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] We summarize the main contributions in the introduction and abstract.
   (b) Did you describe the limitations of your work? [Yes] Please see Section 3.4.1.
   (c) Did you discuss any potential negative societal impacts of your work? [N/A]
   (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] The code is included in the supplemental material.
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes] Please see Appendix A.5.
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [Yes] Please see Appendix A.4.
   (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] Please see Appendix A.5.
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] We give the assets sources in the footnotes.
   (b) Did you mention the license of the assets? [Yes] The assets are public.
   (c) Did you include any new assets either in the supplemental material or as a URL? [Yes] We include our code in the supplemental material.
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A]
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [Yes] We use public data that contains no identifiable information or offensive content.
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]

A Appendix

A.1 Pseudo Code

The pseudo code of the proposed PET is displayed in Algorithm 1.

Algorithm 1: PET

Input: Target data instance (row) $X^t = \{x_i^t | i = 1, \ldots, F\}$, Retrieval Pool D_{ret}.

Retrieve K the most relevant data instances for X^t from the retrieval pool D_{ret} to obtain $\{X^{r_1}, \cdots, X^{r_K}\}$ according to Equation (1).

Construct a hypergraph $G = (V_D, V_F, E)$ from hyperedges $\{X^t, X^{r_1}, \cdots, X^{r_K}\}$.

Initialize the node and edge embeddings according to Equations (3), (4), and (5).

for $l \in \{1, \ldots, L\}$ **do**

  for $j \in V_D \cup V_F$ **do**

    $n_j^{(l)} = AttenAGG^{(l)}\left(\{(e_{ij}^{(l-1)} \odot h_i^{(l-1)}) \| h_i^{(l-1)} | \forall i \in N(j)\}\right)$

    $h_j^{(l)} = \sigma(W_N^{(l)}(h_j^{(l-1)} \| n_j^{(l)}))$

  end

  for $(i, j) \in E$ **do**

    $e_{ij}^{(l)} = \sigma(W_E^{(l)}(h_i^{(l)} \| h_j^{(l)} \| e_{ij}^{(l-1)}))$

  end

end

$\hat{y_t} \leftarrow MLP(h_t^{(L)})$

A.2 Visualization Results on More Datasets

In Section 4.3.2, we have presented the t-SNE visualization of the embeddings for Tmall dataset. Here we provide the visualization results for Taobao and Alipay in Figures 5 and 6.

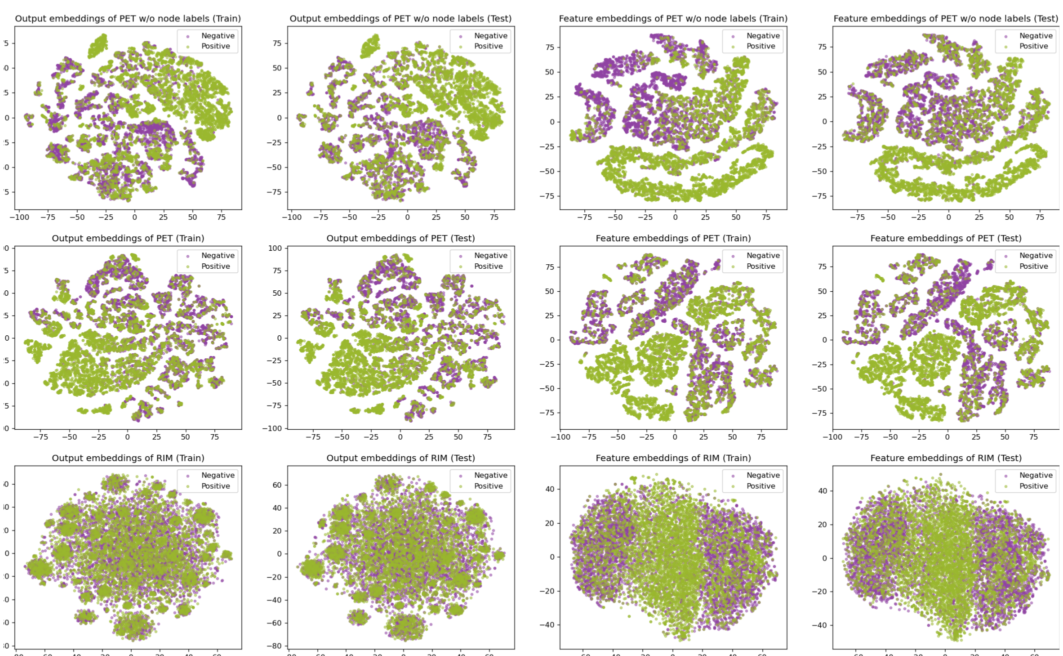

Figure 5: The t-SNE visualization of data and feature embeddings on Taobao.

Similarly, we randomly sample 10,000 data instances from the train data and 10,000 data instances from the test data. Then we visualize the data instance node embeddings (output embeddings) and

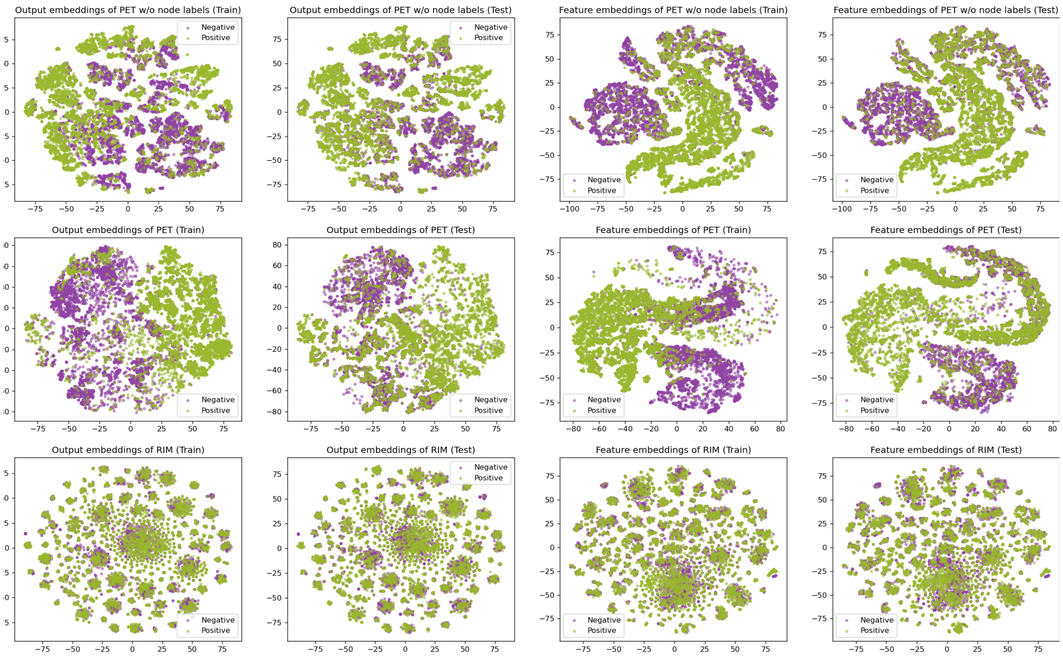

Figure 6: The t-SNE visualization of data and feature embeddings on Alipay.

the feature embeddings with t-SNE. The embeddings of positive data samples are visualized in green, while those of negative data samples are visualized in purple.

From the visualization results, we can see that PET yields more informative embeddings, i.e., the representation of tabular data, that separate positive data and negative data better.

A.3 Impact of Different Retrieval Sizes

We further study the impact of different retrieval sizes. Results are displayed in Figure 7.

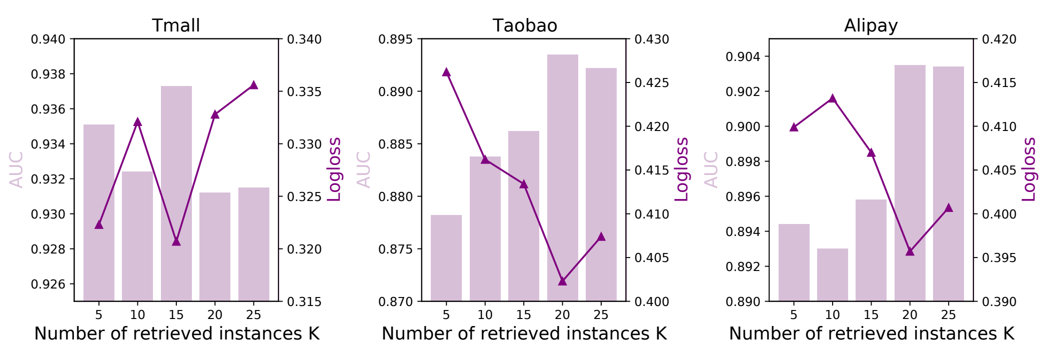

Figure 7: Performance of PET under different retrieval sizes K.

From the results, we can see that the optimal retrieval sizes are similar for different datasets. Generally, more retrieved instances can contain more auxiliary information and give better results, but too many retrieved instances may introduce noises.

A.4 Error Bars

Due to the page limit, we report the standard errors for PET and the second best model here. The results are summarized in Tables 6 and 7.

Table 6: Error bars for the CTR prediction task.

		Tmall		Taobao		Alipay	
		AUC	LogLoss	AUC	LogLoss	AUC	LogLoss
RIM	Mean	0.9138	0.3804	0.8563	0.4644	0.8006	0.5615
	Std.	0.0023	0.0027	0.0021	0.0029	0.0026	0.0031
PET	Mean	0.9324	0.3321	0.8838	0.4162	0.8930	0.4132
	Std.	0.0030	0.0036	0.0022	0.0031	0.0024	0.0027

Table 7: Error bars for the top-N recommendation task.

			HR@1	HR@5	HR@10	NDCG@5	NDCG@10	MRR
ML-1M	RIM	Mean	0.0645	0.2515	0.4014	0.1577	0.2059	0.1704
		Std.	0.0032	0.0063	0.0061	0.0057	0.0048	0.0043
	PET	Mean	0.0904	0.2889	0.4404	0.1903	0.2390	0.2006
		Std.	0.0037	0.0057	0.0073	0.0045	0.0048	0.0041
LastFM	RIM	Mean	0.0915	0.3468	0.5780	0.2165	0.2911	0.2210
		Std.	0.0063	0.0081	0.0137	0.0043	0.0062	0.0042
	PET	Mean	0.1149	0.3621	0.6033	0.2381	0.3156	0.2492
		Std.	0.0098	0.0075	0.0074	0.0092	0.0059	0.0088

From the results, we can see that the performance of PET is stable. In addition, PET consistently outperforms other baselines.

A.5 Experiment Settings

In this section, we offer the detailed hyperparameters settings to reproduce the results. The hyperparameters of PET for each dataset are summarized in Table 8.

Table 8: Hyperparameters for different datasets.

Hyperparameters	Tmall	Taobao	Alipay	ML-1M	LastFM
Embedding Size	16	16	32	16	16
# GNN Layers	3	3	3	2	2
MLP	$[200, 80, 1]$	$[200, 80, 1]$	$[200, 80, 1]$	$[200, 80, 1]$	$[200, 80, 1]$
K	10	10	10	10	10
Batch Size	100	200	100	100	500
Optimizer	Adam	Adam	Adam	Adam	Adam
Learning Rate	5e-4	1e-4	5e-4	1e-3	1e-3
L2 Regularization	1e-4	5e-4	5e-4	1e-4	1e-4

The hyperparameters of PET are inherited from RIM. For the baseline tuning, we follow previous work. Concretely, we use the optimal hyperparameters reported by RIM (Qin et al., 2021) for the non-graph methods. For the graph methods, the learning rate is selected from {1e-4, 3e-4, 5e-4, 1e-3}, and l2 regularization is selected from {1e-4, 1e-5, 5e-5}. The embedding sizes of all the models are consistent to ensure the fair comparison.

We use ElasticSearch[8] to retrieve the relevant data instances. The model is implemented based on DGL[9]. All the experiments are run on Tesla T4 instances.

[8]https://www.elastic.co/

[9]https://www.dgl.ai/