDIGO also significantly outperforms the baselines on almost all the additional metrics. In Table 6, we also report the variance of each metric. Recall that, when computing r-Hits@k and r-MRR, we took all possible relevant samples for constructing negative examples; so the variance is not applicable in these cases and we omitted it in the table. As we can see, the variance was very small for all the metrics on the benchmarks.

E Training and Testing Time

Table 7 and Table 8 show the statistics of training time and testing time for all the systems on the benchmarks. Compared with Hamaguchi et al. and GraIL, INDIGO takes less time to both train and test; R-GCN is also slower that INDIGO in testing but comparable in training (note that testing time is much smaller that training and thus should not be generally considered as a limiting factor for a system).

F Potential Negative Societal Impacts

KG completion may potentially lead to harmful leakage of information in applications involving datasets containing sensitive data. Preventing such situations should be a concern when applying this technology in practice in some use cases, and further research in this direction might be needed.

Furthermore, predictions made by a KG completion system may be unknowingly biased, which may be an issue in certain applications such as recommendation systems. This is related to the broader issues of fairness and explainability of ML-based systems. In this regard, however, we see the ability of extracting human-readable rules as an advantage of GNN-based KG completion approaches over alternative solutions.

Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
   (b) Did you describe the limitations of your work? [Yes] See Section 5.
   (c) Did you discuss any potential negative societal impacts of your work? [Yes] See the supplemental material.
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [N/A]
   (b) Did you include complete proofs of all theoretical results? [N/A]
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [Yes] See the supplemental material.
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes] See Section 4.
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [N/A] We test the systems with 10 runs and report the variance of the result.
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [Yes] See Section 4.
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? [Yes]
   (b) Did you mention the license of the assets? [Yes] See the supplemental material.
   (c) Did you include any new assets either in the supplemental material or as a URL? [Yes] For benchmarks, we provide all the existing benchmarks we used, and the INDIGO-BM we created. For the code, we provide all the code of INDIGO.
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A] All the data we used are publicly available.
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A] We didn't contain that information.
5. If you used crowdsourcing or conducted research with human subjects...
   (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]
   (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]

Table 5: Additional ranking-based metric results on the benchmarks in %; R, G, H, and I stand for R-GCN, GraIL, the system of Hamaguchi et al., and INDIGO, respectively

Bench-			e-Hits@1				e-Hits@10				e-MRR			
mark			R	G	H	I	R	G	H	I	R	G	H	I
GraIL-BM	FB15K-237	v1	7.3	**34.4**	-	32.7	47.3	55.4	-	**55.6**	19.1	**42.0**	-	41.6
		v2	7.6	**48.5**	-	25.8	52.2	**82.9**	-	48.9	20.4	**60.8**	-	34.6
		v3	4.4	56.3	-	**46.4**	46.7	**85.0**	-	46.9	16.8	**66.1**	-	32.9
		v4	5.8	**44.4**	-	26.1	49.1	**84.5**	-	49.7	18.5	**56.9**	-	35.0
	NELL-995	v1	3.0	**49.0**	-	35.5	22.1	**57.0**	-	44.0	10.3	**52.6**	-	40.2
		v2	3.7	**53.3**	-	29.6	40.1	**90.8**	-	56.6	15.3	**66.9**	-	39.8
		v3	7.2	**65.0**	-	35.5	46.7	**93.6**	-	56.1	18.7	**75.8**	-	43.3
		v4	3.0	**51.8**	-	34.3	31.7	57.3	-	**67.8**	12.6	**55.4**	-	46.3
	WN18RR	v1	6.1	**71.3**	-	3.5	72.1	**84.0**	-	30.3	22.2	**77.3**	-	13.0
		v2	7.1	**78.1**	-	8.2	65.4	**81.6**	-	41.2	21.8	**80.2**	-	18.8
		v3	6.2	**50.3**	-	21.9	61.2	**58.4**	-	55.5	20.0	**54.2**	-	32.4
		v4	2.7	**74.7**	-	9.5	29.0	**76.3**	-	25.0	12.0	**76.2**	-	16.0
Hamaguchi-BM		h-1K	20.7	4.9	**33.1**	21.1	44.7	44.0	**78.5**	49.7	29.7	14.8	**48.2**	31.2
		h-3K	15.8	8.8	**28.5**	22.2	45.8	49.3	**72.9**	52.6	26.7	18.8	**42.8**	32.8
		h-5K	10.6	11.8	**30.1**	21.8	42.0	50.6	**73.4**	50.3	21.8	21.6	**44.1**	31.7
		t-1K	9.1	14.8	**25.8**	22.7	34.7	50.8	**68.4**	42.4	18.1	24.1	**39.3**	30.2
		t-3K	9.4	14.0	17.9	**24.9**	46.1	31.9	**56.9**	45.8	21.9	20.7	30.1	**32.8**
		t-5K	7.8	17.0	18.2	**27.5**	50.8	50.6	**56.2**	49.4	20.6	26.3	30.2	**35.7**
		b-1K	30.8	17.0	28.9	**32.5**	**62.5**	26.5	46.7	43.2	**49.7**	26.7	44.0	40.5
		b-3K	12.7	12.7	21.3	**29.5**	36.0	22.3	37.1	**37.5**	31.4	21.6	34.8	**37.4**
		b-5K	8.5	15.5	17.0	**29.3**	24.3	26.3	30.6	**39.9**	24.3	25.6	29.7	**37.8**
INDIGO-BM			18.0	**53.7**	-	43.1	64.5	**78.8**	-	61.2	32.4	**63.3**	-	50.3

Bench-			r-Hits@1				r-Hits@10				r-MRR			
mark			R	G	H	I	R	G	H	I	R	G	H	I
GraIL-BM	FB15K-237	v1	1.4	0.5	-	**36.4**	4.4	6.3	-	**75.2**	4.0	3.5	-	**48.9**
		v2	1.1	0.2	-	**42.0**	6.7	4.6	-	**85.3**	4.3	2.6	-	**57.1**
		v3	0.6	1.8	-	**48.1**	6.5	20.7	-	**84.1**	3.8	8.5	-	**59.8**
		v4	1.5	0.8	-	**42.0**	6.6	10.1	-	**83.0**	4.2	5.0	-	**56.4**
	NELL-995	v1	**16.0**	0.0	-	13.0	92.0	87.0	-	**100.0**	32.5	14.5	-	**50.0**
		v2	0.2	2.7	-	**46.5**	4.8	17.2	-	**80.3**	3.4	9.6	-	**56.2**
		v3	0.5	0.7	-	**43.4**	8.5	10.1	-	**81.2**	3.5	4.7	-	**56.2**
		v4	2.1	0.0	-	**40.1**	10.0	7.3	-	**81.7**	6.1	3.9	-	**50.8**
	WN18RR	v1	0.5	0.5	-	**67.5**	**100.0**	**100.0**	-	**100.0**	16.1	17.6	-	**82.8**
		v2	6.6	2.0	-	**51.7**	**100.0**	**100.0**	-	**100.0**	20.7	23.0	-	**74.2**
		v3	15.5	3.5	-	**64.2**	95.9	98.3	-	**99.8**	30.6	22.4	-	**78.2**
		v4	3.1	16.3	-	**37.0**	**100.0**	**100.0**	-	**100.0**	19.4	29.8	-	**66.6**
Hamaguchi-BM		h-1K	7.9	0.2	19.6	**45.4**	94.1	97.0	97.8	**100.0**	28.6	20.5	39.3	**64.7**
		h-3K	7.8	1.0	14.4	**44.7**	95.3	97.8	98.2	**99.8**	27.6	22.0	35.5	**65.2**
		h-5K	9.1	0.8	11.8	**38.4**	95.1	89.7	98.8	**99.9**	29.0	21.1	33.4	**63.1**
		t-1K	7.2	0.2	14.8	**36.3**	96.6	97.3	97.6	**100.0**	26.6	21.7	36.1	**62.5**
		t-3K	10.0	0.3	12.1	**45.0**	94.3	97.2	98.4	**100.0**	29.1	21.3	32.4	**67.6**
		t-5K	6.6	0.7	12.7	**42.7**	93.6	86.3	98.1	**99.9**	26.0	20.7	33.9	**66.6**
		b-1K	7.4	0.2	13.2	**50.2**	96.2	96.6	98.7	**100.0**	25.8	21.9	33.7	**72.1**
		b-3K	5.7	0.3	14.6	**43.1**	95.4	88.6	97.1	**99.8**	24.6	20.4	35.6	**67.4**
		b-5K	6.2	1.2	10.8	**49.7**	96.0	88.8	97.3	**100.0**	25.4	20.8	31.0	**71.3**
INDIGO-BM			17.6	3.4	-	**54.1**	54.1	10.2	-	**92.4**	29.3	6.4	-	**67.8**

Table 6: Variance of results on the benchmarks in %; R, G, H, and I stand for R-GCN, GraIL, the system of Hamaguchi et al., and INDIGO, respectively

Bench-			Accuracy				AUC				Precision				Recall			
mark			R	G	H	I	R	G	H	I	R	G	H	I	R	G	H	I
GraIL-BM	FB15K-237	v1	.00	.04	-	.00	.00	.03	-	.00	.00	.17	-	.02	.00	.05	-	.00
		v2	.00	.00	-	.00	.00	.00	-	.00	.00	.00	-	.00	.00	.01	-	.00
		v3	.00	.00	-	.00	.00	.00	-	.00	.00	.00	-	.01	.00	.01	-	.00
		v4	.00	.00	-	.01	.00	.01	-	.01	.00	.00	-	.04	.00	.01	-	.00
	NELL-995	v1	.00	.01	-	.05	.00	.01	-	.02	.00	.01	-	.07	.00	.01	-	.00
		v2	.00	.02	-	.00	.00	.01	-	.00	.00	.02	-	.01	.00	.01	-	.00
		v3	.00	.00	-	.00	.00	.00	-	.00	.00	.00	-	.01	.00	.00	-	.00
		v4	.00	.01	-	.01	.00	.00	-	.01	.00	.00	-	.02	.00	.00	-	.00
	WN18RR	v1	.00	.00	-	.05	.00	.01	-	.05	.00	.00	-	.07	.00	.01	-	.00
		v2	.00	.00	-	.00	.00	.00	-	.01	.00	.00	-	.01	.00	.00	-	.00
		v3	.00	.00	-	.01	.00	.01	-	.00	.00	.00	-	.03	.00	.02	-	.00
		v4	.00	.00	-	.02	.00	.00	-	.02	.00	.00	-	.06	.00	.00	-	.00
Hamaguchi-BM		h-1K	.00	.01	.00	.02	.00	.00	.01	.03	.00	.00	.01	.04	.00	.02	.00	.00
		h-3K	.00	.00	.00	.01	.00	.00	.01	.01	.00	.00	.01	.04	.00	.00	.00	.00
		h-5K	.00	.00	.00	.01	.00	.00	.01	.00	.00	.00	.01	.02	.00	.00	.00	.00
		t-1K	.00	.01	.01	.02	.00	.01	.01	.04	.00	.00	.02	.04	.00	.02	.01	.00
		t-3K	.00	.01	.00	.09	.00	.00	.00	.01	.00	.00	.01	.03	.00	.01	.00	.00
		t-5K	.00	.00	.02	.06	.00	.00	.03	.00	.00	.00	.05	.01	.00	.00	.00	.00
		b-1K	.00	.01	.00	.03	.00	.01	.01	.02	.00	.00	.01	.06	.00	.02	.00	.00
		b-3K	.00	.00	.01	.03	.00	.00	.01	.01	.00	.00	.02	.08	.00	.00	.00	.00
		b-5K	.00	.00	.01	.03	.00	.00	.01	.00	.00	.00	.01	.08	.00	.00	.00	.00
INDIGO-BM			.00	.00	-	.00	.00	.00	-	.00	.00	.00	-	.00	.00	.00	-	.00

Bench-			e-Hits@1				e-Hits@3				e-Hits@10				e-MRR			
mark			R	G	H	I	R	G	H	I	R	G	H	I	R	G	H	I
GraIL-BM	FB15K-237	v1	.00	.01	-	.02	.00	.00	-	.02	.00	.00	-	.01	.00	.00	-	.01
		v2	.00	.01	-	.05	.00	.00	-	.03	.00	.00	-	.03	.00	.00	-	.02
		v3	.00	.01	-	.00	.00	.00	-	.00	.00	.00	-	.00	.00	.00	-	.00
		v4	.00	.01	-	.00	.00	.00	-	.00	.00	.00	-	.00	.00	.00	-	.01
	NELL-995	v1	.00	.01	-	.00	.00	.01	-	.00	.00	.04	-	.00	.00	.01	-	.00
		v2	.00	.01	-	.02	.00	.00	-	.01	.00	.00	-	.01	.00	.00	-	.01
		v3	.00	.00	-	.00	.00	.01	-	.00	.00	.01	-	.00	.00	.00	-	.00
		v4	.00	.01	-	.00	.00	.00	-	.01	.00	.00	-	.00	.00	.00	-	.00
	WN18RR	v1	.00	.01	-	.03	.00	.00	-	.02	.00	.00	-	.01	.00	.00	-	.01
		v2	.00	.00	-	.00	.00	.00	-	.00	.00	.00	-	.00	.00	.00	-	.00
		v3	.00	.00	-	.00	.00	.00	-	.00	.00	.00	-	.00	.00	.00	-	.00
		v4	.00	.00	-	.00	.00	.00	-	.00	.00	.00	-	.00	.00	.00	-	.00
Hamaguchi-BM		h-1K	.00	.00	.03	.02	.00	.00	.02	.01	.00	.00	.01	.00	.00	.00	.01	.00
		h-3K	.00	.00	.01	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		h-5K	.00	.00	.00	.01	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		t-1K	.00	.00	.02	.00	.00	.00	.01	.00	.00	.00	.01	.00	.00	.00	.01	.00
		t-3K	.00	.00	.00	.01	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		t-5K	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.01
		b-1K	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
		b-3K	.00	.00	.00	.02	.00	.00	.00	.01	.00	.00	.00	.01	.00	.00	.00	.01
		b-5K	.00	.00	.01	.00	.00	.00	.00	.00	.00	.00	.01	.01	.00	.00	.00	.00
INDIGO-BM			.00	.00	.00	.01	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00

Table 7: Training and testing time on GraIL-BM (in hours)

	Model	GraIL-BM / FB15K-237 v1	v2	v3	v4	GraIL-BM / NELL-995 v1	v2	v3	v4	GraIL-BM / WN18RR v1	v2	v3	v4
Train	R-GCN	**1.11**	**2.61**	6.97	17.33	1.41	3.34	6.78	3.50	0.56	1.13	2.59	0.83
	GraIL	6.63	47.11	334.83	471.97	1.92	31.32	254.15	172.96	0.72	1.55	2.98	1.68
	INDIGO	1.31	3.13	**6.64**	**9.95**	**1.21**	**2.34**	**4.83**	**2.25**	**0.32**	**0.66**	**0.89**	**0.45**
Test	R-GCN	**0.002**	**0.004**	**0.01**	**0.02**	**0.004**	**0.01**	**0.02**	**0.02**	**0.003**	**0.01**	**0.02**	**0.03**
	GraIL	0.96	2.95	9.24	12.97	0.05	1.01	3.73	1.22	0.08	0.12	0.18	0.28
	INDIGO	0.04	0.13	0.29	0.44	0.03	0.12	0.26	0.24	0.02	0.04	0.11	0.11

Table 8: Training and testing time on Hamaguchi-BM and INDIGO-BM (in hours)

	Model	Hamaguchi-BM / h 1k	3k	5k	Hamaguchi-BM / t 1k	3k	5k	Hamaguchi-BM / b 1k	3k	5k	INDIGO-BM
Train	R-GCN	1.76	1.31	1.21	1.18	1.09	0.90	1.23	0.78	0.55	45.83
	GraIL	9.38	7.72	7.47	8.59	5.87	3.75	9.52	3.68	4.52	69.59
	Hamaguchi et al.	1.33	1.40	1.08	1.07	1.09	0.94	1.54	1.01	0.92	-
	INDIGO	**1.14**	**1.06**	**0.95**	**0.79**	**0.48**	**0.41**	**0.74**	**0.45**	**0.37**	**39.68**
Test	R-GCN	**0.02**	**0.06**	0.12	**0.03**	**0.10**	**0.17**	**0.04**	**0.12**	**0.18**	**2.46**
	GraIL	0.08	0.34	0.44	0.09	0.31	0.56	0.10	0.33	0.59	49.65
	Hamaguchi et al.	0.07	0.21	0.35	0.08	0.22	0.37	0.08	0.21	0.35	-
	INDIGO	**0.02**	**0.06**	**0.11**	0.09	0.18	0.21	0.10	0.19	0.20	3.56