One for All, All for One: Learning and Transferring User Embeddings for Cross-Domain Recommendation

Chenglin Li
University of Alberta
Edmonton, AB, Canada

Yuanzhen Xie
Platform and Content Group, Tencent
Shenzhen, China

Chenyun Yu
Sun Yat-sen University
Shenzhen, China

Bo Hu
Platform and Content Group, Tencent
Shenzhen, China

Zang Li
Platform and Content Group, Tencent
Shenzhen, China

Guoqiang Shu
Platform and Content Group, Tencent
Shenzhen, China

Xiaohu Qie
Platform and Content Group, Tencent
Shenzhen, China

Di Niu
University of Alberta
Edmonton, AB, Canada

https://github.com/Chain123/CAT-ART

——WSDM 2023

gesis
Leibniz-Institut
für Sozialwissenschaften

Forschungszentrum · Research Center
L3S

Reported by Yabo Yin

1.Introduction

2.Method

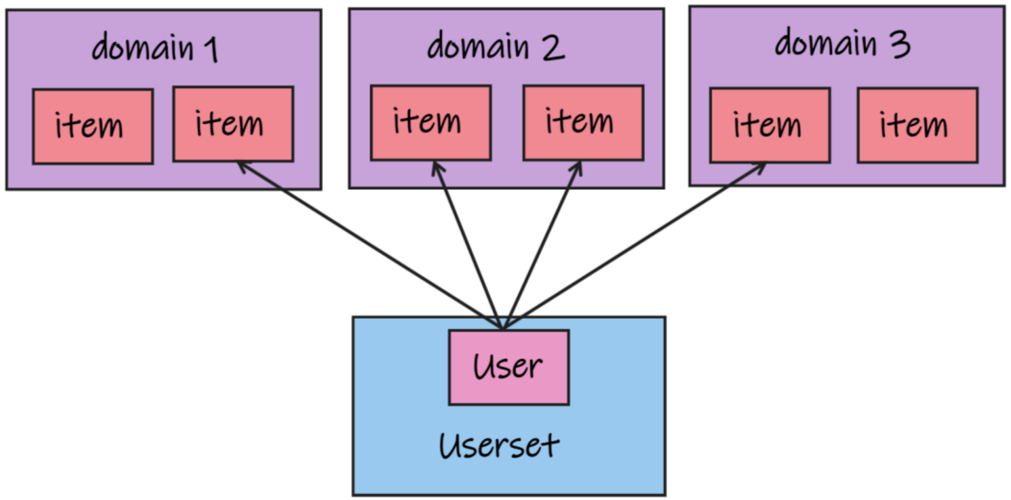
3.Experiments

gesis
Leibniz-Institut
für Sozialwissenschaften

Introduction

1. The cross-domain user representation, extracted directly on the collected data from all domains, may be severely biased by the domains with richer data and may fail to model the user preferences in sparse domains.

2. Directly transferring these features may cause performance degradation in the target domain due to various reasons, e.g., low-quality embeddings transferred from irrelevant domains.



Method

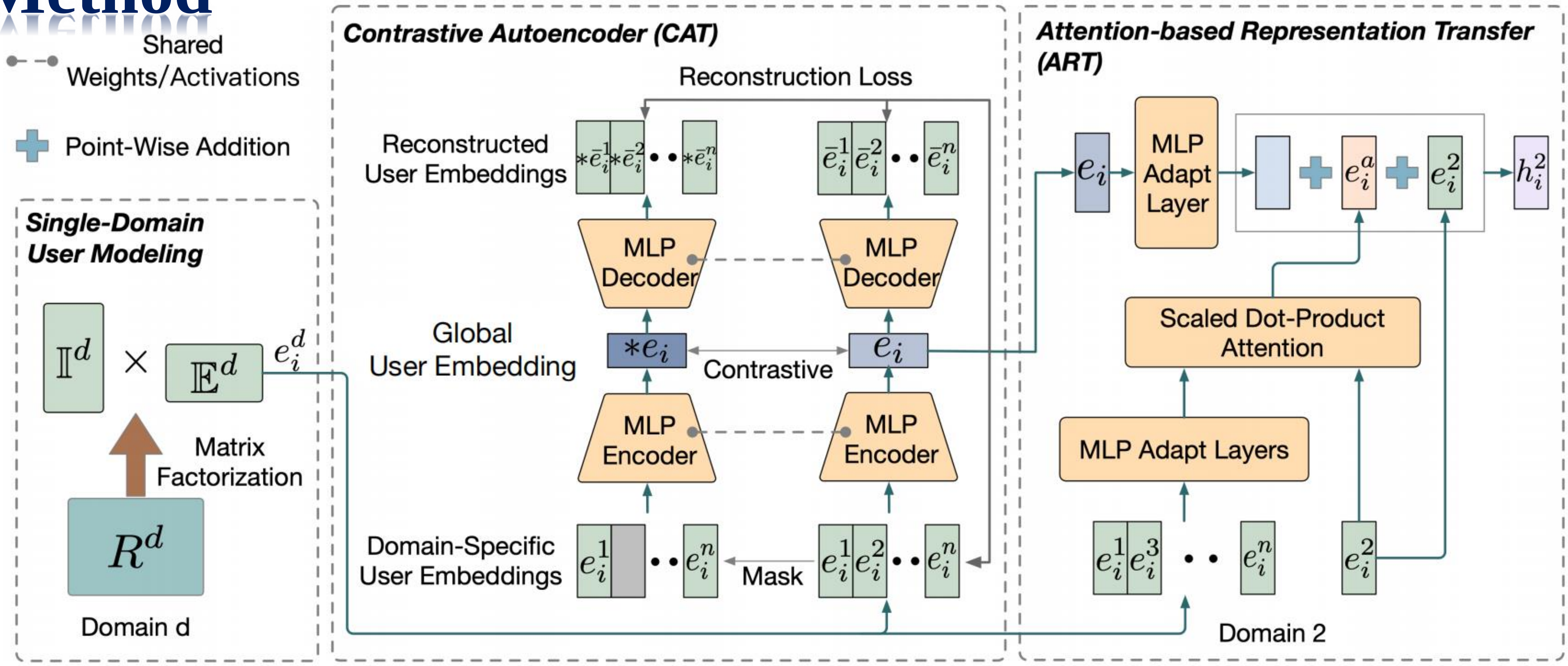


Figure 1: The architecture of the CAT-ART model. The CAT module takes domain-specific user embeddings as input and generates global user representation in a self-supervised manner. Then, the global user embedding e_i and the domain-specific embeddings from all the other domains are transferred to a target domain, e.g., domain 2, for boosted recommendations.

Method

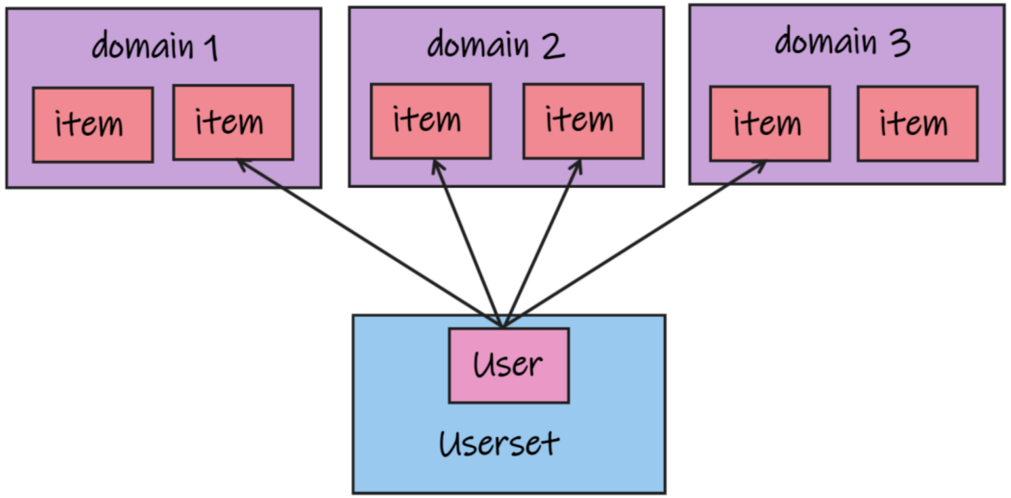
Preliminaries

a global user set U, and item sets $\{V_1, \cdots V_n\}$ in $n \geq 3$ domains.

matrix R^d shape $|U| \times |V_d|$

for user i and item j in domain d, $r_{ij}^d \in [0, 1]$,

Our goal is to improve the recommendation accuracy in all n domains simultaneously based on the interaction matrices.



Method

Domain-specific User Embedding

$$r_{ij}^d = \boldsymbol{e}_i^d \boldsymbol{I}_j^d .$$

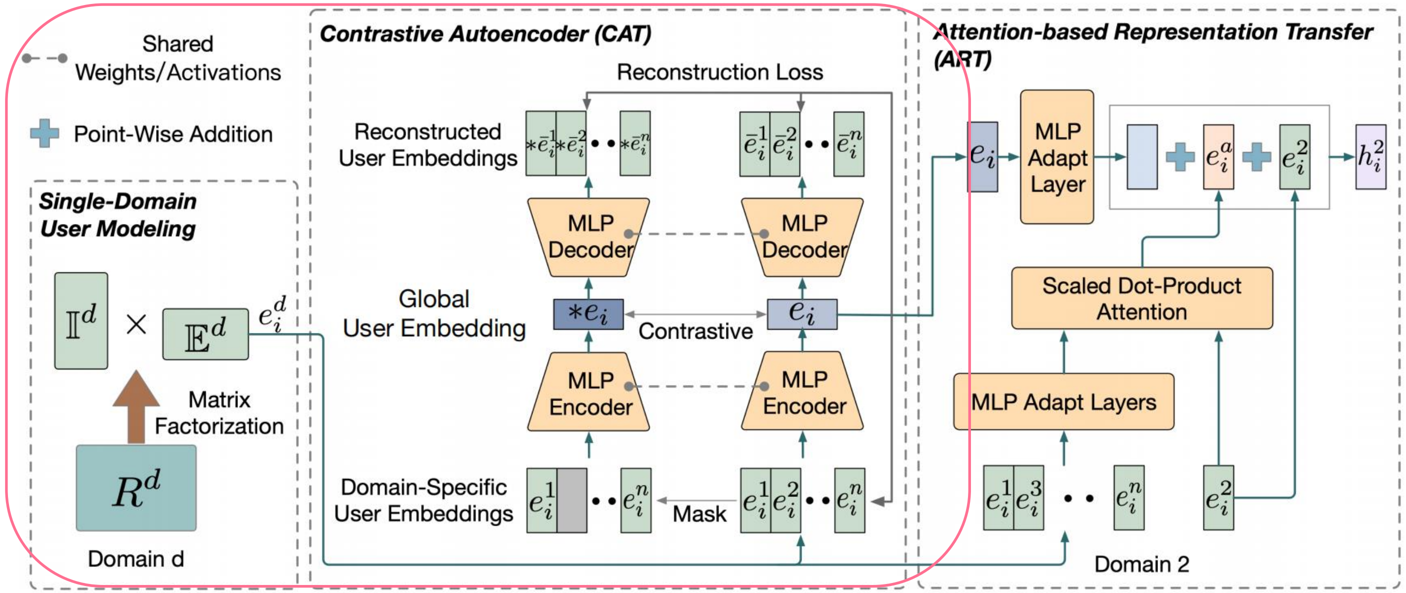
$$\mathcal{L}_{bpr}^d = -\sum_{i\in U}\sum_{j\in p_i^d}\sum_{l\notin p_i^d} \log\sigma(r_{ij}^d - r_{il}^d), \quad (1)$$

Contrastive Autoencoder

$$\boldsymbol{e}_i = \mathrm{MLP}_{\mathrm{enc}}(\boldsymbol{e}_i^1 \frown \boldsymbol{e}_i^2 \frown \cdots \frown \boldsymbol{e}_i^n)$$
$$\bar{\boldsymbol{e}}_i^1 \frown \bar{\boldsymbol{e}}_i^2 \frown \cdots \frown \bar{\boldsymbol{e}}_i^n = \mathrm{MLP}_{\mathrm{dec}}(\boldsymbol{e}_i), \quad (2)$$

$$\mathcal{L}_{rec} = \frac{1}{|U|}\sum_{i\in U}\sum_{d=1}^{n} ||\boldsymbol{e}_i^d - \bar{\boldsymbol{e}}_i^d||_2 . \quad (3)$$

$$\boldsymbol{e}_i^1 \frown \boldsymbol{e}_m \frown \cdots \frown \boldsymbol{e}_i^n = \mathrm{Mask}(\boldsymbol{e}_i^1 \frown \boldsymbol{e}_i^2 \frown \cdots \frown \boldsymbol{e}_i^n)$$
$$*\boldsymbol{e}_i = \mathrm{MLP}_{\mathrm{enc}}(\boldsymbol{e}_i^1 \frown \boldsymbol{e}_m \frown \cdots \frown \boldsymbol{e}_i^n) \quad (4)$$



$$\phi(\boldsymbol{e}_i, *\boldsymbol{e}_i) = \frac{\boldsymbol{e}_i * \boldsymbol{e}_i^T}{|\boldsymbol{e}_i||*\boldsymbol{e}_i|}$$

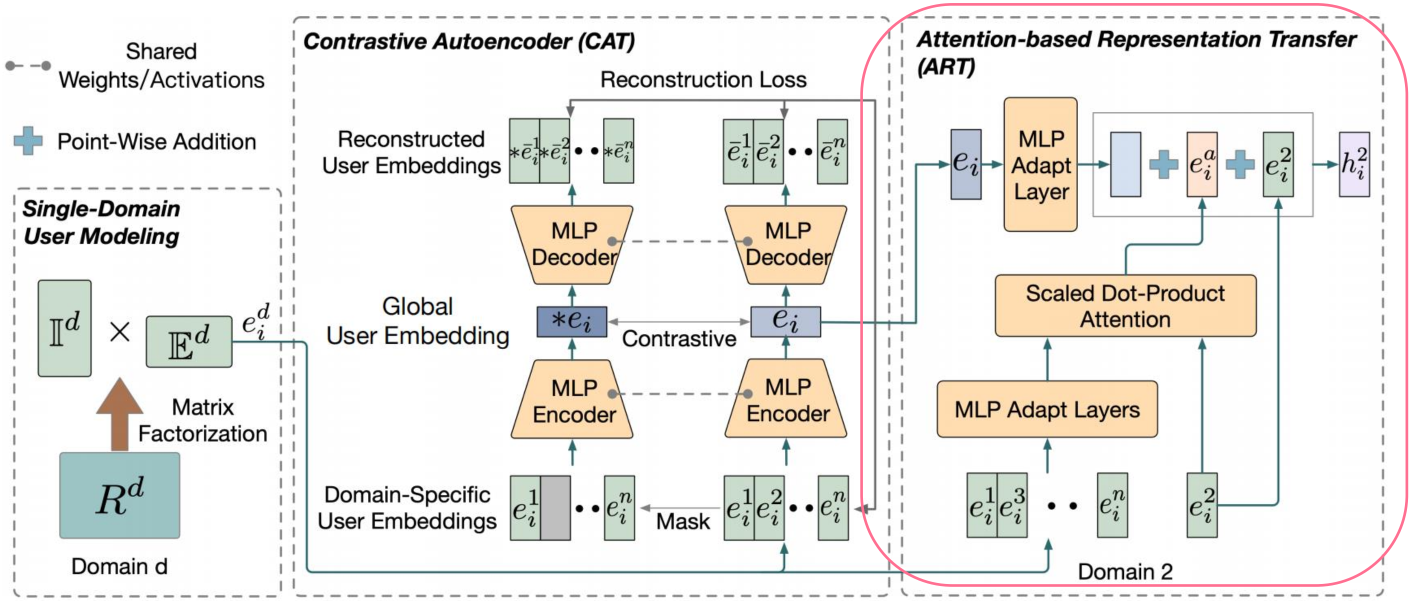
$$l_i = -log\frac{exp(\phi(\boldsymbol{e}_i, *\boldsymbol{e}_i)/\tau)}{\sum_{k=1}^{N} exp(\phi(\boldsymbol{e}_i, *\boldsymbol{e}_k)/\tau)} - log\frac{exp(\phi(*\boldsymbol{e}_i, \boldsymbol{e}_i)/\tau)}{\sum_{k=1}^{N} exp(\phi(*\boldsymbol{e}_i, \boldsymbol{e}_k)/\tau)}, \quad (5)$$

Method

$$*\bar{\boldsymbol{e}}_i^1 \frown * \bar{\boldsymbol{e}}_i^2 \frown \cdots \frown * \bar{\boldsymbol{e}}_i^n = \mathrm{MLP}_{\mathrm{dec}}(*\boldsymbol{e}_i),$$

$$\mathcal{L}_{rec}^* = \frac{1}{|U|} \sum_{i \in U} \sum_{d=1}^{n} ||e_i^d - *\bar{\boldsymbol{e}}_i^d||_2. \tag{6}$$

$$\mathcal{L}_{cat} = \alpha_1 \mathcal{L}_{rec} + \alpha_2 \mathcal{L}_{rec}^* + (1 - \alpha_1 - \alpha_2) \sum_{i=1}^{|U|} l_i, \tag{7}$$



Attention-based Representation Transfer

$$Q = \boldsymbol{e}_i^d$$

$$K = V = \mathrm{MLP}_{\mathrm{adapt}}(\{\boldsymbol{e}_i^k\}, k \neq d) \tag{8}$$

$$\boldsymbol{e}_i^a = Attention(Q, K, V) = \mathrm{softmax}(\frac{QK^T}{\sqrt{m}})V,$$

$$\boldsymbol{h}_i^d = \boldsymbol{e}_i^d + \mathrm{MLP}_{\mathrm{ind}}(\boldsymbol{e}_i) + \boldsymbol{e}_i^a, \tag{9}$$

Model Training

1. Train multiple single domain with eq(1)
2. Train CAT with eq(7)
3. Fixed single domain embedding and global embedding to train ART

Experiments

Table 1: Statistics of the Collected Dataset with 5 Domains.

Domain	#Users	#Items	#Interactions	Density(‰)
App-Ins	1,166,552	100,000	101,981,793	0.874
APP-Use		100,000	18,156,535	0.155
Articles		50,000	102,832,656	1.763
Video-S		50,000	74,911,020	1.284
Video-L		50,000	11,412,988	0.196

Experiments

Table 2: Results (in %) of the Proposed Method and Baselines. The ↓ represents negative transfer compared with SMF.

Model	Domain	Precision		Recall		NDCG	
		@10	@20	@10	@20	@10	@20
SMF	APP-Ins	33.82±0.70	25.46±0.88	21.51±0.39	31.91±1.22	32.56±0.43	32.53±0.89
	APP-Use	20.91±0.23	12.21±0.26	65.5±0.89	75±1.50	57.39±1.46	60.81±1.72
	Article	16.02±0.73	12.05±0.58	16.64±0.43	23.25±0.40	21.59±1.30	21.93±1.06
	Video-S	3.9±0.03	3.86±0.02	3.59±0.44	6.9±0.77	3.83±0.13	4.84±0.25
	Video-L	5.98±0.20	3.91±0.10	26.73±0.87	34.6±0.88	20.37±1.19	22.91±1.2
CMF	APP-Ins	33.57±0.37↓	25.19±0.37↓	21.8±0.19	32.05±0.43	32.39±0.29↓	32.45±0.27↓
	APP-Use	20.41±0.11↓	12.17±0.05↓	64.91±0.27↓	75.54±0.16	43.99±0.78↓	47.89±0.74↓
	Article	10.29±0.27↓	8.37±0.19↓	8.83±0.23↓	13.79±0.28↓	11.24±0.34↓	12.07±0.31↓
	Video-S	3.87±0.12	3.81±0.12↓	4.08±0.17	7.6±0.29	4.00±0.14	5.04±0.18
	Video-L	4.74±0.03↓	3.26±0.01↓	21.44±0.12↓	29.14±0.06↓	12.67±0.07↓	15.14±0.05↓
MPF	APP-Ins	36.08±1.53	27.11±0.41	23.28±0.99	34.29±0.44	36.95±5.56	36.53±4.02
	APP-Use	20.95±0.12	12.26±0.16	65.55±0.44	75.18±0.84	55.67±2.71↓	59.14±2.52↓
	Article	14.55±0.16↓	11.14±0.11↓	15.35±0.07↓	21.72±0.12↓	20.96±0.63↓	21.29±0.52↓
	Video-S	3.63±0.29↓	3.67±0.13↓	3.71±0.30	7.16±0.68	3.85±0.40	4.91±0.11
	Video-L	2.74±0.95↓	2.09±0.52↓	11.96±4.31↓	18.2±4.66↓	8.03±3.65↓	10.01±3.79↓
GA-MTCDR	APP-Ins	16.77±0.05↓	10.35±0.02↓	11.7±0.01↓	14.37±0.03↓	17.81±0.08↓	16.01±0.03↓
	APP-Use	13.88±0.05↓	10.46±0.01↓	45.44±0.13↓	67.2±0.16↓	32.35±0.13↓	40.16±0.1↓
	Article	4.62±0.13↓	3.73±0.03↓	4.12±0.14↓	6.37±0.11↓	6.22±0.18↓	6.36±0.13↓
	Video-S	3.44±0.03↓	3.1±0.02↓	3.48±0.08↓	6.03±0.06↓	4.22±0.05	4.69±0.04
	Video-L	3.18±0.15↓	2.22±0.07↓	14.21±0.74↓	19.76±0.54↓	10.46±0.63↓	12.23±0.49↓
HeroGRAPH-L	APP-Ins	34.05±2.01	24.47±1.16↓	22.34±1.14	31.61±1.35↓	40.5±1.91	38.12±1.51
	APP-Use	20.68±0.36↓	11.98±0.15↓	66.11±0.83	74.96±0.61↓	59.51±1.08	62.74±0.98
	Article	11.27±0.12↓	8.61±0.12↓	15.01±0.2↓	20.68±0.33↓	18.19±0.16↓	18.86±0.23↓
	Video-S	**3.99**±0.14	3.7±0.15	**5.29**±0.21	**8.97**±0.34	**5.31**±0.18	**6.2**±0.23
	Video-L	5.42±0.29↓	3.65±0.15↓	24.62±1.22↓	32.84±1.29↓	18.71±1.21↓	21.35±1.24↓
CAT-ART	APP-Ins	**38.36**±0.58	**27.96**±0.31	**24.86**±0.34	**35.46**±0.39	**43.47**±1.23	**41.55**±0.94
	APP-Use	**21.23**±0.18	**12.33**±0.18	**66.53**±0.65	**75.66**±1.02	**59.98**±0.86	**63.27**±1.02
	Article	**16.82**±0.21	**12.4**±0.13	**18.76**±0.56	**25.47**±0.6	**25.97**±0.61	**25.79**±0.58
	Video-S	3.93±0.08	**3.93**±0.06	3.83±0.50	7.35±0.82	3.93±0.14	5.05±0.24
	Video-L	**6.08**±0.09	**3.96**±0.08	**27.18**±0.39	**35.01**±0.67	**21.03**±0.38	**23.54**±0.86

Experiments

Table 3: Results (in %) of ablation studies. The ↓ represents negative transfer compared with the SMF model.

Domain	Metric	SMF	+Autoencoder	+Contrastive	+ART	-Attention
App-Ins	Precision@10	33.82±0.70	37.64±1.17	37.95±0.45	**38.36**±0.58	36.24±0.26
	Recall@10	21.51±0.39	24.35±0.76	24.58±0.35	**24.86**±0.34	23.35±0.23
	NDCG@10	32.56±0.43	41.34±3.75	42.56±2.02	**43.47**±1.23	36.08±1.54
APP-Use	Precision@10	20.91±0.23	21.00±0.11	21.08±0.23	**21.23**±0.18	21.01±0.07
	Recall@10	65.50±0.89	65.77±0.33	66.01±0.88	**66.53**±0.65	65.92±0.41
	NDCG@10	57.39±1.46	59.09±0.37	58.61±0.40	**59.98**±0.86	59.28±0.24
Article	Precision@10	16.02±0.73	16.54±0.46	16.46±0.34	**16.82**±0.21	15.88 ± 0.15↓
	Recall@10	16.64±0.43	17.48±1.21	17.19±1.13	**18.76**±0.56	15.89±0.38↓
	NDCG@10	21.59±1.30	23.98±2.28	23.54±2.75	**25.97**±0.61	22.25±1.71
Video-S	Precision@10	3.89±0.025	3.91±0.08	**3.97**±0.13	3.93±0.08	3.82±0.28↓
	Recall@10	3.59±0.44	3.71±0.40	3.72±0.37	**3.83**±0.50	3.46±0.25↓
	NDCG@10	3.83±0.13	3.87±0.08	3.91±0.05	**3.93**±0.14	3.73±0.18↓
Video-L	Precision@10	5.98±0.20	6.04±0.01	6.07±0.04	**6.08**±0.09	5.86±0.03↓
	Recall@10	26.73±0.87	27.00±0.08	27.17±0.20	**27.18**±0.39	26.27±0.09↓
	NDCG@10	20.37±1.19	21.00±0.14	21.12±0.21	**21.03**±0.38	20.26±0.15↓

- **SMF**: The single-domain Matrix Factorization (MF) model.
- **+Autoencoder**: We add the original autoencoder to extract global representations for CDR.
- **+Contrastive**: We further add the contrastive loss for the training of the autoencoder, i.e., the CAT module.
- **+ART**: The ART module is further incorporated to integrate domain-specific user embedding.
- **-Attention**: We remove the attention from the ART and only use MLP layers to integrate domain-specific features.

Experiments

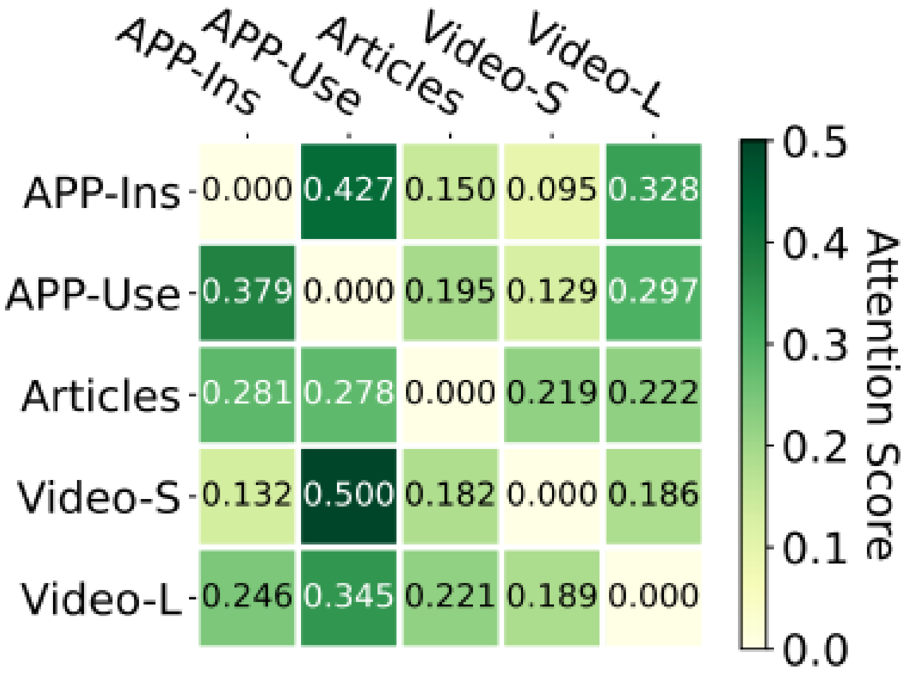


Figure 2: Averaged attention scores on the test set.

Thank you!