Cold-start Sequential Recommendation via Meta Learner

Advisor: Jia-Ling, Koh

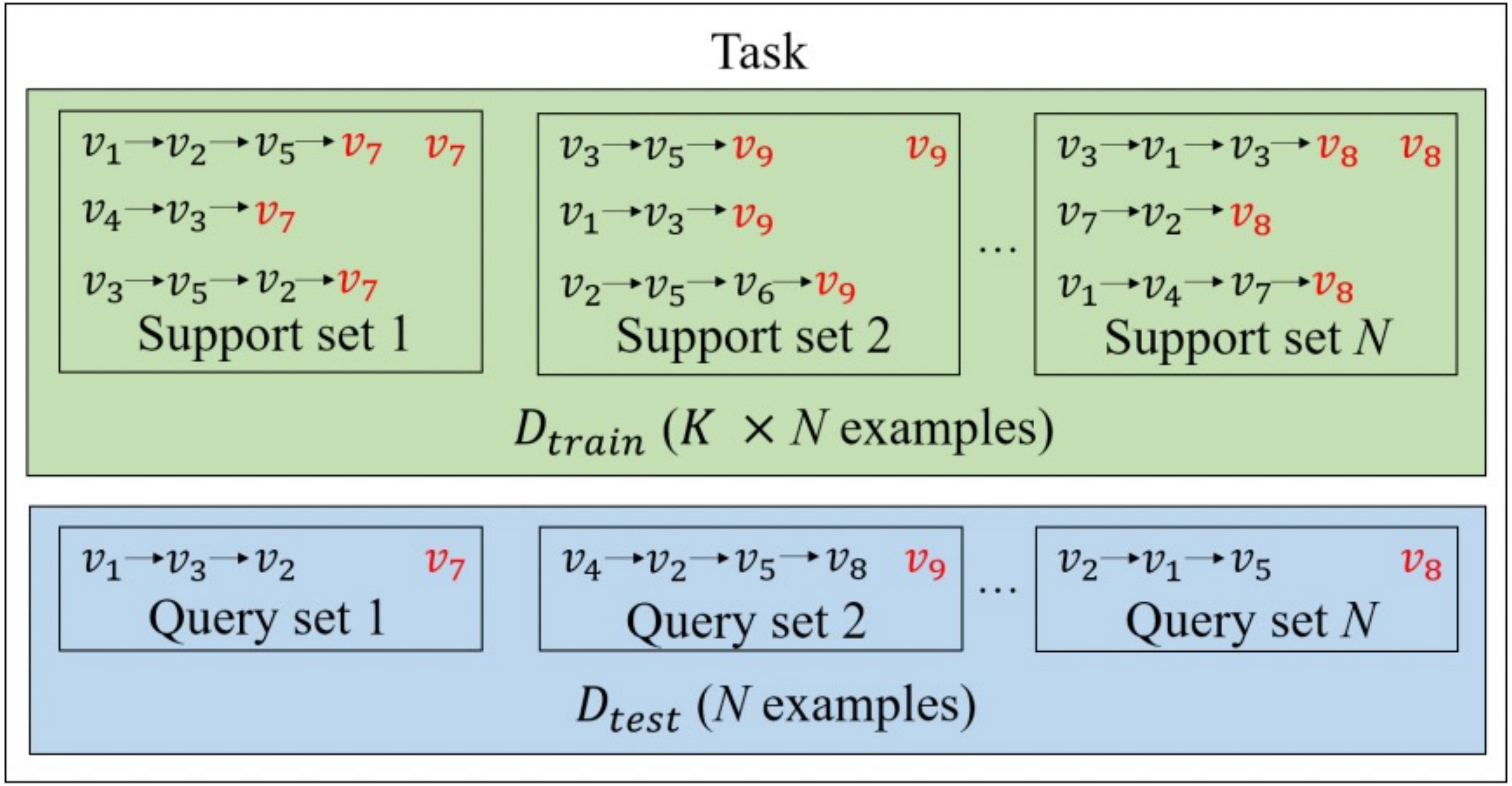
Speaker: Hsiang-Hui Chou

Date: 2023/02/21

Source: AAAI-21

Outline

- Introduction
- Method
- Experiment
- Conclusion


Task
$v_1 \to v_2 \to v_5 \to v_7$ v_7
$v_4 \to v_3 \to v_7$
$v_3 \to v_5 \to v_2 \to v_7$
Support set 1
$v_3 \to v_5 \to v_9$ v_9
$v_1 \to v_3 \to v_9$
$v_2 \to v_5 \to v_6 \to v_9$
Support set 2
…
$v_3 \to v_1 \to v_3 \to v_8$ v_8
$v_7 \to v_2 \to v_8$
$v_1 \to v_4 \to v_7 \to v_8$
Support set N
D_{train} ($K \times N$ examples)
$v_1 \to v_3 \to v_2$ v_7
Query set 1
$v_4 \to v_2 \to v_5 \to v_8$ v_9
Query set 2
…
$v_2 \to v_1 \to v_5$ v_8
Query set N
D_{test} (N examples)

Sequential recommendation

Input
U and V : user set and item set
User U_i interaction sequence $\zeta_i = (v_{i,1}, v_{i,2}, ..., v_{i,n})$

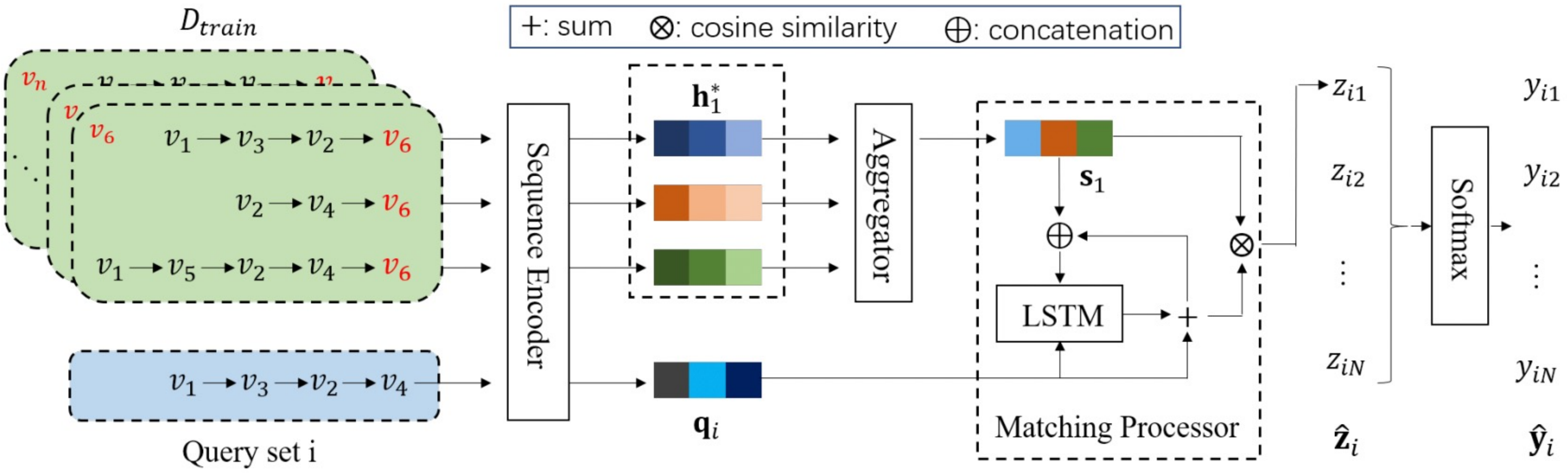
Output
Recommend next item $v_{i,n+1}$

(a score for each candidate item in the item set V)

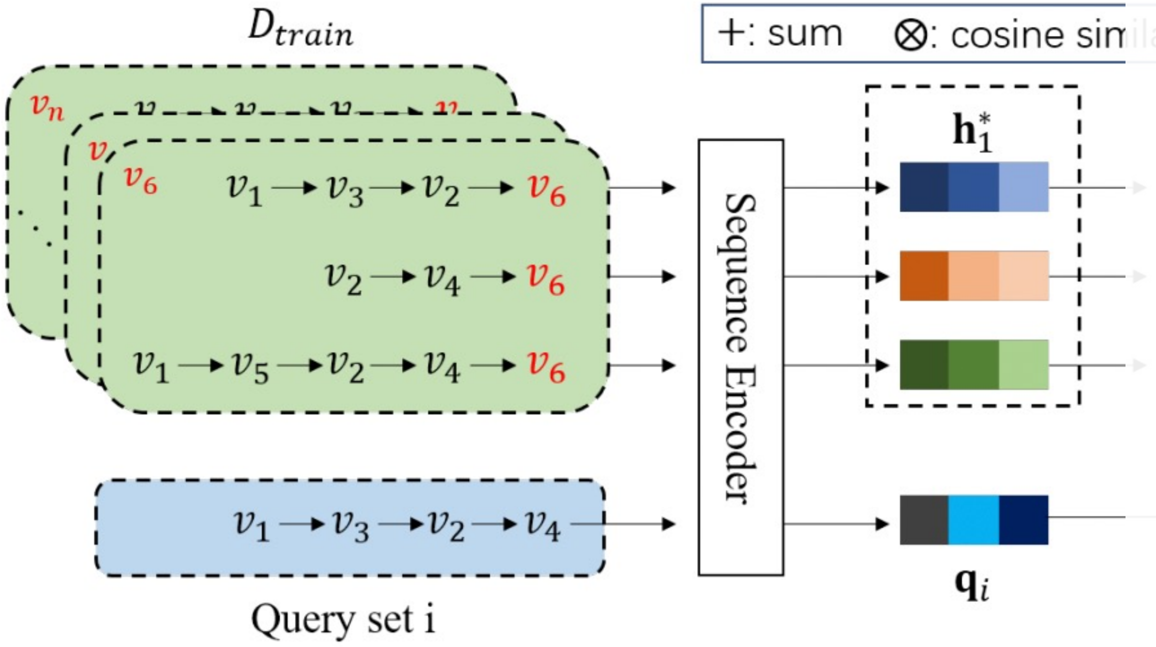
Outline

- Introduction
- Method
- Experiment
- Conclusion

MEta-learning-based COld-start Sequential recommendation framework(Mecos)



Encoding Support and Query Set



$$\mathbf{e}_j = \mathbf{p}^T \left(\mathbf{W}_a^1 \mathbf{v}_{i,n} + \mathbf{W}_a^2 \mathbf{v}_{i,j} + \mathbf{W}_a^3 \mathbf{v}_{i,avg} + \mathbf{b} \right),$$

$$\alpha_j = \frac{exp(\mathbf{e}_j)}{\sum_{k=1}^{n} exp(\mathbf{e}_k)}, \tag{1}$$

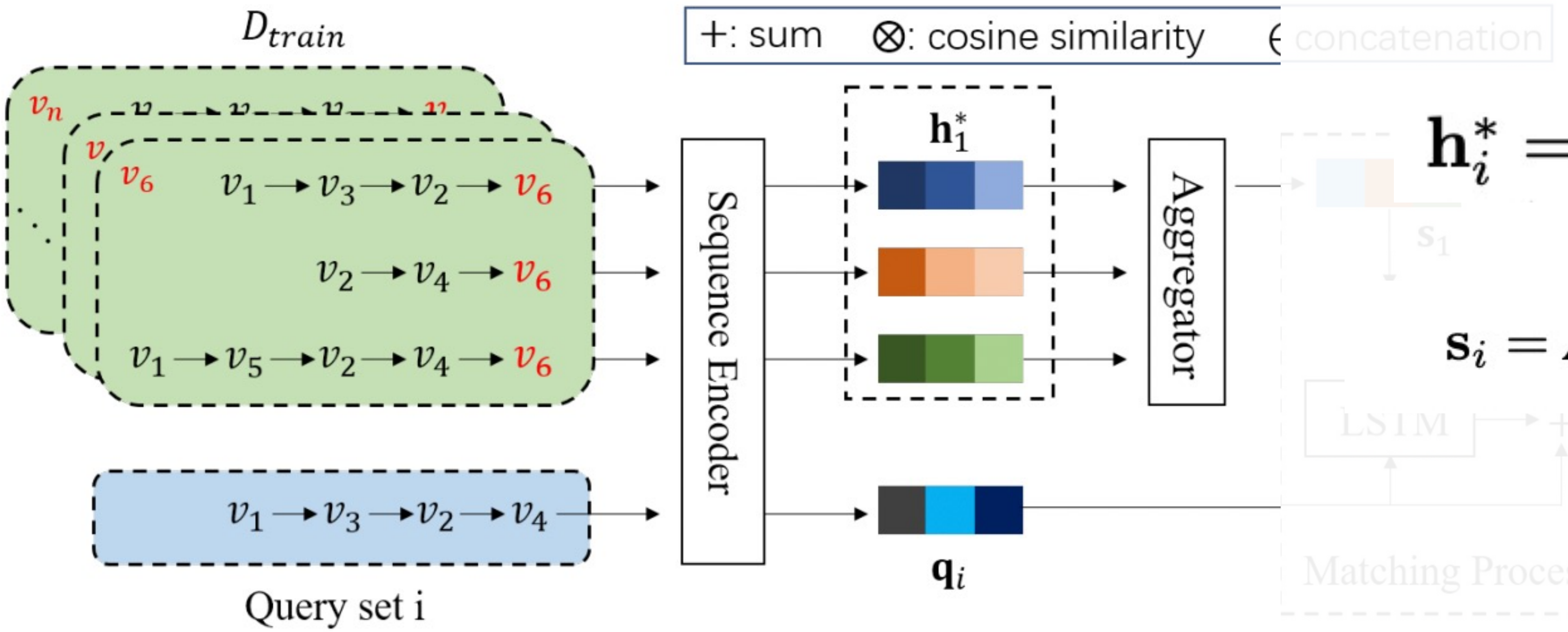
$$\mathbf{R}(\zeta_i) = \sum_{j=1}^{n} \alpha_j \mathbf{v}_{i,j},$$

$$\mathbf{h}_i^0 = \left[\mathbf{R}(\zeta_i); \mathbf{v}_{i,n+1} \right],$$

$$\mathbf{h}_i^l = \mathrm{ReLU} \left(\mathbf{W}^l \mathbf{h}_i^{l-1} + \mathbf{b}^l \right), \tag{2}$$

$$\mathbf{h}_i = \mathbf{h}_i^0 + \mathbf{h}_i^l,$$

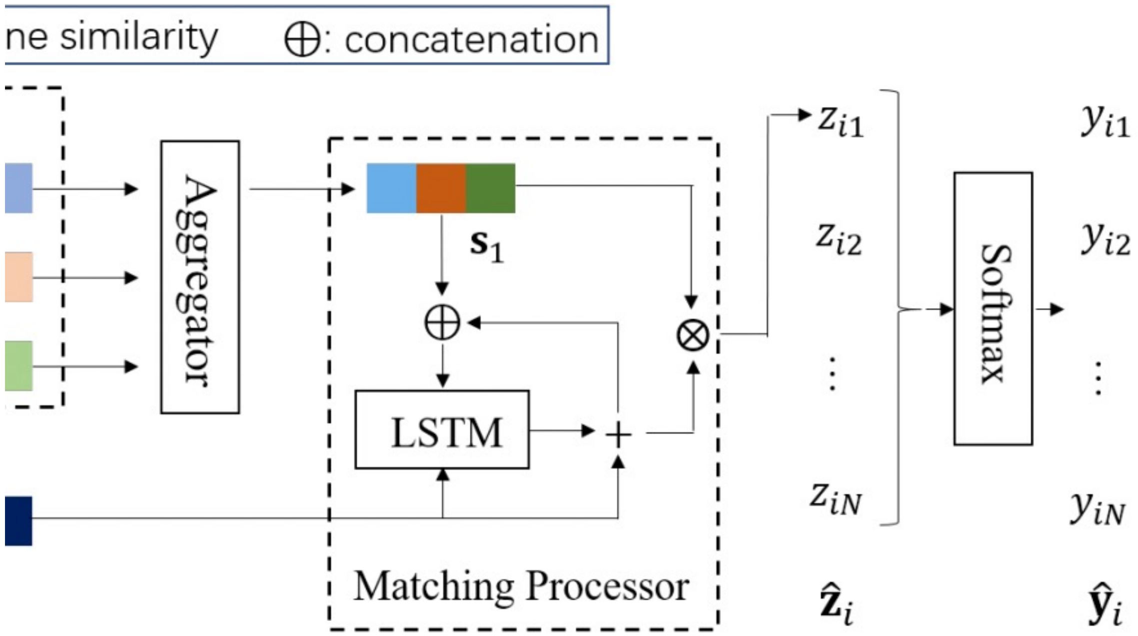
Encoding Support and Query Set



$$\mathbf{h}_i^* = \{\mathbf{h}_{i,1}, \mathbf{h}_{i,2}, ..., \mathbf{h}_{i,K}\}$$

$$\mathbf{s}_i = Aggr(\mathbf{h}_i^*), \quad (3)$$

Matching Support and Query Set



$$\begin{aligned} \hat{\mathbf{q}}_i^t, \mathbf{c}^t &= \mathrm{LSTM}\left(\mathbf{q}_i, \left[\mathbf{q}_i^{t-1}; \mathbf{s}_j\right], \mathbf{c}^{t-1}\right), \\ \mathbf{q}_i^t &= \hat{\mathbf{q}}_i^t + \mathbf{q}_i, \end{aligned} \tag{4}$$

$$z_{ij} = \frac{\mathbf{q}_i^t \mathbf{s}_j}{\|\mathbf{q}_i^t\| \times \|\mathbf{s}_j\|}, \tag{5}$$

$$\hat{\mathbf{y}}_i = \mathrm{softmax}(\hat{\mathbf{z}}_i), \tag{6}$$

loss function

$$\mathcal{L}_T = -\sum_{i=1}^{N}\sum_{j=1}^{N} y_{ij}\log\left(\hat{y}_{ij}\right) + \left(1 - y_{ij}\right)\log\left(1 - \hat{y}_{ij}\right), \tag{7}$$

Outline

- Introduction
- Method
- Experiment
- Conclusion

Dataset

Datasets	**Steam**	**Electronic**	**Tmall**
# sequences	$358,228$	$443,589$	$918,358$
# items	$11,969$	$63,002$	$624,221$
# ground-truth items	$9,496$	$50,331$	$219,560$
Propotion of meta-sequences	0.12	0.31	0.20

Dataset	Metric	GRU4Rec		NARM		Caser		STAMP		SASRec		SR-GNN		Avg.
		w/o	w	w/o	w	w/o	w	w/o	w	w/o	w	w/o	w	Improv.
Steam	HR@5	0.0860	0.2018	0.0862	0.1850	0.0839	**0.2029**	0.0780	0.1861	0.0621	0.1431	0.1119	0.1878	121.33%
	HR@10	0.1498	0.3066	0.1578	0.3002	0.1508	**0.3136**	0.1392	0.2969	0.1113	0.2358	0.1885	0.3085	98.61%
	HR@20	0.2567	0.4544	0.2785	**0.4583**	0.2634	0.4500	0.2441	0.4539	0.2074	0.3751	0.3137	0.4561	70.70%
	NDCG@5	0.0533	**0.1308**	0.0531	0.1206	0.0518	0.1304	0.0486	0.1184	0.0400	0.0955	0.0701	0.1183	129.23%
	NDCG@10	0.0737	0.1639	0.0760	0.1561	0.0732	**0.1651**	0.0681	0.1529	0.0557	0.1245	0.0947	0.1650	112.60%
	NDCG@20	0.1005	0.1991	0.1062	0.1951	0.1014	**0.1997**	0.0944	0.1920	0.0798	0.1588	0.1261	0.1945	89.23%
	MRR	0.0720	0.1412	0.0543	0.1380	0.0728	**0.1456**	0.0689	0.1353	0.0676	0.1138	0.0936	0.1454	95.05%
Electronic	HR@5	0.0745	0.1389	0.0843	0.1479	0.0438	0.1300	0.0464	0.0888	0.0394	0.1029	0.1215	**0.1619**	107.41%
	HR@10	0.1357	0.2431	0.1604	0.2546	0.0870	0.2285	0.0924	0.1652	0.0796	0.1858	0.1935	**0.2591**	91.10%
	HR@20	0.2470	0.3970	0.2949	**0.4158**	0.1740	0.3878	0.1794	0.2995	0.1573	0.3217	0.3236	0.4138	70.65%
	NDCG@5	0.0457	0.0856	0.0507	0.0931	0.0255	0.0807	0.0286	0.0546	0.0230	0.0618	0.0655	**0.0975**	115.98%
	NDCG@10	0.0652	0.1180	0.0751	0.1277	0.0393	0.0975	0.0432	0.0817	0.0359	0.0890	0.0950	**0.1324**	95.92%
	NDCG@20	0.0930	0.1565	0.1088	**0.1682**	0.0610	0.1522	0.0650	0.1122	0.0553	0.1231	0.1202	0.1678	84.53%
	MRR	0.0684	0.1061	0.0699	0.1131	0.0440	0.0903	0.0515	0.0736	0.0514	0.0915	0.0921	**0.1243**	63.01%
Tmall	HR@5	0.2863	0.3947	0.3105	0.4005	0.1006	0.3384	0.2791	0.3719	0.0862	0.3215	0.3263	**0.4030**	105.49%
	HR@10	0.3437	0.4666	0.3860	**0.4727**	0.1724	0.4023	0.3373	0.4386	0.1307	0.3694	0.4081	0.4624	69.59%
	HR@20	0.4371	0.5632	0.4832	**0.5703**	0.2966	0.5070	0.4252	0.5348	0.2086	0.4487	0.5133	0.5616	44.68%
	NDCG@5	0.2470	0.3409	0.2747	**0.3450**	0.0640	0.2914	0.2409	0.3245	0.0578	0.2883	0.2548	0.3375	147.48%
	NDCG@10	0.2655	0.3641	0.2957	**0.3680**	0.0870	0.3107	0.2597	0.3456	0.0721	0.3040	0.2921	0.3609	116.16%
	NDCG@20	0.2889	0.3883	0.3201	**0.3927**	0.1181	0.3339	0.2817	0.3697	0.0915	0.3232	0.3264	0.3848	90.36%
	MRR	0.2580	0.3477	0.2848	0.3510	0.0799	0.2996	0.2520	0.3327	0.0659	0.2984	0.2956	**0.3635**	123.46%

Dataset	Metric	Mecos_R	Variant_1	Variant_2	Variant_3
Steam	HR@10	**0.1606**	0.0933	0.1163	0.0716
	NDCG@10	**0.0805**	0.0445	0.0555	0.0339
	MRR	**0.0551**	0.0405	0.0476	0.0294
Electronic	HR@10	**0.1382**	0.0832	0.1207	0.0770
	NDCG@10	**0.0665**	0.0382	0.0572	0.0329
	MRR	**0.0658**	0.0355	0.0505	0.0314
Tmall	HR@10	**0.3481**	0.3135	0.3230	0.2921
	NDCG@10	**0.2915**	0.2787	0.2737	0.2552
	MRR	**0.2891**	0.2461	0.2503	0.2324

Table 3: Performance of ablation on different components.

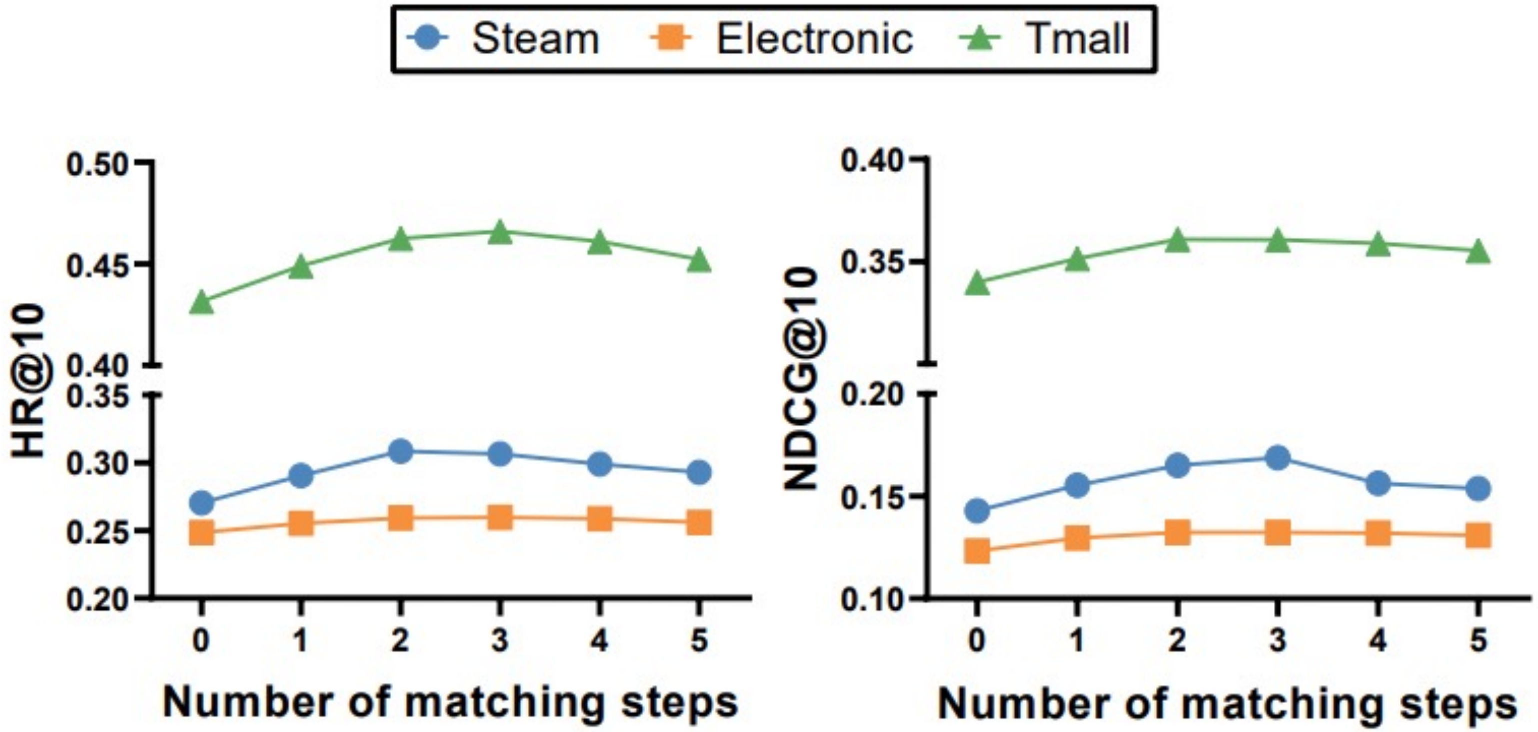
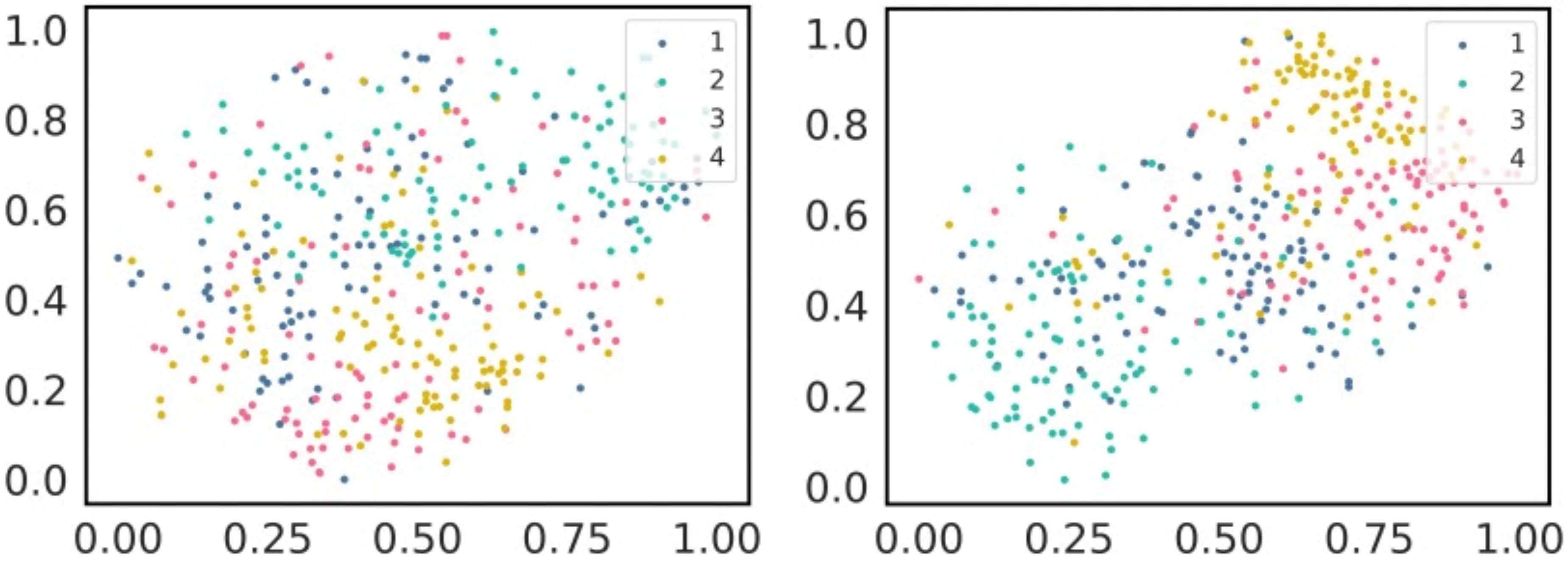


Figure 3: Performance w.r.t. the number of matching steps t.



(a) SR-GNN (b) SR-GNN with Mecos

Outline

- Introduction
- Method
- Experiment
- Conclusion

Conclusion

- Mecos can alleviate the item cold-start problem in sequential recommendation

- With the pre-trained item embeddings as the input, Mecos could be painlessly integrated with other state-of-the-art models