

Annex: Dataset Recommendation for Heterogeneous and Dynamic Scholarly Graphs via Multimodal Representation Learning

Anonymous Author(s)

1 Results

In this section, we expand the analysis by presenting a more comprehensive set of results. Specifically, we compare the performance of MM-SAN against the graph-based baselines at three different cutoff levels: 5, 10, and 20. We also evaluate the models using a broader range of metrics, including hits, recall (R), precision (P), F1-score (F1), and mean reciprocal rank (MRR). In the original paper, we reported results at cutoff 5 using NDCG and recall, as these were representative of the model’s overall performance. On the other hand, in this extended analysis, we provide a more detailed evaluation to better capture the model’s behavior across different recommendation depths and performance dimensions. We report two sets of tables. The first set, Tables 1, 2, 3, 4, 5, 6 reports the performances of GAT, SAGE, HAN, HGT and MM-SAN on MES and PubMed datasets. In the same way as the original paper, we evaluated the methods effectiveness in transductive, semi-inductive and fully inductive setups, in different metadata conditions: with full datasets metadata available (ideal condition) and with 75%, 50%, 25% available metadata. The results confirm the considerations done in the original paper: all models, including MM-SAN, perform best in the transductive setting when full metadata is available (100% row). However, this scenario is least representative of real-world academic contexts. MM-SAN maintains reliable performance even in more challenging semi-inductive and full-inductive settings, where predictions are made for previously unseen nodes. Under more realistic conditions, where metadata availability is limited, MM-SAN, like other models, experiences a slight decline in performance as the proportion of datasets with metadata decreases (from 75% to 25%). This trend underscores the importance of text-based features for enhancing recommendation accuracy and highlights MM-SAN’s flexibility in handling situations with sparse metadata.

In the second set of experiments (Tables 7, 8, 9, 10, 11, 12), we present the performance results for SAGE, GAT, HGNN, their enhanced versions (where the original sampling method is replaced with ours), and MM-SAN. For each baseline and its enhanced version, we report the percentage gain to highlight the impact of our sampling method on each baseline. We observe a significant improvement across all methods when replacing the original sampling approach with ours. This improvement is attributed to our method’s focus on enhancing heterogeneity and selecting the most representative nodes of the target. In contrast, methods like SAGE and GAT, which are designed for homogeneous graphs, sample neighboring nodes randomly, potentially leading to less representative samples due to imbalance node types distribution.

2 Baselines

We applied grid search to systematically tune the hyperparameters of each baseline model for fair comparison. In Table 13, we report

the specific choices and value ranges explored for each hyperparameter and each baseline.

3 Sampling Analysis

In this set of experiments, we investigate the performance of MM-SAN when replacing the original random walk-based sampling method with random sampling. We explored the effect of random sampling at different hop distances (k-hops) to evaluate how sampling neighbors from progressively deeper levels influences the results. The results for MES and PubMed datasets at cutoff 5 are shown in Tables 14, 15 respectively.

We observe that MM-SAN, in its original implementation using random walk-based sampling, outperforms random sampling at every selected hop. However, two opposite tendencies can be observed. If we analyse the results of MM-SAN with random sampling at three different hops (RS-1, RS-2, RS-3 in Table 15) we can see two opposite behaviors of MM-SAN on the two datasets tested. In the MES dataset, which is more connected, well-curated, and smaller, representing an ideal condition, we see that as we increase the number of hops, the sample becomes more representative. Therefore, in ideal cases, it is clear that the deeper we search, the more representative the sample becomes. This is not true for PubMed, which represents a much more sparse and realistic dataset. In this case, there is no significant difference between hop-1 and hop-3 sampling. This can be attributed to the sparsity of the graph, the lack of disambiguation among authors, and the high probability that the sampled nodes are all of the same type and do not contain useful information for building a representative feature.

4 Augmentation Analysis

In this section, we expand on the component-based analysis presented in the original paper and report results on the MES and PubMed datasets in Tables 16 and 17. We evaluate five graph-based baselines—GAT, SAGE, HAN, HGT, and HGNN—along with modified versions that incorporate our sampling strategy (denoted as GAT-RW, SAGE-RW, and HGNN-RW).

Consistent with findings from the original study, we observe that the augmentation phase improves the performance of HGNN and MM-SAN. This improvement can be attributed to their sensitivity to sample heterogeneity; the inclusion of new node types enhances their ability to capture complex relationships. In contrast, the original implementations of SAGE and GAT do not benefit from node augmentation, likely because they are designed for homogeneous graphs and struggle to exploit the added heterogeneity.

Interestingly, in the methods implementing our sampling mechanism – with the suffix -RW, we see that the augmentation improves the overall performances, suggesting that our method allows these models to better accommodate heterogeneous graph structures. As

Table 1: Recall@5 (R@5), nDG@5 (N@5), Precision@5 (P@5), f1-score (f1@5), Mean reciprocal rank (MRR@5), Hits@5 scores over the MES dataset in transductive (Tran), semi-inductive (Semi), and inductive (Ind) settings. 100% indicates that all the datasets have textual metadata. [75%, 50%, 25%] indicates the percentage of datasets having textual metadata. The best-performing method is indicated in boldface.

MES	Setting	Metric	GAT	SAGE	HAN	HGT	HGNN	MM-SAN
100%	Tran	R@10	0.748	0.715	0.008	0.023	0.656	0.796
		N@10	0.641	0.573	0.009	0.020	0.455	0.566
		P@10	0.020	0.045	0.001	0.001	0.069	0.084
		F1@10	0.085	0.126	0.002	0.003	0.124	0.151
		Hits@10	0.765	0.794	0.010	0.033	0.660	0.803
		MRR@10	0.473	0.469	0.010	0.011	0.392	0.493
	Semi	R@10	0.773	0.712	0.000	0.007	0.655	0.799
		N@10	0.528	0.504	0.000	0.003	0.439	0.551
		P@10	0.074	0.042	0.000	0.000	0.066	0.083
		F1@10	0.111	0.110	0.000	0.000	0.120	0.147
		Hits@10	0.692	0.720	0.000	0.009	0.659	0.811
		MRR@10	0.523	0.485	0.000	0.002	0.370	0.472
	Ind	R@10	0.716	0.738	0.000	0.000	0.658	0.743
		N@10	0.526	0.519	0.000	0.000	0.447	0.538
		P@10	0.045	0.031	0.000	0.000	0.067	0.086
		F1@10	0.127	0.110	0.000	0.000	0.122	0.137
		Hits@10	0.726	0.767	0.000	0.000	0.658	0.743
		MRR@10	0.439	0.437	0.000	0.000	0.379	0.472
75%	Tran	R@10	0.568	0.522	0.008	0.022	0.534	0.704
		N@10	0.454	0.422	0.006	0.011	0.363	0.480
		P@10	0.020	0.019	0.000	0.000	0.057	0.074
		F1@10	0.057	0.096	0.000	0.000	0.103	0.164
		Hits@10	0.565	0.594	0.010	0.033	0.544	0.707
		MRR@10	0.395	0.324	0.001	0.002	0.309	0.408
	Semi	R@10	0.480	0.505	0.000	0.000	0.598	0.723
		N@10	0.442	0.393	0.000	0.000	0.369	0.441
		P@10	0.041	0.023	0.000	0.000	0.061	0.127
		F1@10	0.022	0.022	0.000	0.000	0.110	0.132
		Hits@10	0.577	0.509	0.000	0.000	0.598	0.727
		MRR@10	0.349	0.398	0.000	0.000	0.295	0.405
	Ind	R@10	0.510	0.525	0.000	0.000	0.562	0.701
		N@10	0.434	0.426	0.000	0.000	0.397	0.455
		P@10	0.048	0.007	0.000	0.000	0.057	0.071
		F1@10	0.042	0.088	0.000	0.000	0.104	0.129
		Hits@10	0.596	0.509	0.010	0.020	0.566	0.704
		MRR@10	0.344	0.381	0.000	0.000	0.347	0.383
50%	Tran	R@10	0.312	0.317	0.007	0.015	0.473	0.639
		N@10	0.308	0.234	0.006	0.012	0.329	0.405
		P@10	0.020	0.028	0.000	0.000	0.050	0.067
		F1@10	0.040	0.049	0.000	0.000	0.090	0.120
		Hits@10	0.322	0.265	0.010	0.020	0.476	0.646
		MRR@10	0.303	0.196	0.002	0.001	0.286	0.334
	Semi	R@10	0.320	0.397	0.000	0.000	0.527	0.670
		N@10	0.253	0.304	0.000	0.000	0.333	0.404
		P@10	0.020	0.010	0.000	0.000	0.053	0.067
		F1@10	0.008	0.063	0.000	0.000	0.096	0.122
		Hits@10	0.347	0.382	0.000	0.000	0.530	0.674
		MRR@10	0.293	0.244	0.000	0.000	0.273	0.322
	Ind	R@10	0.363	0.347	0.000	0.000	0.543	0.625
		N@10	0.272	0.236	0.000	0.000	0.367	0.409
		P@10	0.010	0.020	0.000	0.000	0.055	0.063
		F1@10	0.001	0.000	0.000	0.000	0.100	0.115
		Hits@10	0.360	0.352	0.000	0.000	0.546	0.632
		MRR@10	0.238	0.221	0.000	0.000	0.315	0.345
25%	Tran	R@10	0.164	0.159	0.003	0.007	0.376	0.595
		N@10	0.134	0.107	0.002	0.007	0.225	0.371
		P@10	0.010	0.000	0.000	0.000	0.039	0.061
		F1@10	0.036	0.000	0.000	0.000	0.071	0.111
		Hits@10	0.189	0.138	0.003	0.008	0.388	0.599
		MRR@10	0.139	0.089	0.003	0.008	0.181	0.300
	Semi	R@10	0.181	0.206	0.000	0.000	0.466	0.561
		N@10	0.148	0.170	0.000	0.000	0.307	0.344
		P@10	0.020	0.020	0.000	0.000	0.047	0.057
		F1@10	0.002	0.002	0.000	0.000	0.085	0.103
		Hits@10	0.210	0.180	0.000	0.000	0.470	0.561
		MRR@10	0.180	0.107	0.000	0.000	0.259	0.275
	Ind	R@10	0.134	0.207	0.000	0.000	0.451	0.536
		N@10	0.146	0.158	0.000	0.000	0.270	0.322
		P@10	0.020	0.009	0.000	0.000	0.045	0.054
		F1@10	0.040	0.000	0.000	0.000	0.082	0.098
		Hits@10	0.184	0.205	0.000	0.000	0.454	0.539
		MRR@10	0.187	0.124	0.000	0.000	0.216	0.255

a result, the incorporation of diverse node types becomes beneficial, further validating the effectiveness of our sampling approach.

Table 2: Recall@5 (R@5), nDG@5 (N@5), Precision@5 (P@5), f1-score (f1@5), Mean reciprocal rank (MRR@5), Hits@5 scores over the PubMed dataset in transductive (Tran), semi-inductive (Semi), and inductive (Ind) settings. 100% indicates that all the datasets have textual metadata. [75%, 50%, 25%] indicates the percentage of datasets having textual metadata. The best-performing method is indicated in boldface.

PubMed	Setting	Metric	GAT	SAGE	HGT	HAN	HGNN	MM-SAN
100%	Tran	R@5	0.241	0.239	0.016	0.030	0.284	0.380
		N@5	0.181	0.179	0.017	0.023	0.202	0.270
		P@5	0.069	0.076	0.003	0.005	0.058	0.077
		F1@5	0.121	0.121	0.005	0.007	0.096	0.127
		HIT@5	0.300	0.267	0.016	0.030	0.286	0.381
		MRR@5	0.233	0.304	0.019	0.023	0.209	0.234
	Semi	R@5	0.231	0.221	0.006	0.010	0.246	0.374
		N@5	0.171	0.161	0.005	0.008	0.198	0.267
		P@5	0.070	0.075	0.003	0.005	0.055	0.076
		F1@5	0.123	0.125	0.005	0.007	0.091	0.126
		HIT@5	0.254	0.297	0.015	0.022	0.273	0.376
		MRR@5	0.222	0.254	0.009	0.013	0.222	0.232
	Ind	R@5	0.196	0.187	0.000	0.000	0.271	0.344
		N@5	0.134	0.133	0.000	0.000	0.199	0.246
		P@5	0.062	0.069	0.003	0.005	0.050	0.070
		F1@5	0.109	0.115	0.005	0.007	0.083	0.116
		HIT@5	0.289	0.209	0.014	0.021	0.301	0.347
		MRR@5	0.301	0.213	0.009	0.013	0.187	0.214
75%	Tran	R@5	0.201	0.183	0.006	0.012	0.241	0.319
		N@5	0.149	0.139	0.005	0.009	0.171	0.227
		P@5	0.052	0.062	0.003	0.005	0.048	0.064
		F1@5	0.094	0.104	0.004	0.006	0.079	0.107
		HIT@5	0.300	0.298	0.013	0.019	0.252	0.319
		MRR@5	0.196	0.196	0.008	0.012	0.149	0.197
	Semi	R@5	0.186	0.174	0.000	0.007	0.237	0.318
		N@5	0.140	0.133	0.000	0.007	0.174	0.228
		P@5	0.052	0.062	0.003	0.005	0.048	0.065
		F1@5	0.093	0.103	0.004	0.006	0.080	0.107
		HIT@5	0.296	0.202	0.013	0.019	0.289	0.319
		MRR@5	0.198	0.198	0.008	0.012	0.150	0.199
	Ind	R@5	0.145	0.136	0.000	0.000	0.202	0.295
		N@5	0.104	0.097	0.000	0.000	0.199	0.210
		P@5	0.047	0.057	0.002	0.003	0.043	0.060
		F1@5	0.084	0.094	0.003	0.005	0.072	0.099
		HIT@5	0.273	0.199	0.010	0.015	0.215	0.297
		MRR@5	0.182	0.190	0.007	0.011	0.135	0.183
50%	Tran	R@5	0.156	0.145	0.005	0.007	0.203	0.272
		N@5	0.116	0.111	0.005	0.007	0.140	0.192
		P@5	0.034	0.044	0.002	0.003	0.041	0.055
		F1@5	0.062	0.072	0.004	0.006	0.069	0.091
		HIT@5	0.206	0.216	0.011	0.017	0.206	0.274
		MRR@5	0.155	0.163	0.007	0.011	0.131	0.166
	Semi	R@5	0.144	0.129	0.000	0.001	0.201	0.268
		N@5	0.110	0.100	0.000	0.002	0.142	0.190
		P@5	0.031	0.041	0.002	0.003	0.040	0.054
		F1@5	0.059	0.068	0.004	0.006	0.066	0.090
		HIT@5	0.200	0.206	0.011	0.017	0.243	0.270
		MRR@5	0.140	0.150	0.005	0.007	0.124	0.164
	Ind	R@5	0.096	0.057	0.000	0.000	0.129	0.246
		N@5	0.086	0.079	0.000	0.000	0.120	0.140
		P@5	0.010	0.020	0.002	0.003	0.029	0.041
		F1@5	0.024	0.034	0.003	0.005	0.048	0.068
		HIT@5	0.091	0.101	0.008	0.012	0.144	0.204
		MRR@5	0.063	0.073	0.005	0.007	0.087	0.121

Table 3: Recall@10 (R@10), nDG@10 (N@10), Precision@10 (P@10), f1-score (f1@10), Mean reciprocal rank (MRR@10), Hits@10 scores over the MES dataset in transductive (Tran), semi-inductive (Semi), and inductive (Ind) settings. 100% indicates that all the datasets have textual metadata. [75%, 50%, 25%] indicates the percentage of datasets having textual metadata. Methods with the -RW suffix indicate that their original sampling mechanism has been replaced by our proposed sampling strategy. The best-performing method is indicated in boldface.

MES	Setting	Metric	GAT	SAGE	HAN	HGT	HGNN	MM-SAN
100%	Tran	R@10	0.748	0.715	0.000	0.023	0.656	0.796
		N@10	0.641	0.573	0.000	0.020	0.455	0.566
		P@10	0.020	0.045	0.000	0.001	0.069	0.084
		F1@10	0.085	0.126	0.000	0.003	0.124	0.151
		Hits@10	0.765	0.794	0.000	0.033	0.660	0.803
		MRR@10	0.473	0.469	0.000	0.011	0.392	0.493
	Semi	R@10	0.773	0.712	0.000	0.007	0.655	0.799
		N@10	0.528	0.504	0.000	0.003	0.439	0.551
		P@10	0.074	0.042	0.000	0.000	0.066	0.083
		F1@10	0.111	0.110	0.000	0.000	0.120	0.147
		Hits@10	0.692	0.720	0.000	0.009	0.659	0.811
		MRR@10	0.523	0.485	0.000	0.002	0.370	0.472
	Ind	R@10	0.716	0.738	0.000	0.000	0.658	0.743
		N@10	0.526	0.519	0.000	0.000	0.447	0.538
		P@10	0.045	0.031	0.000	0.000	0.067	0.086
		F1@10	0.127	0.110	0.000	0.000	0.122	0.137
		Hits@10	0.726	0.767	0.000	0.000	0.658	0.743
		MRR@10	0.439	0.437	0.000	0.000	0.379	0.472
75%	Tran	R@10	0.568	0.522	0.000	0.022	0.534	0.704
		N@10	0.454	0.422	0.000	0.011	0.363	0.480
		P@10	0.020	0.019	0.000	0.000	0.057	0.074
		F1@10	0.057	0.096	0.000	0.000	0.103	0.164
		Hits@10	0.565	0.594	0.000	0.033	0.544	0.707
		MRR@10	0.395	0.324	0.001	0.002	0.309	0.408
	Semi	R@10	0.480	0.505	0.000	0.000	0.598	0.723
		N@10	0.442	0.393	0.000	0.000	0.369	0.441
		P@10	0.041	0.023	0.000	0.000	0.061	0.127
		F1@10	0.022	0.022	0.000	0.000	0.110	0.132
		Hits@10	0.577	0.509	0.000	0.000	0.598	0.727
		MRR@10	0.349	0.398	0.000	0.000	0.295	0.405
	Ind	R@10	0.510	0.525	0.000	0.000	0.562	0.701
		N@10	0.434	0.426	0.000	0.000	0.397	0.455
		P@10	0.048	0.007	0.000	0.000	0.057	0.071
		F1@10	0.042	0.088	0.000	0.000	0.104	0.129
		Hits@10	0.596	0.509	0.010	0.020	0.566	0.704
		MRR@10	0.344	0.381	0.002	0.000	0.347	0.383
50%	Tran	R@10	0.312	0.317	0.002	0.015	0.473	0.639
		N@10	0.308	0.234	0.006	0.012	0.329	0.405
		P@10	0.020	0.028	0.000	0.003	0.050	0.067
		F1@10	0.040	0.049	0.002	0.006	0.090	0.120
		Hits@10	0.322	0.265	0.010	0.020	0.476	0.646
		MRR@10	0.303	0.196	0.002	0.001	0.286	0.334
	Semi	R@10	0.320	0.397	0.000	0.000	0.527	0.670
		N@10	0.253	0.304	0.000	0.000	0.333	0.404
		P@10	0.020	0.010	0.000	0.000	0.053	0.067
		F1@10	0.008	0.063	0.000	0.000	0.096	0.122
		Hits@10	0.347	0.382	0.010	0.020	0.530	0.674
		MRR@10	0.293	0.244	0.002	0.000	0.273	0.322
	Ind	R@10	0.363	0.347	0.000	0.000	0.543	0.625
		N@10	0.272	0.236	0.000	0.000	0.367	0.409
		P@10	0.010	0.020	0.000	0.000	0.055	0.063
		F1@10	0.001	0.000	0.000	0.000	0.100	0.115
		Hits@10	0.360	0.352	0.010	0.000	0.546	0.632
		MRR@10	0.238	0.221	0.002	0.000	0.315	0.345
25%	Tran	R@10	0.164	0.159	0.000	0.007	0.376	0.595
		N@10	0.134	0.107	0.000	0.008	0.225	0.371
		P@10	0.010	0.000	0.000	0.000	0.039	0.061
		F1@10	0.036	0.000	0.000	0.000	0.071	0.111
		Hits@10	0.189	0.138	0.003	0.008	0.388	0.599
		MRR@10	0.139	0.089	0.003	0.008	0.181	0.300
	Semi	R@10	0.181	0.206	0.000	0.000	0.466	0.561
		N@10	0.148	0.170	0.000	0.000	0.307	0.344
		P@10	0.020	0.020	0.000	0.000	0.047	0.057
		F1@10	0.002	0.002	0.000	0.000	0.085	0.103
		Hits@10	0.210	0.180	0.000	0.000	0.470	0.561
		MRR@10	0.180	0.107	0.000	0.000	0.259	0.275
	Ind	R@10	0.134	0.207	0.000	0.000	0.451	0.536
		N@10	0.146	0.158	0.000	0.000	0.270	0.322
		P@10	0.020	0.009	0.000	0.000	0.045	0.054
		F1@10	0.040	0.000	0.000	0.000	0.082	0.098
		Hits@10	0.184	0.205	0.000	0.000	0.454	0.539
		MRR@10	0.187	0.124	0.000	0.000	0.216	0.255

Table 4: Recall@10 (R@10), nDG@10 (N@10), Precision@10 (P@10), f1-score (f1@10), Mean reciprocal rank (MRR@10), Hits@10 scores over the PubMed dataset in transductive (Tran), semi-inductive (Semi), and inductive (Ind) settings. 100% indicates that all the datasets have textual metadata. [75%, 50%, 25%] indicates the percentage of datasets having textual metadata. Methods with the -RW suffix indicate that their original sampling mechanism has been replaced by our proposed sampling strategy. The best-performing method is indicated in boldface.

MES	Setting	Metric	GAT	SAGE	HAN	HGT	HGNN	MM-SAN
100%	Tran	R@10	0.393	0.435	0.016	0.023	0.342	0.466
		N@10	0.215	0.266	0.017	0.020	0.224	0.298
		P@10	0.046	0.000	0.003	0.001	0.035	0.047
		F1@10	0.047	0.069	0.005	0.003	0.063	0.085
		Hits@10	0.406	0.458	0.016	0.033	0.342	0.467
		MRR@10	0.185	0.235	0.019	0.011	0.187	0.245
	Semi	R@10	0.348	0.349	0.000	0.007	0.335	0.442
		N@10	0.227	0.214	0.000	0.003	0.221	0.289
		P@10	0.000	0.033	0.000	0.000	0.034	0.045
		F1@10	0.038	0.068	0.000	0.000	0.062	0.081
		Hits@10	0.352	0.388	0.000	0.009	0.336	0.444
		MRR@10	0.229	0.219	0.000	0.002	0.185	0.242
	Ind	R@10	0.413	0.392	0.000	0.000	0.302	0.426
		N@10	0.269	0.250	0.000	0.000	0.196	0.273
		P@10	0.000	0.025	0.000	0.000	0.031	0.043
		F1@10	0.051	0.009	0.000	0.000	0.056	0.078
		Hits@10	0.344	0.342	0.000	0.000	0.304	0.428
		MRR@10	0.184	0.235	0.000	0.000	0.163	0.225
75%	Tran	R@10	0.328	0.404	0.008	0.022	0.289	0.406
		N@10	0.186	0.173	0.006	0.011	0.188	0.255
		P@10	0.000	0.000	0.003	0.000	0.029	0.041
		F1@10	0.013	0.000	0.004	0.006	0.053	0.075
		Hits@10	0.352	0.360	0.010	0.009	0.290	0.408
		MRR@10	0.117	0.127	0.008	0.156	0.156	0.209
	Semi	R@10	0.310	0.316	0.000	0.000	0.295	0.391
		N@10	0.234	0.207	0.000	0.000	0.190	0.252
		P@10	0.000	0.031	0.000	0.000	0.030	0.040
		F1@10	0.000	0.067	0.000	0.000	0.054	0.072
		Hits@10	0.330	0.349	0.010	0.000	0.296	0.393
		MRR@10	0.152	0.160	0.008	0.000	0.157	0.209
	Ind	R@10	0.289	0.295	0.000	0.000	0.262	0.366
		N@10	0.157	0.176	0.000	0.000	0.169	0.233
		P@10	0.000	0.016	0.000	0.000	0.027	0.037
		F1@10	0.016	0.031	0.000	0.000	0.041	0.058
		Hits@10	0.351	0.287	0.010	0.000	0.263	0.368
		MRR@10	0.118	0.154	0.007	0.000	0.141	0.192
50%	Tran	R@10	0.266	0.278	0.005	0.015	0.247	0.353
		N@10	0.150	0.191	0.005	0.012	0.158	0.218
		P@10	0.022	0.000	0.002	0.003	0.025	0.036
		F1@10	0.000	0.008	0.004	0.006	0.045	0.065
		Hits@10	0.348	0.336	0.011	0.017	0.248	0.354
		MRR@10	0.163	0.118	0.007	0.011	0.130	0.176
	Semi	R@10	0.281	0.266	0.000	0.000	0.251	0.340
		N@10	0.199	0.204	0.000	0.000	0.164	0.213
		P@10	0.000	0.000	0.000	0.024	0.025	0.034
		F1@10	0.058	0.057	0.000	0.043	0.046	0.062
		Hits@10	0.261	0.266	0.011	0.234	0.253	0.342
		MRR@10	0.130	0.075	0.005	0.168	0.137	0.174
	Ind	R@10	0.281	0.248	0.000	0.000	0.223	0.317
		N@10	0.182	0.136	0.000	0.000	0.139	0.194
		P@10	0.000	0.000	0.000	0.023	0.023	0.032
		F1@10	0.015	0.031	0.000	0.041	0.041	0.058
		Hits@10	0.290	0.228	0.000	0.226	0.224	0.319
		MRR@10	0.064	0.064	0.000	0.146	0.114	0.155
25%	Tran	R@10	0.277	0.236	0.003	0.007	0.193	0.311
		N@10	0.112	0.115	0.002	0.007	0.120	0.184
		P@10	0.015	0.000	0.000	0.000	0.020	0.031
		F1@10	0.015	0.000	0.000	0.000	0.035	0.057
		Hits@10	0.258	0.240	0.003	0.008	0.194	0.312
		MRR@10	0.122	0.111	0.003	0.008	0.097	0.145
	Semi	R@10	0.264	0.228	0.000	0.024	0.187	0.293
		N@10	0.128	0.109	0.000	0.000	0.117	0.174
		P@10	0.000	0.000	0.000	0.000	0.019	0.030
		F1@10	0.000	0.000	0.000	0.022	0.034	0.054
		Hits@10	0.275	0.288	0.000	0.121	0.189	0.295
		MRR@10	0.093	0.105	0.000	0.080	0.095	0.138
	Ind	R@10	0.223	0.234	0.000	0.003	0.182	0.270
		N@10	0.144	0.063	0.000	0.000	0.113	0.162
		P@10	0.000	0.000	0.000	0.011	0.019	0.027
		F1@10	0.028	0.000	0.000	0.000	0.034	0.050
		Hits@10	0.225	0.175	0.008	0.012	0.184	0.272
		MRR@10	0.080	0.102	0.005	0.007	0.093	0.130

Table 5: Recall@20 (R@20), nDG@20 (N@20), Precision@20 (P@20), f1-score (f1@20), Mean reciprocal rank (MRR@20), Hits@20 scores over the MES dataset in transductive (Tran), semi-inductive (Semi), and inductive (Ind) settings. 100% indicates that all the datasets have textual metadata. [75%, 50%, 25%] indicates the percentage of datasets having textual metadata. Methods with the -RW suffix indicate that their original sampling mechanism has been replaced by our proposed sampling strategy. The best-performing method is indicated in boldface.

MES	Setting	Metric	GAT	SAGE	HAN	HGT	HGNN	MM-SAN
100%	Tran	R@20	0.769	0.747	0.016	0.023	0.707	0.867
		N@20	0.563	0.570	0.017	0.020	0.468	0.584
		P@20	0.039	0.041	0.003	0.001	0.037	0.046
		F1@20	0.031	0.043	0.005	0.003	0.070	0.086
		Hits@20	0.811	0.824	0.016	0.033	0.714	0.871
		MRR@20	0.486	0.491	0.019	0.011	0.395	0.498
	Semi	R@20	0.730	0.768	0.000	0.007	0.723	0.886
		N@20	0.568	0.542	0.000	0.003	0.457	0.573
		P@20	0.004	0.000	0.000	0.000	0.037	0.045
		F1@20	0.080	0.077	0.000	0.000	0.070	0.086
		Hits@20	0.735	0.777	0.000	0.009	0.727	0.886
		MRR@20	0.476	0.476	0.000	0.002	0.375	0.477
	Ind	R@20	0.738	0.778	0.000	0.000	0.711	0.809
		N@20	0.554	0.554	0.000	0.000	0.460	0.555
		P@20	0.030	0.013	0.000	0.000	0.036	0.041
		F1@20	0.033	0.009	0.000	0.000	0.069	0.078
		Hits@20	0.758	0.766	0.000	0.000	0.711	0.809
		MRR@20	0.436	0.472	0.000	0.000	0.383	0.477
75%	Tran	R@20	0.558	0.540	0.008	0.022	0.609	0.765
		N@20	0.444	0.449	0.006	0.011	0.381	0.495
		P@20	0.000	0.000	0.003	0.000	0.032	0.040
		F1@20	0.000	0.000	0.004	0.006	0.061	0.076
		Hits@20	0.605	0.532	0.010	0.009	0.619	0.769
		MRR@20	0.411	0.400	0.008	0.003	0.314	0.412
	Semi	R@20	0.569	0.556	0.000	0.000	0.674	0.818
		N@20	0.410	0.393	0.000	0.000	0.388	0.466
		P@20	0.022	0.000	0.000	0.000	0.034	0.041
		F1@20	0.021	0.046	0.000	0.000	0.065	0.079
		Hits@20	0.516	0.542	0.000	0.000	0.674	0.826
		MRR@20	0.360	0.327	0.000	0.000	0.300	0.361
	Ind	R@20	0.613	0.617	0.000	0.000	0.658	0.799
		N@20	0.479	0.436	0.000	0.000	0.422	0.480
		P@20	0.000	0.000	0.000	0.000	0.034	0.040
		F1@20	0.041	0.016	0.000	0.000	0.064	0.077
		Hits@20	0.594	0.568	0.000	0.000	0.658	0.803
		MRR@20	0.388	0.383	0.000	0.000	0.354	0.477
50%	Tran	R@20	0.313	0.337	0.005	0.015	0.561	0.779
		N@20	0.263	0.217	0.005	0.012	0.351	0.440
		P@20	0.000	0.000	0.000	0.000	0.030	0.041
		F1@20	0.027	0.011	0.004	0.006	0.056	0.077
		Hits@20	0.340	0.345	0.011	0.017	0.565	0.789
		MRR@20	0.230	0.201	0.007	0.011	0.292	0.344
	Semi	R@20	0.349	0.382	0.000	0.024	0.629	0.758
		N@20	0.291	0.333	0.000	0.000	0.360	0.427
		P@20	0.014	0.000	0.000	0.000	0.032	0.039
		F1@20	0.021	0.041	0.000	0.000	0.061	0.073
		Hits@20	0.357	0.333	0.011	0.000	0.629	0.758
		MRR@20	0.218	0.212	0.005	0.000	0.280	0.328
	Ind	R@20	0.321	0.290	0.000	0.000	0.582	0.691
		N@20	0.270	0.213	0.000	0.000	0.376	0.426
		P@20	0.000	0.000	0.000	0.000	0.026	0.035
		F1@20	0.001	0.012	0.000	0.000	0.049	0.062
		Hits@20	0.321	0.169	0.000	0.000	0.513	0.651
		MRR@20	0.199	0.169	0.000	0.000	0.317	0.349
25%	Tran	R@20	0.113	0.150	0.000	0.000	0.460	0.667
		N@20	0.132	0.066	0.000	0.000	0.247	0.389
		P@20	0.033	0.000	0.000	0.000	0.024	0.051
		F1@20	0.040	0.000	0.000	0.000	0.046	0.065
		Hits@20	0.139	0.126	0.003	0.008	0.563	0.667
		MRR@20	0.129	0.064	0.003	0.008	0.187	0.305
	Semi	R@20	0.160	0.196	0.000	0.000	0.530	0.705
		N@20	0.112	0.121	0.000	0.000	0.324	0.381
		P@20	0.022	0.008	0.000	0.000	0.027	0.036
		F1@20	0.040	0.000	0.000	0.026	0.051	0.068
		Hits@20	0.245	0.176	0.000	0.000	0.530	0.705
		MRR@20	0.125	0.141	0.000	0.000	0.264	0.285
	Ind	R@20	0.217	0.134	0.000	0.000	0.513	0.645
		N@20	0.147	0.129	0.000	0.000	0.286	0.350
		P@20	0.010	0.010	0.000	0.000	0.026	0.033
		F1@20	0.015	0.012	0.000	0.000	0.049	0.062
		Hits@20	0.184	0.169	0.000	0.000	0.513	0.651
		MRR@20	0.118	0.087	0.000	0.000	0.220	0.263

Table 6: Recall@20 (R@20), nDG@20 (N@20), Precision@20 (P@20), f1-score (f1@20), Mean reciprocal rank (MRR@20), Hits@20 scores over the PubMed dataset in transductive (Tran), semi-inductive (Semi), and inductive (Ind) settings. 100% indicates that all the datasets have textual metadata. [75%, 50%, 25%] indicates the percentage of datasets having textual metadata. Methods with the -RW suffix indicate that their original sampling mechanism has been replaced by our proposed sampling strategy. The best-performing method is indicated in boldface.

PubMed	Setting	Metric	GAT	SAGE	HAN	HGT	HGNN	MM-SAN
100%	Tran	R@20	0.513	0.505	0.016	0.023	0.396	0.528
		N@20	0.303	0.300	0.017	0.020	0.238	0.313
		P@20	0.011	0.000	0.003	0.001	0.020	0.027
		F1@20	0.032	0.016	0.005	0.003	0.038	0.077
		Hits@20	0.596	0.556	0.016	0.033	0.398	0.529
		MRR@20	0.248	0.242	0.019	0.011	0.191	0.290
	Semi	R@20	0.506	0.484	0.000	0.007	0.388	0.511
		N@20	0.264	0.264	0.000	0.003	0.234	0.307
		P@20	0.019	0.000	0.000	0.000	0.020	0.026
		F1@20	0.022	0.001	0.000	0.000	0.037	0.049
		Hits@20	0.521	0.497	0.000	0.009	0.390	0.513
		MRR@20	0.226	0.216	0.000	0.000	0.189	0.247
	Ind	R@20	0.439	0.416	0.000	0.000	0.352	0.496
		N@20	0.280	0.267	0.000	0.308	0.209	0.291
		P@20	0.014	0.022	0.000	0.000	0.018	0.025
		F1@20	0.043	0.039	0.000	0.000	0.034	0.048
		Hits@20	0.453	0.480	0.000	0.000	0.394	0.499
		MRR@20	0.235	0.235	0.000	0.259	0.167	0.230
75%	Tran	R@20	0.312	0.301	0.000	0.000	0.342	0.486
		N@20	0.260	0.182	0.000	0.000	0.201	0.276
		P@20	0.021	0.000	0.000	0.000	0.017	0.025
		F1@20	0.023	0.002	0.000	0.000	0.033	0.047
		Hits@20	0.401	0.387	0.000	0.000	0.343	0.488
		MRR@20	0.149	0.140	0.000	0.000	0.160	0.214
	Semi	R@20	0.327	0.286	0.000	0.000	0.335	0.462
		N@20	0.228	0.213	0.000	0.000	0.200	0.290
		P@20	0.008	0.000	0.000	0.000	0.017	0.023
		F1@20	0.041	0.010	0.000	0.000	0.032	0.045
		Hits@20	0.329	0.284	0.000	0.000	0.335	0.463
		MRR@20	0.205	0.205	0.000	0.000	0.160	0.234
	Ind	R@20	0.320	0.267	0.000	0.000	0.306	0.427
		N@20	0.225	0.194	0.000	0.000	0.180	0.249
		P@20	0.002	0.000	0.000	0.000	0.016	0.025
		F1@20	0.017	0.000	0.000	0.000	0.030	0.037
		Hits@20	0.410	0.328	0.000	0.000	0.309	0.429
		MRR@20	0.136	0.111	0.000	0.000	0.144	0.197
50%	Tran	R@20	0.290	0.170	0.000	0.000	0.305	0.424
		N@20	0.153	0.109	0.000	0.000	0.173	0.236
		P@20	0.016	0.011	0.000	0.000	0.015	0.051
		F1@20	0.018	0.012	0.000	0.000	0.029	0.051
		Hits@20	0.272	0.195	0.000	0.000	0.306	0.426
		MRR@20	0.097	0.062	0.000	0.000	0.134	0.182
	Semi	R@20	0.367	0.168	0.000	0.000	0.294	0.415
		N@20	0.210	0.161	0.000	0.000	0.175	0.232
		P@20	0.002	0.000	0.000	0.000	0.015	0.021
		F1@20	0.031	0.000	0.000	0.000	0.028	0.040
		Hits@20	0.413	0.213	0.000	0.000	0.296	0.418
		MRR@20	0.096	0.095	0.000	0.000	0.140	0.179
	Ind	R@20	0.281	0.107	0.000	0.000	0.269	0.382
		N@20	0.113	0.094	0.000	0.000	0.151	0.210
		P@20	0.008	0.000	0.000	0.000	0.014	0.019
		F1@20	0.031	0.000	0.000	0.000	0.026	0.037
		Hits@20	0.186	0.154	0.000	0.000	0.270	0.385
		MRR@20	0.159	0.136	0.000	0.000	0.117	0.160
25%	Tran	R@20	0.202	0.101	0.000	0.000	0.236	0.382
		N@20	0.192	0.061	0.000	0.000	0.131	0.202
		P@20	0.001	0.000	0.000	0.000	0.012	0.019
		F1@20	0.023	0.000	0.000	0.000	0.023	0.037
		Hits@20	0.144	0.037	0.003	0.008	0.237	0.383
		MRR@20	0.051	0.000	0.003	0.008	0.100	0.150
	Semi	R@20	0.293	0.049	0.000	0.000	0.226	0.370
		N@20	0.145	0.067	0.000	0.000	0.126	0.194
		P@20	0.004	0.000	0.000	0.000	0.011	0.019
		F1@20	0.023	0.004	0.000	0.000	0.022	0.036
		Hits@20	0.242	0.097	0.000	0.000	0.228	0.372
		MRR@20	0.070	0.040	0.000	0.000	0.098	0.144
	Ind	R@20	0.120	0.107	0.000	0.000	0.224	0.345
		N@20	0.119	0.004	0.000	0.000	0.124	0.181
		P@20	0.003	0.000	0.000	0.000	0.011	0.017
		F1@20	0.031	0.000	0.000	0.000	0.022	0.033
		Hits@20	0.342	0.029	0.000	0.000	0.226	0.346
		MRR@20	0.010	0.000	0.000	0.000	0.095	0.135

Table 7: We report the performances of SAGE, GAT, HGNN, their enhanced versions with our sampling strategy, and MM-SAN. We report recall@5 (R@5), nDG@5 (N@5), Precision@5 (P@5), f1-score (f1@5), Mean reciprocal rank (MRR@5), Hits@5 scores over the MES dataset in transductive (Tran), semi-inductive (Semi), and inductive (Ind) settings. 100% indicates that all the datasets have textual metadata. [75%, 50%, 25%] indicates the percentage of datasets having textual metadata. The columns Gain indicate the percentage gain achieved from the original implementation and its enhanced version with our sampling strategy. In boldface we report the highest performances for each setting and metric.

MES	Setting	Metric	GAT	GAT-RW	Gain (GAT)	SAGE	SAGE-RW	Gain (SAGE)	HGNN	HGNN-RW	Gain (HGNN)	MMSAN
0%	Tran	R@5	0.625	0.700	12.000	0.489	0.610	24.700	0.581	0.700	20.500	0.727
		N@5	0.475	0.520	9.500	0.398	0.510	28.100	0.438	0.520	18.800	0.546
		P@5	0.145	0.150	3.400	0.142	0.148	4.200	0.120	0.145	20.800	0.152
		F1@5	0.240	0.246	2.500	0.236	0.244	3.400	0.198	0.240	21.200	0.250
		Hits@5	0.711	0.721	1.400	0.719	0.735	2.200	0.580	0.725	25.100	0.741
		MRR@5	0.482	0.490	1.600	0.448	0.470	4.900	0.378	0.480	26.900	0.484
	Semi	R@5	0.621	0.700	12.700	0.446	0.600	34.400	0.587	0.700	19.300	0.719
		N@5	0.479	0.490	2.300	0.316	0.400	37.800	0.417	0.510	22.300	0.524
		P@5	0.141	0.145	2.800	0.155	0.160	3.200	0.118	0.138	16.900	0.145
		F1@5	0.242	0.241	0.400	0.236	0.238	0.800	0.197	0.230	16.700	0.242
		Hits@5	0.723	0.721	0.300	0.716	0.710	0.900	0.591	0.720	21.900	0.727
		MRR@5	0.454	0.456	0.400	0.447	0.455	1.800	0.362	0.450	24.400	0.461
	Ind	R@5	0.590	0.622	5.400	0.439	0.512	16.700	0.582	0.634	8.900	0.687
		N@5	0.442	0.513	16.000	0.315	0.346	18.700	0.422	0.475	12.600	0.519
		P@5	0.123	0.136	10.700	0.133	0.145	9.200	0.118	0.120	1.700	0.149
		F1@5	0.226	0.232	2.600	0.225	0.230	2.200	0.196	0.202	3.100	0.232
		Hits@5	0.611	0.627	2.600	0.607	0.631	3.900	0.586	0.644	9.900	0.691
		MRR@5	0.434	0.437	0.700	0.437	0.438	0.200	0.370	0.377	1.800	0.466
75%	Tran	R@5	0.481	0.511	6.200	0.352	0.372	5.700	0.573	0.593	3.500	0.623
		N@5	0.370	0.391	5.700	0.294	0.314	6.900	0.356	0.387	8.600	0.447
		P@5	0.118	0.132	11.900	0.119	0.123	3.400	0.101	0.103	2.100	0.128
		F1@5	0.198	0.204	3.000	0.193	0.202	4.700	0.164	0.172	4.900	0.210
		Hits@5	0.484	0.518	7.000	0.472	0.478	1.300	0.556	0.598	7.600	0.659
		MRR@5	0.491	0.499	1.600	0.451	0.493	9.300	0.300	0.309	3.000	0.495
	Semi	R@5	0.483	0.501	3.700	0.332	0.393	18.400	0.515	0.545	5.800	0.598
		N@5	0.404	0.434	7.400	0.247	0.306	23.800	0.373	0.387	3.800	0.442
		P@5	0.101	0.109	7.900	0.106	0.111	4.700	0.103	0.102	1.000	0.106
		F1@5	0.173	0.181	4.600	0.180	0.184	2.300	0.172	0.181	5.300	0.187
		Hits@5	0.516	0.525	1.700	0.449	0.503	11.900	0.515	0.552	7.200	0.630
		MRR@5	0.393	0.398	1.300	0.377	0.391	3.700	0.283	0.292	3.200	0.398
	Ind	R@5	0.409	0.432	5.600	0.325	0.354	9.100	0.523	0.549	5.000	0.598
		N@5	0.324	0.338	4.300	0.235	0.241	2.400	0.383	0.385	0.500	0.468
		P@5	0.111	0.114	2.700	0.110	0.116	5.500	0.096	0.098	1.700	0.116
		F1@5	0.183	0.191	4.400	0.186	0.192	3.200	0.159	0.163	2.500	0.192
		Hits@5	0.458	0.476	3.900	0.463	0.502	8.400	0.580	0.599	3.300	0.609
		MRR@5	0.443	0.449	1.300	0.383	0.417	8.900	0.336	0.355	5.600	0.368
50%	Tran	R@5	0.361	0.399	10.600	0.286	0.309	8.100	0.443	0.453	2.300	0.539
		N@5	0.319	0.337	5.600	0.212	0.260	22.600	0.324	0.367	13.300	0.398
		P@5	0.069	0.071	2.900	0.066	0.072	9.100	0.083	0.093	11.900	0.102
		F1@5	0.111	0.116	4.500	0.117	0.118	0.800	0.136	0.139	2.200	0.168
		Hits@5	0.337	0.360	6.800	0.400	0.401	0.300	0.451	0.460	2.000	0.563
		MRR@5	0.311	0.337	8.400	0.282	0.292	3.600	0.276	0.292	5.800	0.340
	Semi	R@5	0.349	0.380	8.900	0.240	0.263	9.600	0.484	0.500	3.300	0.518
		N@5	0.277	0.302	9.100	0.209	0.256	22.400	0.352	0.371	5.400	0.359
		P@5	0.069	0.076	10.100	0.075	0.083	10.800	0.077	0.081	5.100	0.108
		F1@5	0.123	0.126	2.400	0.130	0.139	6.900	0.129	0.130	0.800	0.179
		Hits@5	0.372	0.389	4.600	0.414	0.417	0.700	0.486	0.502	3.300	0.538
		MRR@5	0.302	0.303	0.300	0.295	0.296	0.300	0.253	0.257	1.600	0.303
	Ind	R@5	0.321	0.376	17.100	0.196	0.211	7.600	0.434	0.457	5.300	0.526
		N@5	0.258	0.276	7.000	0.150	0.177	18.100	0.339	0.357	5.300	0.391
		P@5	0.063	0.067	6.300	0.059	0.066	12.000	0.092	0.100	7.400	0.099
		F1@5	0.108	0.111	2.800	0.100	0.109	9.000	0.153	0.157	2.500	0.164
		Hits@5	0.333	0.386	15.900	0.380	0.399	5.000	0.461	0.471	2.100	0.558
		MRR@5	0.273	0.281	2.900	0.243	0.281	15.700	0.304	0.306	0.700	0.326
25%	Tran	R@5	0.230	0.278	21.000	0.190	0.201	5.800	0.336	0.357	6.300	0.467
		N@5	0.203	0.203	0.000	0.165	0.165	0.000	0.242	0.242	0.000	0.329
		P@5	0.032	0.039	21.900	0.030	0.038	26.700	0.060	0.064	6.600	0.095
		F1@5	0.063	0.064	1.600	0.060	0.062	3.300	0.098	0.100	2.000	0.157
		Hits@5	0.245	0.281	14.700	0.331	0.340	2.800	0.340	0.375	10.100	0.473
		MRR@5	0.250	0.260	4.000	0.200	0.232	16.000	0.169	0.173	2.400	0.283
	Semi	R@5	0.206	0.282	37.100	0.145	0.190	31.000	0.367	0.401	9.200	0.454
		N@5	0.172	0.191	11.100	0.125	0.189	51.100	0.270	0.301	11.400	0.306
		P@5	0.042	0.047	11.900	0.045	0.047	4.400	0.074	0.079	6.800	0.086
		F1@5	0.070	0.078	11.400	0.072	0.078	8.600	0.123	0.129	4.900	0.144
		Hits@5	0.235	0.330	40.300	0.292	0.320	9.400	0.371	0.410	10.600	0.462
		MRR@5	0.177	0.181	2.300	0.186	0.198	6.500	0.247	0.256	4.100	0.256
	Ind	R@5	0.159	0.176	10.700	0.075	0.075	0.000	0.322	0.367	13.900	0.453
		N@5	0.128	0.128	0.000	0.069	0.069	0.000	0.233	0.233	0.000	0.321
		P@5	0.035	0.042	20.000	0.035	0.043	22.900	0.058	0.063	8.600	0.088
		F1@5	0.060	0.070	16.700	0.063	0.072	14.300	0.096	0.103	7.300	0.147
		Hits@5	0.203	0.211	3.900	0.214	0.217	1.400	0.340	0.378	11.100	0.441
		MRR@5	0.186	0.190	2.100	0.174	0.188	8.100	0.194	0.201	3.500	0.241

Table 8: We report the performances of SAGE, GAT, HGNN, their enhanced versions with our sampling strategy, and MM-SAN. We report recall@5 (R@5), nDG@5 (N@5), Precision@5 (P@5), f1-score (f1@5), Mean reciprocal rank (MRR@5), Hits@5 scores over the PubMed dataset in transductive (Tran), semi-inductive (Semi), and inductive (Ind) settings. 100% indicates that all the datasets have textual metadata. [75%, 50%, 25%] indicates the percentage of datasets having textual metadata. The columns Gain indicate the percentage gain achieved from the original implementation and its enhanced version with our sampling strategy. In boldface we report the highest performances for each setting and metric.

PubMed	Setting	Metric	GAT	GAT-RW	Gain (GAT)	SAGE	SAGE-RW	Gain (SAGE)	HGNN	HGNN-RW	Gain (HGNN)	MM-SAN
100%	Tran	R@5	0.241	0.298	23.654	0.239	0.278	16.412	0.284	0.301	5.963	0.380
		N@5	0.181	0.202	11.629	0.179	0.201	12.345	0.202	0.256	26.731	0.270
		P@5	0.069	0.076	10.145	0.076	0.076	0.000	0.058	0.070	20.690	0.077
		F1@5	0.121	0.125	3.289	0.121	0.126	4.090	0.096	0.106	10.417	0.127
		HIT@5	0.300	0.310	3.333	0.267	0.301	12.727	0.286	0.309	8.036	0.381
		MRR@5	0.233	0.254	9.024	0.304	0.369	21.228	0.209	0.255	22.021	0.234
	Semi	R@5	0.231	0.242	4.762	0.221	0.253	14.457	0.246	0.271	10.204	0.374
		N@5	0.171	0.205	19.886	0.161	0.182	13.211	0.198	0.201	1.523	0.267
		P@5	0.070	0.075	7.143	0.075	0.075	0.000	0.055	0.072	30.000	0.076
		F1@5	0.123	0.125	1.626	0.125	0.125	0.000	0.091	0.125	37.360	0.126
		HIT@5	0.254	0.301	18.627	0.297	0.311	4.716	0.273	0.309	13.213	0.376
		MRR@5	0.222	0.245	10.351	0.254	0.302	18.633	0.222	0.243	9.423	0.232
	Ind	R@5	0.196	0.240	22.448	0.187	0.239	27.747	0.271	0.320	18.073	0.344
		N@5	0.134	0.186	38.806	0.133	0.201	51.127	0.199	0.267	34.151	0.246
		P@5	0.062	0.069	11.290	0.069	0.069	0.000	0.050	0.069	37.093	0.070
		F1@5	0.109	0.115	5.514	0.115	0.115	0.000	0.083	0.115	5.221	0.116
		HIT@5	0.289	0.299	3.448	0.209	0.265	26.629	0.301	0.321	6.668	0.347
		MRR@5	0.301	0.319	5.985	0.213	0.274	28.662	0.187	0.221	18.216	0.214
75%	Tran	R@5	0.201	0.240	19.404	0.183	0.198	8.207	0.241	0.267	10.771	0.319
		N@5	0.149	0.167	12.083	0.139	0.145	4.337	0.171	0.193	12.831	0.227
		P@5	0.052	0.063	21.154	0.062	0.063	0.000	0.048	0.063	31.250	0.064
		F1@5	0.094	0.106	12.765	0.104	0.106	1.923	0.079	0.106	34.177	0.107
		HIT@5	0.300	0.318	6.000	0.298	0.310	4.033	0.252	0.278	10.267	0.319
		MRR@5	0.196	0.196	0.000	0.196	0.196	0.000	0.149	0.170	13.781	0.197
	Semi	R@5	0.186	0.201	8.064	0.174	0.199	14.377	0.237	0.258	8.854	0.318
		N@5	0.140	0.167	19.286	0.133	0.167	25.379	0.174	0.192	10.345	0.228
		P@5	0.052	0.064	23.077	0.062	0.064	3.226	0.048	0.064	33.333	0.065
		F1@5	0.093	0.106	13.978	0.103	0.106	2.921	0.080	0.100	25.000	0.107
		HIT@5	0.296	0.318	7.432	0.202	0.265	31.619	0.289	0.295	2.083	0.319
		MRR@5	0.198	0.198	0.000	0.198	0.198	0.000	0.150	0.184	22.667	0.199
	Ind	R@5	0.145	0.175	20.690	0.136	0.190	39.706	0.202	0.232	14.851	0.295
		N@5	0.104	0.165	58.654	0.097	0.102	5.154	0.145	0.188	29.828	0.210
		P@5	0.047	0.059	25.532	0.057	0.059	3.509	0.050	0.059	18.000	0.060
		F1@5	0.084	0.098	16.667	0.094	0.098	4.255	0.072	0.098	26.389	0.099
		HIT@5	0.273	0.288	5.510	0.199	0.229	15.083	0.215	0.227	5.581	0.297
		MRR@5	0.132	0.145	9.848	0.142	0.155	9.157	0.135	0.150	11.111	0.183
50%	Tran	R@5	0.156	0.198	27.027	0.145	0.168	15.862	0.203	0.213	4.933	0.272
		N@5	0.116	0.124	6.897	0.111	0.139	25.232	0.140	0.166	18.571	0.192
		P@5	0.034	0.053	55.882	0.044	0.053	20.455	0.041	0.053	29.268	0.055
		F1@5	0.062	0.075	21.000	0.072	0.083	15.277	0.069	0.089	28.986	0.091
		HIT@5	0.206	0.248	20.370	0.216	0.229	6.018	0.206	0.235	14.084	0.274
		MRR@5	0.155	0.163	5.161	0.163	0.163	0.000	0.131	0.159	21.427	0.166
	Semi	R@5	0.144	0.165	14.583	0.129	0.132	2.325	0.201	0.210	4.474	0.268
		N@5	0.110	0.122	10.909	0.100	0.130	30.000	0.142	0.139	0.000	0.190
		P@5	0.031	0.054	74.194	0.041	0.054	31.707	0.040	0.054	35.000	0.054
		F1@5	0.059	0.071	20.340	0.069	0.090	30.435	0.066	0.090	28.904	0.090
		HIT@5	0.200	0.224	12.000	0.206	0.249	21.008	0.243	0.273	12.343	0.270
		MRR@5	0.140	0.165	17.857	0.150	0.175	16.667	0.124	0.165	33.064	0.164
	Ind	R@5	0.096	0.132	37.500	0.097	0.122	25.773	0.167	0.199	19.212	0.246
		N@5	0.086	0.101	17.441	0.079	0.098	24.051	0.120	0.140	16.667	0.170
		P@5	0.030	0.049	63.333	0.040	0.049	22.500	0.035	0.049	40.000	0.050
		F1@5	0.056	0.070	25.000	0.066	0.082	24.242	0.058	0.082	41.379	0.083
		HIT@5	0.189	0.213	12.672	0.199	0.229	15.076	0.215	0.227	5.887	0.247
		MRR@5	0.132	0.145	9.848	0.142	0.155	9.157	0.135	0.150	11.111	0.183
25%	Tran	R@5	0.113	0.167	47.787	0.108	0.119	10.185	0.156	0.199	27.500	0.234
		N@5	0.084	0.100	19.048	0.089	0.099	11.236	0.108	0.123	13.888	0.159
		P@5	0.010	0.021	110.000	0.020	0.046	130.000	0.030	0.046	53.846	0.047
		F1@5	0.024	0.063	162.500	0.034	0.060	76.471	0.049	0.077	60.870	0.078
		HIT@5	0.120	0.132	10.000	0.110	0.129	17.273	0.140	0.153	9.286	0.235
		MRR@5	0.063	0.099	57.143	0.073	0.091	24.658	0.091	0.102	12.155	0.135
	Semi	R@5	0.098	0.109	11.224	0.096	0.102	6.250	0.152	0.167	10.666	0.211
		N@5	0.077	0.091	18.181	0.075	0.089	18.667	0.107	0.111	3.738	0.147
		P@5	0.011	0.026	136.364	0.021	0.042	100.000	0.029	0.042	44.828	0.043
		F1@5	0.024	0.067	179.167	0.034	0.060	76.471	0.048	0.070	66.667	0.071
		HIT@5	0.178	0.199	11.801	0.155	0.178	14.838	0.205	0.230	12.195	0.213
		MRR@5	0.101	0.107	5.940	0.078	0.105	34.615	0.090	0.114	26.079	0.127
	Ind	R@5	0.053	0.089	68.868	0.057	0.077	35.088	0.129	0.134	3.883	0.202
		N@5	0.038	0.088	131.579	0.043	0.063	46.512	0.090	0.100	10.526	0.140
		P@5	0.010	0.022	120.000	0.020	0.034	70.000	0.029	0.040	37.931	0.041
		F1@5	0.024	0.063	162.500	0.034	0.045	70.588	0.048	0.067	45.833	0.068
		HIT@5	0.091	0.140	53.846	0.101	0.137	35.643	0.144	0.189	31.944	0.204
		MRR@5	0.063	0.112	77.778	0.073	0.118	61.643	0.087	0.117	34.483	0.121

Table 9: We report the performances of SAGE, GAT, HGNN, their enhanced versions with our sampling strategy, and MM-SAN. We report recall@10 (R@10), nDG@10 (N@10), Precision@10 (P@10), f1-score (f1@10), Mean reciprocal rank (MRR@10), Hits@10 scores over the MES dataset in transductive (Tran), semi-inductive (Semi), and inductive (Ind) settings. 100% indicates that all the datasets have textual metadata. [75%, 50%, 25%] indicates the percentage of datasets having textual metadata. The columns Gain indicate the percentage gain achieved from the original implementation and its enhanced version with our sampling strategy. In boldface we report the highest performances for each setting and metric.

MES	Setting	Metric	GAT	GAT-RW	Gain (GAT)	SAGE	SAGE-RW	Gain (SAGE)	HGNN	HGNN-RW	Gain (HGNN)	MM-SAN
100%	Tran	R@10	0.748	0.786	5.089	0.715	0.778	8.817	0.656	0.670	2.141	0.796
		N@10	0.641	0.666	3.907	0.573	0.632	10.290	0.455	0.466	2.415	0.566
		P@10	0.040	0.065	62.000	0.045	0.085	88.889	0.069	0.081	17.391	0.084
		F1@10	0.085	0.153	79.531	0.126	0.153	21.428	0.124	0.127	2.424	0.151
		Hits@10	0.765	0.789	3.122	0.794	0.800	0.756	0.660	0.683	3.439	0.803
		MRR@10	0.473	0.523	10.573	0.469	0.478	1.910	0.392	0.484	23.470	0.493
	Semi	R@10	0.773	0.785	1.552	0.712	0.778	9.259	0.655	0.685	4.854	0.799
		N@10	0.528	0.546	3.409	0.504	0.536	6.349	0.439	0.479	9.115	0.551
		P@10	0.074	0.082	10.811	0.042	0.083	97.619	0.066	0.071	7.578	0.081
		F1@10	0.111	0.139	25.314	0.110	0.141	28.182	0.120	0.135	12.786	0.147
		Hits@10	0.692	0.711	2.747	0.720	0.800	11.111	0.659	0.668	1.364	0.811
		MRR@10	0.523	0.567	8.421	0.485	0.564	16.331	0.370	0.383	3.513	0.472
	Ind	R@10	0.716	0.733	2.376	0.738	0.740	0.270	0.658	0.699	6.232	0.743
		N@10	0.526	0.533	1.333	0.519	0.530	2.114	0.447	0.529	18.338	0.538
		P@10	0.045	0.080	77.778	0.061	0.078	27.800	0.067	0.081	20.895	0.086
		F1@10	0.124	0.127	2.362	0.110	0.132	20.000	0.122	0.130	6.557	0.137
		Hits@10	0.726	0.783	7.841	0.767	0.770	0.391	0.658	0.697	5.923	0.743
		MRR@10	0.439	0.445	1.364	0.437	0.446	2.062	0.379	0.441	16.353	0.472
75%	Tran	R@10	0.568	0.605	6.514	0.522	0.616	17.929	0.534	0.583	9.168	0.704
		N@10	0.454	0.474	4.409	0.422	0.473	11.151	0.363	0.442	21.778	0.480
		P@10	0.050	0.065	30.000	0.049	0.065	32.105	0.057	0.070	22.808	0.074
		F1@10	0.057	0.076	37.509	0.096	0.117	22.917	0.103	0.161	56.341	0.164
		Hits@10	0.565	0.612	8.319	0.594	0.619	4.211	0.544	0.617	13.416	0.707
		MRR@10	0.395	0.401	1.518	0.324	0.369	13.580	0.309	0.366	18.452	0.408
	Semi	R@10	0.480	0.580	20.833	0.505	0.580	14.850	0.598	0.631	5.504	0.723
		N@10	0.440	0.442	0.452	0.393	0.448	13.994	0.369	0.433	17.300	0.441
		P@10	0.041	0.059	43.902	0.043	0.059	37.143	0.061	0.127	107.377	0.073
		F1@10	0.022	0.037	68.364	0.022	0.024	9.090	0.110	0.123	11.818	0.132
		Hits@10	0.577	0.583	1.042	0.509	0.583	14.541	0.598	0.620	3.670	0.727
		MRR@10	0.349	0.403	15.468	0.398	0.406	2.016	0.295	0.381	29.103	0.405
	Ind	R@10	0.510	0.609	19.412	0.525	0.602	14.714	0.562	0.583	3.733	0.701
		N@10	0.434	0.451	3.913	0.426	0.455	6.796	0.397	0.444	11.837	0.455
		P@10	0.048	0.062	29.167	0.007	0.008	14.429	0.057	0.070	22.807	0.071
		F1@10	0.042	0.112	166.667	0.088	0.111	26.136	0.104	0.120	15.384	0.129
		Hits@10	0.596	0.612	2.690	0.509	0.605	18.843	0.566	0.611	7.954	0.704
		MRR@10	0.344	0.355	3.196	0.381	0.421	10.512	0.347	0.417	20.222	0.383
25%	Tran	R@10	0.164	0.195	19.018	0.159	0.188	18.739	0.376	0.458	21.847	0.595
		N@10	0.134	0.158	17.910	0.107	0.152	41.682	0.225	0.308	36.944	0.371
		P@10	0.010	0.011	10.000	0.015	0.020	25.000	0.039	0.54	37.510	0.061
		F1@10	0.036	0.038	5.556	0.000	0.037	0.000	0.071	0.125	76.923	0.111
		Hits@10	0.189	0.204	7.946	0.138	0.197	42.393	0.388	0.480	23.759	0.599
		MRR@10	0.139	0.153	10.145	0.089	0.147	65.167	0.181	0.187	3.315	0.300
	Semi	R@10	0.181	0.250	38.078	0.206	0.250	21.847	0.466	0.489	4.925	0.561
		N@10	0.148	0.200	35.135	0.170	0.205	20.588	0.307	0.402	31.001	0.344
		P@10	0.020	0.026	30.000	0.020	0.026	30.000	0.047	0.141	33.333	0.057
		F1@10	0.002	0.003	50.000	0.002	0.004	100.000	0.085	0.092	8.353	0.103
		Hits@10	0.210	0.250	19.048	0.180	0.250	38.889	0.470	0.489	3.809	0.561
		MRR@10	0.180	0.183	1.667	0.107	0.192	79.907	0.259	0.350	35.221	0.275
	Ind	R@10	0.134	0.217	61.940	0.207	0.217	4.831	0.451	0.474	5.099	0.536
		N@10	0.146	0.194	32.877	0.158	0.192	21.523	0.270	0.370	37.037	0.322
		P@10	0.020	0.022	10.000	0.009	0.022	144.444	0.045	0.061	44.444	0.054
		F1@10	0.040	0.040	0.000	0.040	0.000	0.000	0.082	0.091	11.500	0.098
		Hits@10	0.184	0.224	21.739	0.205	0.224	9.756	0.454	0.511	12.509	0.539
		MRR@10	0.187	0.189	1.070	0.124	0.186	50.000	0.216	0.299	38.725	0.255

Table 10: We report the performances of SAGE, GAT, HGNN, their enhanced versions with our sampling strategy, and MM-SAN. We report recall@10 (R@10), nDG@10 (N@10), Precision@10 (P@10), f1-score (f1@10), Mean reciprocal rank (MRR@10), Hits@10 scores over the PubMed dataset in transductive (Tran), semi-inductive (Semi), and inductive (Ind) settings. 100% indicates that all the datasets have textual metadata. [75%, 50%, 25%] indicates the percentage of datasets having textual metadata. The columns Gain indicate the percentage gain achieved from the original implementation and its enhanced version with our sampling strategy. In boldface we report the highest performances for each setting and metric.

MES	Setting	Metric	GAT	GAT-RW	Gain (GAT)	SAGE	SAGE-RW	Gain (SAGE)	HGNN	HGNN-RW	Gain (HGNN)	MM-SAN
100%	Tran	R@10	0.393	0.394	0.254	0.435	0.455	4.609	0.342	0.382	11.759	0.466
		N@10	0.215	0.236	9.767	0.266	0.321	20.684	0.224	0.265	18.408	0.298
		P@10	0.026	0.046	48.000	0.030	0.046	44.000	0.035	0.040	14.280	0.047
		F1@10	0.038	0.047	19.148	0.069	0.083	20.290	0.063	0.075	19.048	0.085
		Hits@10	0.406	0.456	12.319	0.456	0.458	0.435	0.342	0.392	14.706	0.467
		MRR@10	0.185	0.242	30.302	0.235	0.279	18.723	0.187	0.198	5.856	0.245
	Semi	R@10	0.348	0.388	11.494	0.349	0.447	28.032	0.335	0.382	13.873	0.442
		N@10	0.227	0.247	8.817	0.214	0.324	51.401	0.221	0.258	16.724	0.289
		P@10	0.020	0.032	60.000	0.033	0.045	36.364	0.034	0.054	58.588	0.045
		F1@10	0.038	0.054	42.105	0.068	0.082	20.588	0.062	0.064	3.226	0.081
		Hits@10	0.352	0.425	20.732	0.388	0.448	15.573	0.336	0.351	4.463	0.444
		MRR@10	0.229	0.231	0.873	0.219	0.285	30.145	0.185	0.224	21.108	0.242
	Ind	R@10	0.413	0.457	10.677	0.392	0.418	6.633	0.302	0.417	38.413	0.426
		N@10	0.269	0.272	3.344	0.250	0.294	17.647	0.196	0.234	19.387	0.273
		P@10	0.010	0.019	90.000	0.025	0.042	68.000	0.031	0.061	100.000	0.043
		F1@10	0.051	0.063	94.118	0.009	0.017	100.556	0.056	0.080	36.000	0.078
		Hits@10	0.344	0.352	1.736	0.342	0.421	23.106	0.304	0.341	12.114	0.428
		MRR@10	0.184	0.220	19.565	0.130	0.256	96.923	0.163	0.167	2.448	0.225
75%	Tran	R@10	0.328	0.390	18.902	0.404	0.439	12.646	0.289	0.384	32.893	0.406
		N@10	0.186	0.189	1.613	0.173	0.256	47.950	0.188	0.250	33.620	0.255
		P@10	0.000	0.000	0.000	0.036	0.040	12.300	0.029	0.040	11.976	0.041
		F1@10	0.013	0.030	102.000	0.030	0.065	101.000	0.053	0.059	11.797	0.075
		Hits@10	0.320	0.352	9.091	0.354	0.360	1.667	0.290	0.381	31.414	0.408
		MRR@10	0.117	0.150	28.205	0.127	0.225	76.775	0.133	0.156	14.744	0.209
	Semi	R@10	0.310	0.313	1.290	0.316	0.349	10.426	0.295	0.366	24.151	0.391
		N@10	0.234	0.286	23.936	0.207	0.254	22.715	0.190	0.195	2.632	0.252
		P@10	0.000	0.000	0.000	0.036	0.000	0.000	0.030	0.036	23.452	0.040
		F1@10	0.002	0.005	100.000	0.067	0.074	4.493	0.054	0.072	77.778	0.072
		Hits@10	0.330	0.375	13.636	0.349	0.350	0.286	0.296	0.310	4.735	0.393
		MRR@10	0.152	0.165	24.342	0.160	0.225	40.625	0.157	0.162	3.189	0.209
	Ind	R@10	0.289	0.359	24.207	0.295	0.326	10.517	0.262	0.358	36.518	0.366
		N@10	0.157	0.220	40.761	0.176	0.235	33.333	0.169	0.180	6.526	0.233
		P@10	0.000	0.000	0.000	0.023	0.000	0.000	0.023	0.025	15.203	0.032
		F1@10	0.016	0.043	43.750	0.031	0.041	32.258	0.048	0.085	69.583	0.067
		Hits@10	0.351	0.372	5.397	0.287	0.329	14.674	0.263	0.341	29.680	0.368
		MRR@10	0.118	0.176	49.152	0.154	0.207	34.481	0.141	0.180	29.577	0.225
25%	Tran	R@10	0.277	0.303	9.379	0.236	0.245	10.847	0.193	0.298	54.438	0.311
		N@10	0.112	0.160	42.857	0.115	0.185	60.000	0.120	0.138	15.385	0.184
		P@10	0.015	0.030	100.000	0.006	0.012	100.000	0.020	0.035	75.000	0.031
		F1@10	0.015	0.030	100.000	0.000	0.021	40.000	0.035	0.048	37.142	0.057
		Hits@10	0.258	0.295	14.355	0.240	0.310	50.000	0.194	0.298	53.333	0.312
		MRR@10	0.122	0.157	19.672	0.111	0.143	32.432	0.097	0.153	45.360	0.145
	Semi	R@10	0.264	0.294	15.151	0.228	0.262	47.368	0.187	0.195	4.274	0.293
		N@10	0.128	0.138	7.812	0.109	0.123	12.650	0.117	0.160	36.557	0.174
		P@10	0.000	0.000	0.000	0.012	0.030	106.000	0.010	0.019	50.000	0.030
		F1@10	0.000	0.000	0.000	0.022	0.000	0.000	0.034	0.052	43.294	0.054
		Hits@10	0.275	0.277	1.091	0.288	0.320	57.634	0.253	0.322	27.295	0.342
		MRR@10	0.053	0.091	45.565	0.080	0.105	23.810	0.048	0.095	49.474	0.138
	Ind	R@10	0.223	0.254	13.874	0.234	0.261	11.448	0.182	0.246	35.209	0.270
		N@10	0.144	0.161	11.111	0.063	0.083	31.746	0.113	0.132	16.817	0.162
		P@10	0.000	0.000	0.000	0.011	0.000	0.000	0.019	0.030	28.124	0.027
		F1@10	0.028	0.042	50.000	0.000	0.020	100.000	0.026	0.034	23.529	0.050
		Hits@10	0.225	0.229	1.778	0.175	0.182	36.000	0.184	0.232	26.087	0.272
		MRR@10	0.080	0.111	38.750	0.102	0.175	46.470	0.093	0.120	60.215	0.130

Table 11: We report the performances of SAGE, GAT, HGNN, their enhanced versions with our sampling strategy, and MM-SAN. We report recall@20 (R@20), nDG@20 (N@20), Precision@20 (P@20), f1-score (f1@20), Mean reciprocal rank (MRR@20), Hits@20 scores over the MES dataset in transductive (Tran), semi-inductive (Semi), and inductive (Ind) settings. 100% indicates that all the datasets have textual metadata. [75%, 50%, 25%] indicates the percentage of datasets having textual metadata. The columns Gain indicate the percentage gain achieved from the original implementation and its enhanced version with our sampling strategy. In boldface we report the highest performances for each setting and metric.

MES	Setting	Metric	GAT	GAT-RW	Gain (GAT)	SAGE	SAGE-RW	Gain (SAGE)	HGNN	HGNN-RW	Gain (HGNN)	MM-SAN
100%	Tran	R@20	0.769	0.833	8.319	0.747	0.840	12.465	0.707	0.773	9.339	0.867
		N@20	0.563	0.583	3.550	0.570	0.580	1.754	0.468	0.470	0.426	0.584
		P@20	0.039	0.044	12.821	0.041	0.044	7.317	0.037	0.045	21.622	0.046
		F1@20	0.031	0.083	167.741	0.043	0.084	95.327	0.070	0.082	17.143	0.086
		Hits@20	0.811	0.837	3.211	0.824	0.844	2.422	0.714	0.726	1.679	0.871
		MRR@20	0.486	0.490	0.823	0.491	0.495	0.816	0.395	0.410	3.797	0.498
	Semi	R@20	0.730	0.833	14.113	0.768	0.867	13.001	0.723	0.764	5.658	0.886
		N@20	0.568	0.572	0.701	0.542	0.559	3.134	0.457	0.466	1.975	0.573
		P@20	0.004	0.042	950.000	0.020	0.044	100.000	0.037	0.044	18.920	0.045
		F1@20	0.080	0.081	1.250	0.077	0.084	9.092	0.070	0.083	18.571	0.086
		Hits@20	0.735	0.833	13.333	0.777	0.871	12.062	0.727	0.760	4.536	0.886
		MRR@20	0.476	0.486	2.105	0.476	0.487	2.313	0.375	0.387	3.200	0.477
	Ind	R@20	0.738	0.808	9.495	0.778	0.806	3.594	0.711	0.717	0.843	0.809
		N@20	0.554	0.554	0.000	0.554	0.612	10.477	0.460	0.491	6.452	0.555
		P@20	0.030	0.040	33.333	0.013	0.041	215.385	0.036	0.040	11.111	0.041
		F1@20	0.033	0.077	133.333	0.009	0.079	777.778	0.069	0.077	11.594	0.078
		Hits@20	0.758	0.808	6.603	0.766	0.816	6.537	0.711	0.773	8.703	0.809
		MRR@20	0.436	0.446	2.294	0.472	0.479	1.479	0.383	0.454	18.502	0.477
75%	Tran	R@20	0.558	0.636	14.027	0.540	0.622	15.185	0.609	0.697	14.457	0.765
		N@20	0.444	0.454	2.248	0.449	0.495	10.225	0.381	0.471	23.642	0.495
		P@20	0.000	0.034	3400.000	0.000	0.033	0.000	0.032	0.039	21.429	0.040
		F1@20	0.000	0.064	6400.000	0.040	0.062	48.000	0.061	0.075	22.951	0.076
		Hits@20	0.605	0.639	5.626	0.532	0.626	17.675	0.619	0.718	16.001	0.769
		MRR@20	0.411	0.411	0.000	0.400	0.410	2.500	0.314	0.317	0.955	0.412
	Semi	R@20	0.569	0.598	5.098	0.556	0.614	10.400	0.674	0.740	9.797	0.818
		N@20	0.410	0.450	9.756	0.393	0.457	16.283	0.388	0.465	19.872	0.466
		P@20	0.022	0.031	40.909	0.000	0.031	3100.000	0.034	0.040	17.647	0.041
		F1@20	0.021	0.042	100.000	0.046	0.060	30.435	0.065	0.078	20.000	0.079
		Hits@20	0.516	0.598	15.917	0.542	0.614	13.302	0.674	0.762	13.083	0.826
		MRR@20	0.360	0.360	0.000	0.327	0.407	24.475	0.300	0.360	20.000	0.361
	Ind	R@20	0.613	0.622	1.469	0.617	0.641	3.880	0.658	0.736	11.855	0.799
		N@20	0.479	0.479	0.000	0.436	0.475	8.973	0.422	0.426	0.949	0.480
		P@20	0.020	0.032	43.000	0.000	0.033	0.000	0.034	0.039	14.706	0.041
		F1@20	0.041	0.060	46.341	0.016	0.062	287.500	0.069	0.077	11.594	0.078
		Hits@20	0.594	0.625	5.214	0.568	0.645	13.531	0.711	0.773	8.718	0.809
		MRR@20	0.388	0.388	0.000	0.383	0.424	10.688	0.383	0.454	18.574	0.477
50%	Tran	R@20	0.313	0.374	19.472	0.337	0.395	17.215	0.561	0.653	16.359	0.779
		N@20	0.263	0.331	26.042	0.217	0.311	43.249	0.351	0.374	6.550	0.440
		P@20	0.018	0.020	11.000	0.000	0.021	0.000	0.030	0.040	33.333	0.041
		F1@20	0.027	0.037	37.037	0.011	0.039	254.545	0.056	0.076	34.667	0.077
		Hits@20	0.340	0.381	12.000	0.345	0.401	16.230	0.565	0.577	2.124	0.789
		MRR@20	0.230	0.321	39.565	0.201	0.289	43.286	0.292	0.318	8.849	0.344
	Semi	R@20	0.349	0.436	25.030	0.382	0.458	20.000	0.629	0.696	10.648	0.758
		N@20	0.291	0.337	15.822	0.333	0.336	0.900	0.360	0.426	18.333	0.427
		P@20	0.014	0.022	57.143	0.000	0.023	0.000	0.032	0.038	19.847	0.039
		F1@20	0.021	0.042	100.000	0.041	0.044	7.317	0.061	0.072	18.032	0.073
		Hits@20	0.357	0.439	23.029	0.409	0.462	13.008	0.629	0.697	10.783	0.758
		MRR@20	0.218	0.310	42.227	0.212	0.300	41.509	0.280	0.322	15.000	0.328
	Ind	R@20	0.321	0.368	14.622	0.290	0.368	26.897	0.582	0.623	7.042	0.691
		N@20	0.270	0.304	12.593	0.213	0.290	36.148	0.376	0.425	13.021	0.426
		P@20	0.000	0.019	1900.000	0.000	0.019	0.000	0.030	0.034	13.333	0.035
		F1@20	0.004	0.006	20.000	0.026	0.036	38.462	0.049	0.060	22.449	0.062
		Hits@20	0.321	0.395	23.049	0.377	0.375	-0.531	0.586	0.625	6.667	0.697
		MRR@20	0.199	0.287	44.178	0.169	0.267	58.588	0.317	0.348	9.772	0.349
25%	Tran	R@20	0.113	0.209	85.840	0.150	0.209	39.333	0.460	0.507	10.217	0.667
		N@20	0.132	0.162	22.727	0.066	0.157	137.879	0.247	0.330	33.541	0.389
		P@20	0.033	0.051	54.545	0.000	0.011	0.000	0.024	0.033	37.500	0.034
		F1@20	0.040	0.061	52.500	0.000	0.021	47.222	0.046	0.064	39.167	0.065
		Hits@20	0.139	0.218	56.788	0.126	0.218	73.015	0.563	0.569	0.536	0.667
		MRR@20	0.129	0.154	19.385	0.064	0.148	131.250	0.187	0.234	25.066	0.305
	Semi	R@20	0.160	0.258	61.250	0.196	0.265	35.327	0.530	0.621	17.170	0.705
		N@20	0.112	0.202	80.357	0.121	0.209	73.553	0.324	0.380	17.213	0.381
		P@20	0.022	0.033	50.000	0.008	0.014	75.000	0.027	0.035	29.630	0.036
		F1@20	0.040	0.055	37.500	0.000	0.026	150.000	0.051	0.067	31.429	0.068
		Hits@20	0.245	0.258	5.316	0.176	0.265	51.137	0.530	0.567	7.021	0.705
		MRR@20	0.125	0.184	47.200	0.141	0.193	36.957	0.264	0.284	7.575	0.285
	Ind	R@20	0.217	0.230	5.990	0.134	0.217	62.314	0.513	0.578	12.719	0.645
		N@20	0.147	0.197	34.015	0.129	0.192	48.440	0.286	0.290	1.418	0.350
		P@20	0.010	0.022	120.000	0.010	0.011	10.000	0.026	0.032	23.077	0.033
		F1@20	0.015	0.022	46.667	0.012	0.021	75.000	0.049	0.060	22.448	0.062
		Hits@20	0.184	0.237	28.804	0.169	0.224	32.456	0.513	0.537	4.684	0.651
		MRR@20	0.118	0.190	61.864	0.087	0.186	113.793	0.220	0.262	19.091	0.263

Table 12: We report the performances of SAGE, GAT, HGNN, their enhanced versions with our sampling strategy, and MM-SAN. We report recall@20 (R@20), nDG@20 (N@20), Precision@20 (P@20), f1-score (f1@20), Mean reciprocal rank (MRR@20), Hits@20 scores over the PubMed dataset in transductive (Tran), semi-inductive (Semi), and inductive (Ind) settings. 100% indicates that all the datasets have textual metadata. [75%, 50%, 25%] indicates the percentage of datasets having textual metadata. The columns Gain indicate the percentage gain achieved from the original implementation and its enhanced version with our sampling strategy. In boldface we report the highest performances for each setting and metric.

PubMed	Setting	Metric	GAT	GAT-RW	Gain (GAT)	SAGE	SAGE-RW	Gain (SAGE)	HGNN	HGNN-RW	Gain (HGNN)	MM-SAN
100%	Tran	R@20	0.513	0.525	2.335	0.505	0.516	2.178	0.396	0.427	7.831	0.528
		N@20	0.303	0.304	0.330	0.300	0.336	12.000	0.238	0.267	12.167	0.313
		P@20	0.011	0.026	136.364	0.010	0.026	210.000	0.020	0.041	105.000	0.027
		F1@20	0.032	0.034	6.250	0.016	0.050	212.500	0.038	0.077	102.632	0.051
		Hits@20	0.596	0.650	7.718	0.556	0.567	7.030	0.398	0.472	18.184	0.529
		MRR@20	0.248	0.249	0.403	0.242	0.283	17.000	0.191	0.194	1.576	0.290
	Semi	R@20	0.506	0.510	0.792	0.484	0.493	1.860	0.388	0.391	0.774	0.511
		N@20	0.264	0.295	11.702	0.264	0.335	27.021	0.234	0.297	26.923	0.307
		P@20	0.019	0.021	10.526	0.010	0.025	250.000	0.020	0.042	110.000	0.026
		F1@20	0.022	0.037	68.182	0.021	0.048	470.000	0.037	0.073	97.297	0.049
		Hits@20	0.521	0.552	5.945	0.497	0.503	1.205	0.390	0.475	21.795	0.513
		MRR@20	0.226	0.228	0.885	0.216	0.288	33.333	0.189	0.252	33.510	0.247
	Ind	R@20	0.439	0.457	4.100	0.416	0.473	13.705	0.352	0.440	25.000	0.496
		N@20	0.280	0.284	1.429	0.267	0.308	15.343	0.209	0.210	0.478	0.291
		P@20	0.014	0.020	42.857	0.022	0.024	9.091	0.018	0.076	322.222	0.025
		F1@20	0.043	0.045	4.651	0.039	0.046	17.949	0.034	0.086	153.333	0.048
		Hits@20	0.453	0.487	7.507	0.480	0.495	3.125	0.394	0.461	16.987	0.499
		MRR@20	0.235	0.235	0.000	0.235	0.259	10.213	0.167	0.221	32.312	0.230
75%	Tran	R@20	0.312	0.485	55.206	0.301	0.399	32.241	0.342	0.370	8.088	0.486
		N@20	0.260	0.262	0.769	0.182	0.267	46.708	0.201	0.250	24.388	0.276
		P@20	0.021	0.024	14.286	0.000	0.020	0.000	0.017	0.079	363.158	0.025
		F1@20	0.023	0.024	4.348	0.012	0.038	300.000	0.033	0.042	27.177	0.047
		Hits@20	0.401	0.485	21.009	0.387	0.399	3.105	0.343	0.370	7.883	0.488
		MRR@20	0.149	0.177	18.872	0.140	0.208	48.571	0.160	0.214	33.750	0.214
	Semi	R@20	0.327	0.335	2.448	0.286	0.382	33.679	0.335	0.429	28.434	0.462
		N@20	0.228	0.264	15.789	0.213	0.263	23.506	0.200	0.280	40.000	0.290
		P@20	0.008	0.020	150.000	0.010	0.019	90.000	0.017	0.020	17.647	0.023
		F1@20	0.041	0.042	2.439	0.010	0.037	270.000	0.032	0.040	25.000	0.045
		Hits@20	0.329	0.417	26.740	0.284	0.382	34.646	0.335	0.379	13.170	0.463
		MRR@20	0.205	0.210	2.439	0.205	0.227	10.732	0.160	0.213	33.125	0.234
	Ind	R@20	0.320	0.331	3.438	0.267	0.358	34.114	0.306	0.370	21.008	0.427
		N@20	0.225	0.239	6.222	0.194	0.243	25.257	0.180	0.198	10.000	0.249
		P@20	0.002	0.004	100.000	0.010	0.013	30.000	0.016	0.043	168.750	0.019
		F1@20	0.017	0.030	76.471	0.011	0.035	218.182	0.030	0.041	37.500	0.041
		Hits@20	0.410	0.411	0.244	0.328	0.361	10.048	0.309	0.388	25.532	0.429
		MRR@20	0.136	0.145	6.618	0.111	0.209	88.288	0.144	0.209	45.208	0.197
25%	Tran	R@20	0.202	0.232	14.851	0.101	0.128	26.730	0.236	0.284	20.339	0.382
		N@20	0.192	0.200	4.167	0.061	0.088	44.262	0.131	0.208	58.024	0.202
		P@20	0.001	0.003	200.000	0.001	0.002	100.000	0.012	0.073	508.333	0.019
		F1@20	0.023	0.023	0.000	0.012	0.000	0.000	0.023	0.030	30.435	0.037
		Hits@20	0.144	0.228	58.333	0.037	0.130	251.351	0.237	0.310	30.712	0.383
		MRR@20	0.051	0.142	178.431	0.000	0.076	100.000	0.100	0.108	8.000	0.150
	Semi	R@20	0.293	0.316	7.856	0.049	0.125	155.102	0.226	0.276	22.125	0.370
		N@20	0.145	0.164	13.103	0.067	0.091	35.820	0.126	0.174	38.888	0.194
		P@20	0.004	0.010	150.000	0.000	0.006	50.000	0.011	0.014	27.273	0.019
		F1@20	0.023	0.026	13.043	0.004	0.012	200.000	0.022	0.030	34.788	0.036
		Hits@20	0.242	0.289	19.508	0.097	0.126	29.897	0.228	0.243	6.579	0.372
		MRR@20	0.070	0.080	14.286	0.040	0.081	102.500	0.098	0.137	40.816	0.144
	Ind	R@20	0.120	0.171	42.500	0.107	0.119	11.214	0.224	0.234	4.464	0.345
		N@20	0.119	0.160	34.454	0.044	0.086	50.000	0.124	0.157	26.667	0.181
		P@20	0.003	0.008	166.667	0.000	0.006	100.000	0.011	0.032	54.545	0.017
		F1@20	0.031	0.032	3.225	0.000	0.011	100.000	0.022	0.026	18.182	0.033
		Hits@20	0.342	0.342	0.000	0.029	0.120	314.482	0.226	0.310	37.168	0.346
		MRR@20	0.030	0.053	43.000	0.098	0.115	23.469	0.095	0.134	41.053	0.135

Table 13: Hyperparameters tested via grid search for each baseline. We put "-" when the hyperparameter does not apply to the method.

Model	Learning Rate	Hidden Dimensions	Layers	Heads	Neighbors
GraphSAGE	{1e-6, 1e-5, 5e-5, 1e-4, 1e-3, 1e-2}	{64, 128, 256}	{2, 3}	-	{5,10, 15}
GAT	{1e-6, 1e-5, 5e-5, 1e-4, 1e-3, 1e-2}	{64, 128, 256}	{2, 3}	-	{5,10, 15}
HGT	{1e-6, 1e-5, 5e-5, 1e-4, 1e-3, 1e-2}	{64, 128, 256}	{2, 3}	{1,2,4,8}	-
HAN	{1e-6, 1e-5, 5e-5, 1e-4, 1e-3, 1e-2}	{64, 128, 256}	{2, 3}	{1,2,4,8}	-
HGNN	{1e-6, 1e-5, 5e-5, 1e-4, 1e-3, 1e-2}	{64, 128, 256}	{2, 3}	{1,2,4,8}	{5,10,15}

Table 14: Random sampling component-based analysis. Evaluation metrics over the PubMed dataset in transductive (Tran), semi-inductive (Semi), and inductive (Ind) settings. Metrics include Recall@5 (R@5), nDG@5 (N@5). 100% indicates full textual metadata; [75%, 50%, 25%] indicate partial metadata availability. RS-*i* refers to Random Sampling at *i*-th depth.

MES	Setting	Metric	RS-1	RS-2	RS-3	MMSAN
100%	Tran	R@5	0.602	0.637	0.644	0.727
		N@5	0.434	0.471	0.475	0.546
		P@5	0.127	0.135	0.145	0.152
		F1@5	0.208	0.221	0.221	0.250
		Hits@5	0.605	0.646	0.653	0.741
		MRR@5	0.413	0.419	0.428	0.484
	Semi	R@5	0.669	0.677	0.702	0.719
		N@5	0.519	0.527	0.535	0.524
		P@5	0.138	0.143	0.145	0.145
		F1@5	0.230	0.232	0.237	0.242
		Hits@5	0.669	0.697	0.712	0.727
		MRR@5	0.429	0.440	0.456	0.461
	Ind	R@5	0.651	0.635	0.662	0.687
		N@5	0.494	0.452	0.465	0.519
		P@5	0.132	0.128	0.128	0.149
		F1@5	0.219	0.212	0.222	0.232
		Hits@5	0.658	0.638	0.662	0.691
		MRR@5	0.443	0.452	0.462	0.466
75%	Tran	R@5	0.602	0.615	0.619	0.623
		N@5	0.513	0.513	0.515	0.447
		P@5	0.122	0.120	0.127	0.128
		F1@5	0.203	0.205	0.207	0.210
		Hits@5	0.605	0.617	0.635	0.659
		MRR@5	0.475	0.482	0.493	0.495
	Semi	R@5	0.576	0.627	0.630	0.598
		N@5	0.402	0.437	0.438	0.442
		P@5	0.113	0.131	0.145	0.146
		F1@5	0.222	0.224	0.222	0.227
		Hits@5	0.577	0.628	0.630	0.630
		MRR@5	0.333	0.344	0.372	0.398
	Ind	R@5	0.515	0.525	0.598	0.598
		N@5	0.407	0.410	0.430	0.468
		P@5	0.111	0.111	0.116	0.116
		F1@5	0.191	0.190	0.191	0.192
		Hits@5	0.579	0.575	0.609	0.609
		MRR@5	0.312	0.313	0.330	0.365
50%	Tran	R@5	0.407	0.425	0.439	0.539
		N@5	0.310	0.310	0.358	0.358
		P@5	0.124	0.136	0.132	0.137
		F1@5	0.205	0.224	0.218	0.208
		Hits@5	0.502	0.507	0.516	0.538
		MRR@5	0.408	0.404	0.402	0.416
	Semi	R@5	0.616	0.623	0.636	0.518
		N@5	0.413	0.414	0.458	0.359
		P@5	0.118	0.131	0.137	0.138
		F1@5	0.229	0.230	0.232	0.239
		Hits@5	0.619	0.627	0.636	0.638
		MRR@5	0.409	0.455	0.399	0.403
	Ind	R@5	0.525	0.525	0.531	0.535
		N@5	0.438	0.442	0.451	0.465
		P@5	0.126	0.136	0.154	0.199
		F1@5	0.110	0.150	0.160	0.164
		Hits@5	0.625	0.622	0.672	0.558
		MRR@5	0.256	0.285	0.286	0.326
25%	Tran	R@5	0.431	0.418	0.429	0.467
		N@5	0.337	0.371	0.391	0.328
		P@5	0.088	0.090	0.090	0.095
		F1@5	0.140	0.143	0.147	0.157
		Hits@5	0.439	0.426	0.419	0.473
		MRR@5	0.218	0.253	0.223	0.283
	Semi	R@5	0.425	0.427	0.429	0.454
		N@5	0.236	0.221	0.271	0.306
		P@5	0.082	0.075	0.072	0.086
		F1@5	0.119	0.124	0.120	0.144
		Hits@5	0.429	0.427	0.429	0.462
		MRR@5	0.197	0.172	0.109	0.256
	Ind	R@5	0.409	0.428	0.438	0.453
		N@5	0.256	0.281	0.252	0.321
		P@5	0.078	0.072	0.080	0.088
		F1@5	0.105	0.121	0.100	0.147
		Hits@5	0.412	0.434	0.439	0.441
		MRR@5	0.206	0.224	0.203	0.241

Table 15: Random sampling component-based analysis. Evaluation metrics over the PubMed dataset in transductive (Tran), semi-inductive (Semi), and inductive (Ind) settings. Metrics include Recall@5 (R@5), nDG@5 (N@5). 100% indicates full textual metadata; [75%, 50%, 25%] indicate partial metadata availability. RS-*i* refers to Random Sampling at *i*-th depth.

PubMed	Setting	Metric	RS-1	RS-2	RS-3	MM-SAN
100%	Tran	R@5	0.314	0.313	0.295	0.380
		N@5	0.243	0.225	0.212	0.270
		P@5	0.063	0.063	0.060	0.077
		F1@5	0.105	0.105	0.099	0.127
		Hits@5	0.315	0.314	0.295	0.381
		MRR@5	0.220	0.197	0.184	0.234
	Semi	R@5	0.301	0.307	0.290	0.374
		N@5	0.231	0.217	0.207	0.267
		P@5	0.061	0.062	0.059	0.076
		F1@5	0.101	0.103	0.098	0.126
		Hits@5	0.302	0.309	0.294	0.376
		MRR@5	0.208	0.189	0.181	0.232
	Ind	R@5	0.256	0.271	0.251	0.344
		N@5	0.191	0.190	0.179	0.246
		P@5	0.052	0.055	0.051	0.070
		F1@5	0.086	0.091	0.084	0.116
		Hits@5	0.258	0.274	0.253	0.347
		MRR@5	0.170	0.164	0.156	0.214
75%	Tran	R@5	0.317	0.318	0.318	0.314
		N@5	0.145	0.130	0.137	0.159
		P@5	0.064	0.064	0.064	0.067
		F1@5	0.106	0.107	0.106	0.098
		Hits@5	0.319	0.319	0.318	0.317
		MRR@5	0.204	0.204	0.214	0.216
	Semi	R@5	0.281	0.259	0.285	0.292
		N@5	0.132	0.122	0.127	0.151
		P@5	0.061	0.063	0.062	0.065
		F1@5	0.101	0.104	0.103	0.104
		Hits@5	0.305	0.311	0.308	0.326
		MRR@5	0.211	0.196	0.204	0.211
	Ind	R@5	0.281	0.299	0.302	0.262
		N@5	0.133	0.129	0.133	0.142
		P@5	0.061	0.063	0.062	0.065
		F1@5	0.102	0.105	0.103	0.104
		Hits@5	0.302	0.301	0.318	0.321
		MRR@5	0.209	0.186	0.201	0.210
50%	Tran	R@5	0.312	0.311	0.312	0.317
		N@5	0.233	0.220	0.232	0.235
		P@5	0.062	0.062	0.062	0.052
		F1@5	0.104	0.103	0.104	0.096
		Hits@5	0.312	0.311	0.312	0.317
		MRR@5	0.220	0.192	0.206	0.196
	Semi	R@5	0.292	0.308	0.300	0.316
		N@5	0.223	0.222	0.222	0.225
		P@5	0.060	0.062	0.061	0.071
		F1@5	0.099	0.103	0.101	0.104
		Hits@5	0.293	0.309	0.301	0.318
		MRR@5	0.201	0.192	0.196	0.196
	Ind	R@5	0.270	0.265	0.268	0.277
		N@5	0.200	0.191	0.198	0.202
		P@5	0.055	0.054	0.055	0.058
		F1@5	0.091	0.089	0.090	0.092
		Hits@5	0.271	0.266	0.269	0.278
		MRR@5	0.180	0.158	0.169	0.185
25%	Tran	R@5	0.221	0.220	0.221	0.230
		N@5	0.244	0.236	0.240	0.246
		P@5	0.064	0.064	0.064	0.068
		F1@5	0.107	0.106	0.107	0.098
		Hits@5	0.321	0.320	0.321	0.322
		MRR@5	0.224	0.213	0.219	0.231
	Semi	R@5	0.215	0.214	0.210	0.217
		N@5	0.232	0.231	0.231	0.239
		P@5	0.062	0.063	0.063	0.067
		F1@5	0.103	0.105	0.104	0.110
		Hits@5	0.306	0.316	0.311	0.319
		MRR@5	0.212	0.204	0.208	0.188
	Ind	R@5	0.217	0.217	0.212	0.217
		N@5	0.183	0.182	0.188	0.189
		P@5	0.053	0.055	0.054	0.056
		F1@5	0.087	0.091	0.089	0.092
		Hits@5	0.267	0.278	0.273	0.279
		MRR@5	0.163	0.165	0.164	0.166

Table 16: Recall@5 (R@5) and nDG@5 (N@5) scores over the MES dataset in transductive and 100% metadata available setting. The column "no augmentation" refers to the models run excluding topics and entities; "augmentation" reports the results on the enriched graph.

Method	Augmentation		No Augmentation	
	R@5	N@5	R@5	N@5
GAT	0.578	0.422	0.625	0.475
GAT-RW	0.710	0.506	0.700	0.432
SAGE	0.472	0.312	0.489	0.398
SAGE-RW	0.655	0.519	0.614	0.510
HGT	0.020	0.020	0.023	0.020
HGNN	0.601	0.488	0.581	0.438
HGNN-RW	0.711	0.531	0.700	0.520
HAN	0.008	0.009	0.008	0.009
MM-SAN	0.727	0.546	0.642	0.496

Table 17: Recall@5 (R@5) and nDG@5 (N@5) scores over the PubMed dataset in transductive and 100% metadata available setting. The column "no augmentation" refers to the models run excluding topics and entities; "augmentation" reports the results on the enriched graph.

Method	Augmentation		No Augmentation	
	R@5	N@5	R@5	N@5
GAT	0.228	0.165	0.241	0.181
GAT-RW	0.314	0.219	0.298	0.202
SAGE	0.217	0.172	0.239	0.179
SAGE-RW	0.288	0.218	0.278	0.201
HGNN	0.294	0.222	0.284	0.202
HGNN-RW	0.321	0.267	0.301	0.256
HGT	0.014	0.015	0.016	0.017
HAN	0.026	0.021	0.030	0.023
MM-SAN	0.380	0.272	0.361	0.261