MISSRec: Pre-training and Transferring Multi-modal Interest-aware Sequence Representation for Recommendation

Jinpeng Wang, Ziyun Zeng, Yunxiao Wang, Yuting Wang, Xingyu Lu, Tianxiang Li, Jun Yuan, Rui Zhang, Hai-Tao Zheng, and Shu-Tao Xia.
Tsinghua Shenzhen International Graduate School, Tsinghua University Huawei Noah's Ark Lab Peng Cheng Laboratory
Data & Code: https://github.com/gimpong/MM23-MISSRec Contact: wjp20@mails.tsinghua.edu.cn

Highlights

★ A multi-modal pre-training and transfer learning framework for sequential recommendation.

★ A clustering-based multi-modal interest discovery module.

★ An interest-aware transformer-based encoder-decoder model for multi-modal sequence representation.

Introduction

➢ **Sequential recommendation**: Predicting a user's potential interested items based on her / his historical interactions.

➢ Most existing SR approaches are developed based on ID features, despite the widespread application, suffers from

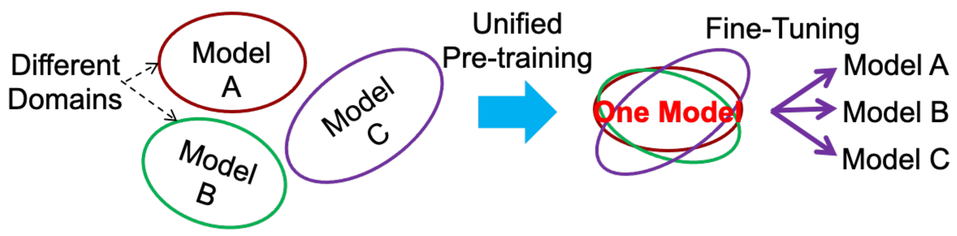
- Under-fitting sparse IDs;
- Cold-start problem;
- Limited transferability across different domains.

➢ Multi-modal information (e.g., texts and images) is ubiquitous and plays an important role in attracting user attention.

➢ Our solution

✦ Use multi-modal features for universal representation.

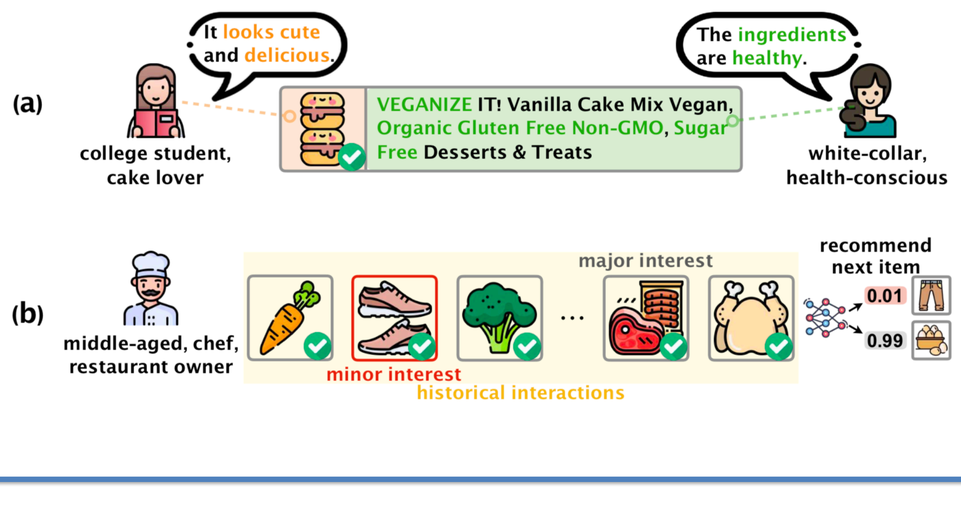
✦ Pre-train a unified model and fine-tune it for transferable recommendation.



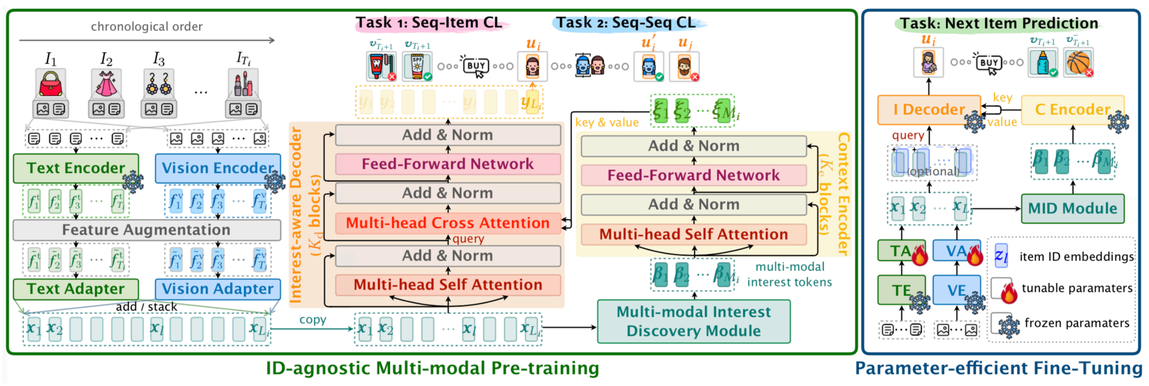
➢ Challenges

✶ The multi-modal synergy is dynamic and user-dependent in the user-item interaction process.

✶ Long-tailed proportion of user interests in the historical interaction.



Our Approach (MISSRec)



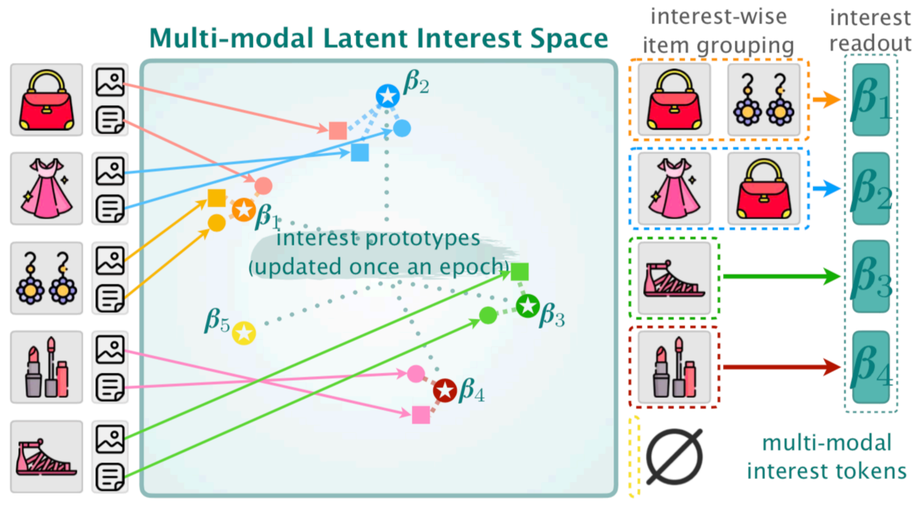
➢ Two stages

1. ID-agnostic multi-modal pre-training;
2. Parameter-efficient fine-tuning in target domains.

➢ Multi-modal Interest Discovery (MID)

✦ Clustering for index construction: *k*-nearest neighbor-based density peaks clustering (DPC-KNN) algorithm;

✦ Interest discovery via indexing.



Experiments

➢ Comparison with state-of-the-arts

Input Type & Model →		ID		T+ID			T+V+ID	*Improv.*	T			T+V	*Improv.*
Dataset	Metric	SASRec	BERT4Rec	FDSA	S^3-Rec	UniSRec	MISSRec	w/ ID	SASRec	ZESRec	UniSRec	MISSRec	w/o ID
Scientific	R@10	0.1080	0.0488	0.0899	0.0525	0.1235	**0.1360**	10.12%	0.0994	0.0851	0.1188	**0.1278**	7.58%
	N@10	0.0553	0.0243	0.0580	0.0275	0.0634	**0.0753**	18.77%	0.0561	0.0475	0.0641	**0.0658**	2.65%
	R@50	0.2042	0.1185	0.1732	0.1418	**0.2473**	0.2431	-	0.2162	0.1746	**0.2394**	0.2375	-
	N@50	0.0760	0.0393	0.0759	0.0468	0.0904	**0.0983**	8.74%	0.0815	0.0670	**0.0903**	0.0893	-
Pantry	R@10	0.0501	0.0308	0.0395	0.0444	0.0693	**0.0779**	12.41%	0.0585	0.0454	0.0636	**0.0771**	21.23%
	N@10	0.0218	0.0152	0.0209	0.0214	0.0311	**0.0365**	17.36%	0.0285	0.0230	0.0306	**0.0345**	12.75%
	R@50	0.1322	0.1030	0.1151	0.1315	0.1827	**0.1875**	2.63%	0.1647	0.1141	0.1658	**0.1833**	10.55%
	N@50	0.0394	0.0305	0.0370	0.0400	0.0556	**0.0598**	7.55%	0.0523	0.0378	0.0527	**0.0571**	8.35%
Instruments	R@10	0.1118	0.0813	0.1070	0.1056	0.1267	**0.1300**	2.60%	0.1127	0.0783	0.1189	**0.1201**	1.01%
	N@10	0.0612	0.0620	0.0796	0.0713	0.0748	**0.0843**	5.90%	0.0661	0.0497	0.0680	**0.0771**	13.38%
	R@50	0.2106	0.1454	0.1890	0.1927	**0.2387**	0.2370	-	0.2104	0.1387	**0.2255**	0.2218	-
	N@50	0.0826	0.0756	0.0972	0.0901	0.0991	**0.1071**	8.07%	0.0873	0.0627	0.0912	**0.0988**	8.33%
Arts	R@10	0.1108	0.0722	0.1002	0.1003	0.1239	**0.1314**	6.05%	0.0977	0.0664	0.1066	**0.1119**	4.97%
	N@10	0.0587	0.0479	0.0714	0.0601	0.0712	**0.0767**	7.42%	0.0562	0.0375	0.0586	**0.0625**	6.66%
	R@50	0.2030	0.1367	0.1779	0.1888	0.2347	**0.2410**	2.68%	0.1916	0.1323	0.2049	**0.2100**	2.49%
	N@50	0.0788	0.0619	0.0883	0.0793	0.0955	**0.1002**	4.92%	0.0766	0.0518	0.0799	**0.0836**	4.63%
Office	R@10	0.1056	0.0825	0.1118	0.1030	**0.1280**	0.1275	-	0.0929	0.0641	0.1013	**0.1038**	2.47%
	N@10	0.0710	0.0634	**0.0868**	0.0653	0.0831	0.0856	-	0.0582	0.0391	0.0619	**0.0666**	7.59%
	R@50	0.1627	0.1227	0.1665	0.1613	**0.2016**	0.2005	-	0.1580	0.1113	**0.1702**	0.1701	-
	N@50	0.0835	0.0721	0.0987	0.0780	0.0991	**0.1012**	2.12%	0.0723	0.0493	0.0769	**0.0808**	5.07%

➢ Influence of multi-modal information

		Scientific				Office			
Input Type	Variant	R@10	R@50	N@10	N@50	R@10	R@50	N@10	N@50
ID+T	UniSRec	0.1676	0.3200	0.0872	0.1208	0.1376	0.2152	0.0937	0.1104
	MISSRec	**0.1697**	**0.3249**	**0.0890**	**0.1227**	**0.1398**	**0.2195**	**0.0953**	**0.1122**
ID+V	UniSRec	0.1631	0.3112	0.0857	0.1188	0.1324	0.2071	0.0921	0.1086
	MISSRec	**0.1657**	**0.3165**	**0.0869**	**0.1198**	**0.1375**	**0.2152**	**0.0946**	**0.1113**
ID+T+V	UniSRec	0.1680	0.3209	0.0865	0.1194	0.1407	0.2203	0.0957	0.1133
	MISSRec	**0.1711**	**0.3265**	**0.0890**	**0.1227**	**0.1421**	**0.2223**	**0.0966**	**0.1138**
T	UniSRec	0.1602	0.3202	0.0806	0.1148	0.1183	0.1893	0.0773	0.0926
	MISSRec	**0.1619**	**0.3239**	**0.0814**	**0.1165**	**0.1186**	**0.1899**	**0.0787**	**0.0943**
V	UniSRec	0.1568	0.3135	0.0782	0.1118	0.1140	0.1823	0.0763	0.0910
	MISSRec	**0.1575**	**0.3144**	**0.0789**	**0.1118**	**0.1152**	**0.1848**	**0.0771**	**0.0923**
T+V	UniSRec	0.1609	0.3216	0.0808	0.1141	0.1205	0.1929	0.0792	0.0945
	MISSRec	**0.1633**	**0.3265**	**0.0812**	**0.1160**	**0.1209**	**0.1936**	**0.0798**	**0.0953**

(Data & Code)