Sequential Recommendation with Probabilistic Logical Reasoning

Huanhuan Yuan[1], **Pengpeng Zhao**[1*], **Xuefeng Xian**[2*] and **Guanfeng Liu**[3] and **Victor S. Sheng**[4] and **Lei Zhao**[1]

[1]Soochow University

[2]Suzhou Vocational University

[3] Macquarie University

[4] Texas Tech University

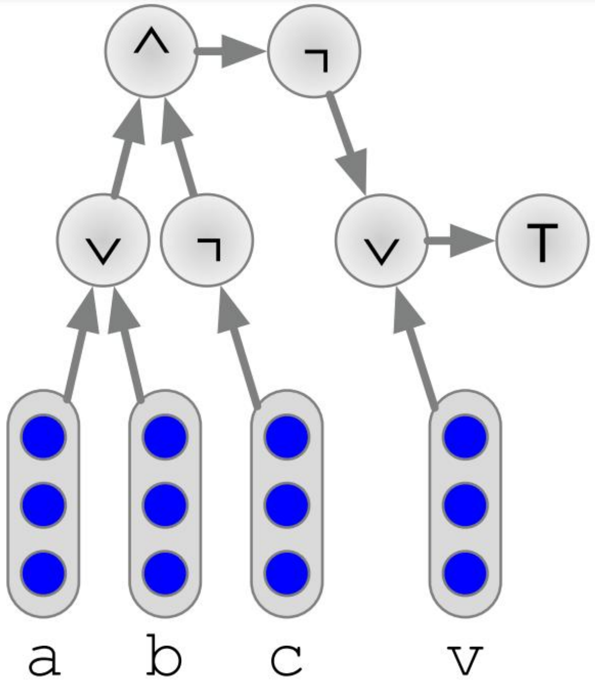
hhyuan@stu.suda.edu.cn, ppzhao@suda.edu.cn, xianxuefeng@jssvc.edu.cn, guanfeng.liu@mq.edu.au, Victor.Sheng@ttu.edu, zhaol@suda.edu.cn

code: https://github.com/Huanhuaneryuan/SR-PLR.

IJCAI 2023

Reported by Minqin Li

Introduction



u: $(a \vee b) \wedge \neg c \rightarrow v$ => $\neg((a \vee b) \wedge \neg c) \vee v$

蕴涵等值式 $A \rightarrow B \Leftrightarrow \neg A \vee B$

The most recent models for logical reasoning are embedding-based. The feature description and logical representation are coupled in the same framework, which makes it hard to distinguish which latent feature contributes to feature representation or logical reasoning.

Most of them assume user preferences are static and embed users and items in a deterministic manner, but ignore that the user's tastes are full of uncertainty and evolving by nature, which incurs inaccurate recommendations.

Method

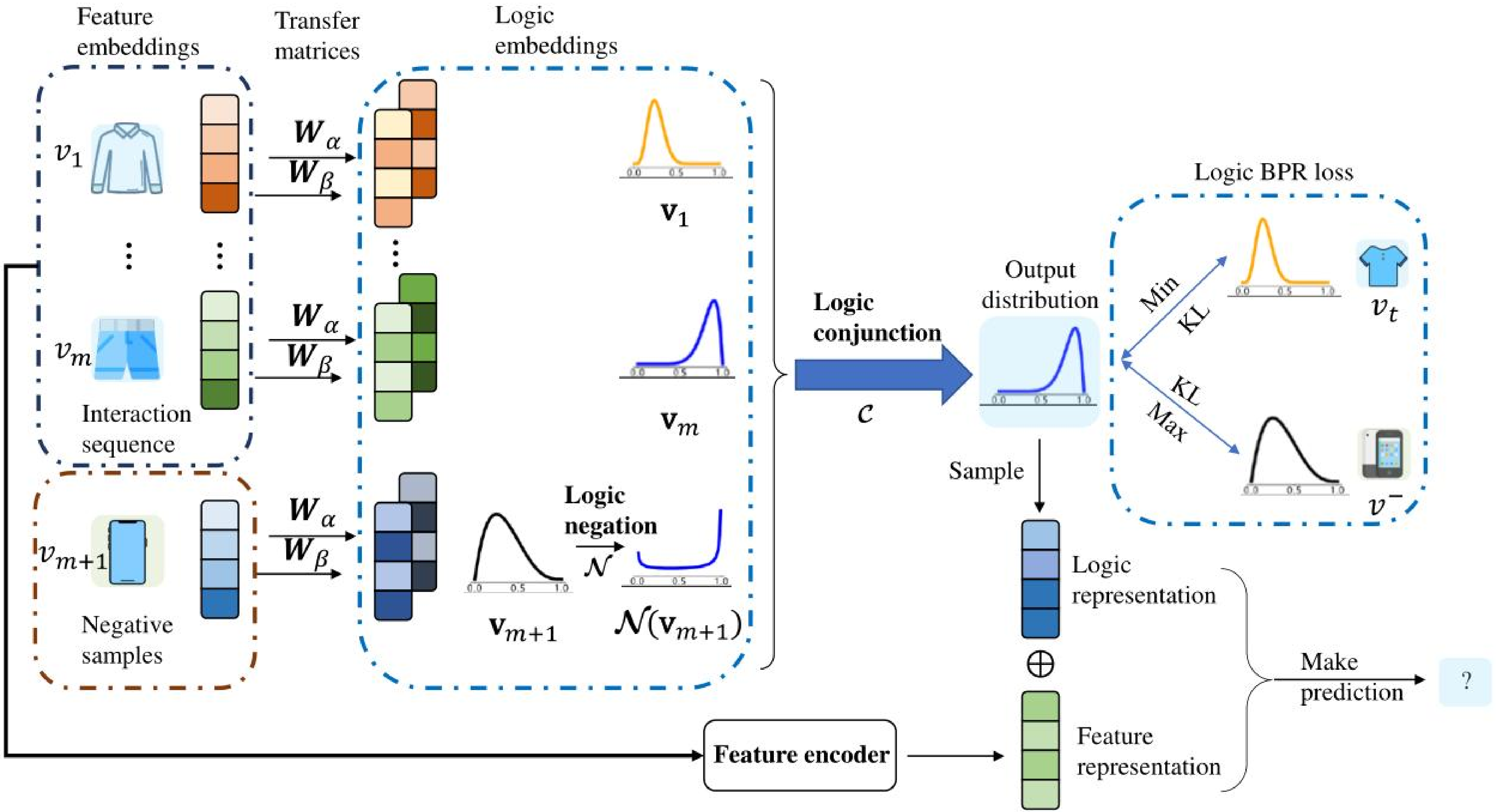
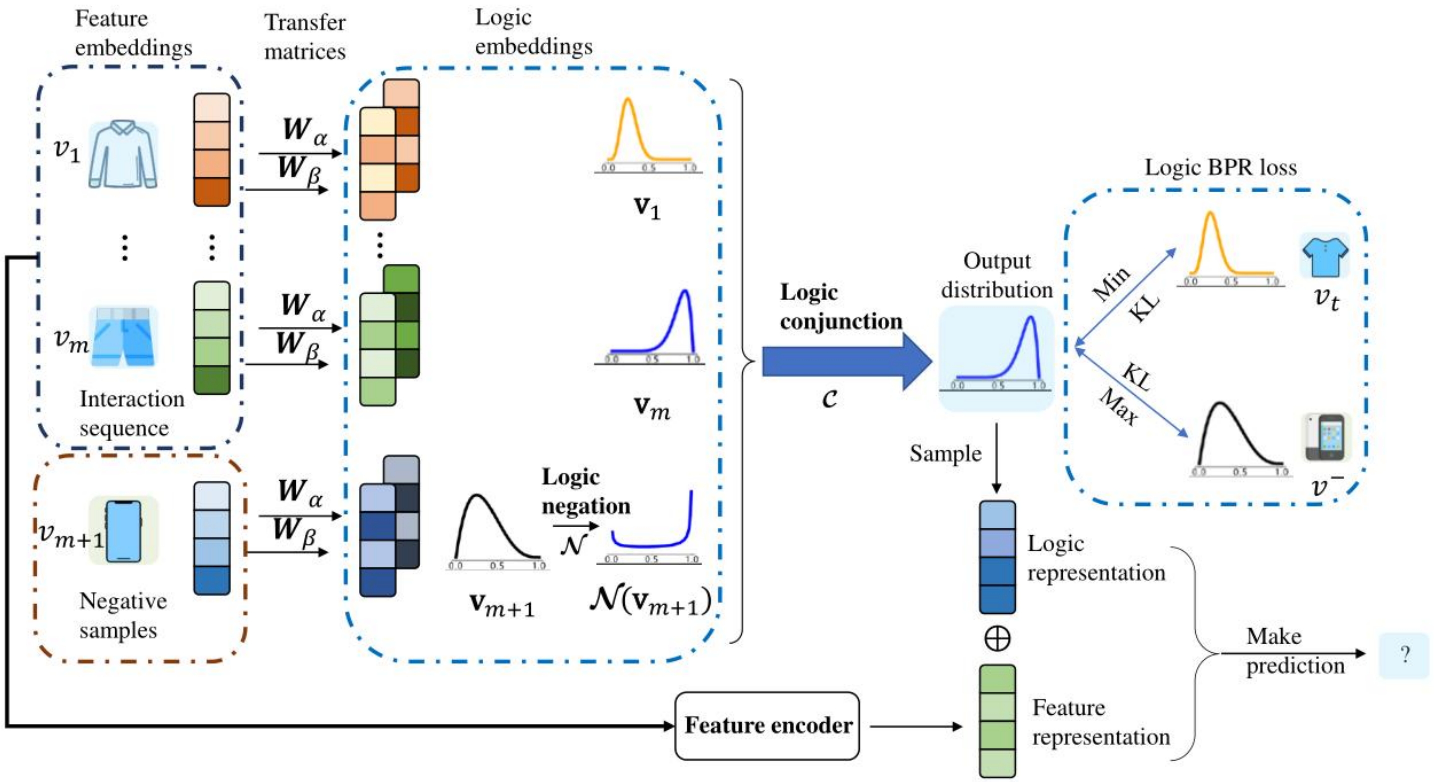


Figure 1: The framework of SR-PLR. $v_1, v_2, \cdots, v_m$ are u's historical items, v_t is a target item and $\mathbf{v}_i$ is their corresponding distributions. v_{m+1} and v^- represent the sampled items that the user does not interact with.

Method



$$\boldsymbol{\alpha} = \mathbf{M}\mathbf{W}_\alpha, \boldsymbol{\beta} = \mathbf{M}\mathbf{W}_\beta \tag{1}$$

$$\mathcal{N}(\mathbf{v}_i) = [(\frac{1}{\alpha_{i,1}}, \frac{1}{\beta_{i,1}}), (\frac{1}{\alpha_{i,2}}, \frac{1}{\beta_{i,2}}), \cdots, (\frac{1}{\alpha_{i,d}}, \frac{1}{\beta_{i,d}})] \tag{2}$$

$$\overline{\mathbf{v}} = [(\sum_{i=1}^{m} \mathbf{w}_i \odot \boldsymbol{\alpha}_i, \sum_{i=1}^{m} \mathbf{w}_i \odot \boldsymbol{\beta}_i)] \tag{3}$$

$$p_{\overline{\mathbf{v}}}(x) = \prod p_{\mathbf{v}_1}^{\mathbf{w}_1}(x) p_{\mathbf{v}_2}^{\mathbf{w}_2}(x) \cdots p_{\mathbf{v}_m}^{\mathbf{w}_m}(x) \tag{4}$$

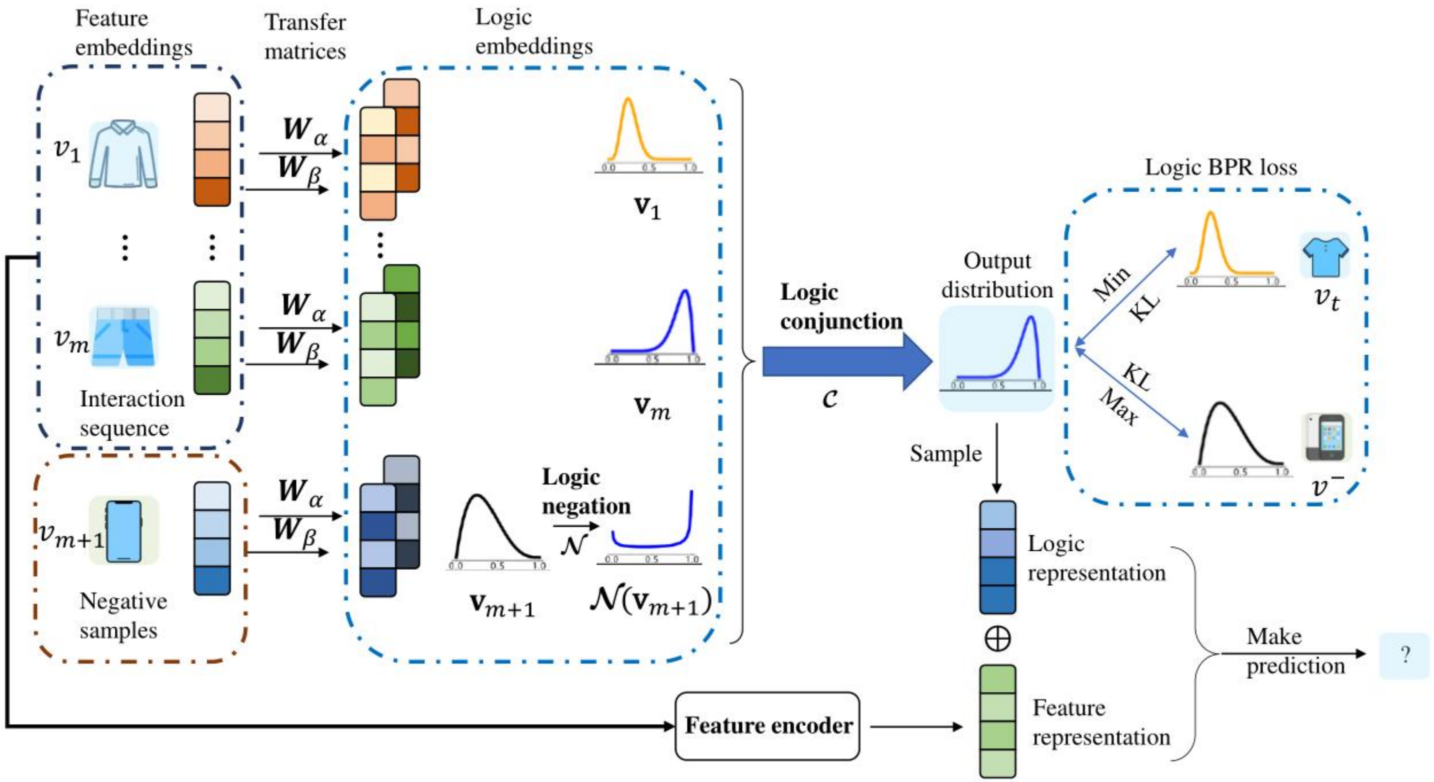
$$\mathbf{w}_i = \frac{\exp(MLP(\boldsymbol{\alpha}_i \oplus \boldsymbol{\beta}_i))}{\sum_j \exp(MLP(\boldsymbol{\alpha}_j \oplus \boldsymbol{\beta}_j))} \tag{5}$$

$$\overline{\mathbf{v}}_u = \mathcal{C}(\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_m, \mathcal{N}(\mathbf{v}_{m+1}), \cdots, \mathcal{N}(\mathbf{v}_n)\}) \tag{6}$$

$$\mathrm{Dist}(\overline{\mathbf{v}}_u, \mathbf{v}_t) = \sum_{k=1}^{d} \mathrm{KL}(\mathbf{P}_k(\mathbf{v}_t), \mathbf{P}_k(\overline{\mathbf{v}}_u)) \tag{7}$$

$$\mathcal{L}_l = \log(\sigma(\mathrm{Dist}(\overline{\mathbf{v}}_u, \mathbf{v}_t) - \mathrm{Dist}(\overline{\mathbf{v}}_u, \mathbf{v}^-))) \tag{8}$$

Method



$$\mathbf{H}_l^u = \frac{\overline{\boldsymbol{\alpha}}_u}{\overline{\boldsymbol{\alpha}}_u + \overline{\boldsymbol{\beta}}_u} \tag{9}$$

$$\hat{\mathbf{y}} = (\mathbf{H}_f^u \oplus \mathbf{H}_l^u)(\mathbf{M} \oplus \mathbf{E})^\top \tag{10}$$

$$\mathcal{L}_{Rec} = -\sum_{i=1}^{|\mathcal{V}|} y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i) \tag{11}$$

$$\mathcal{L} = \mathcal{L}_{Rec} + \lambda \mathcal{L}_l \tag{12}$$

Experiments

Datasets	Users	Items	Ratings	Avg. Len.	Sparsity
Sports	35,598	18,357	296,337	8.3	99.95%
Toys	19,413	11,925	167,597	8.6	99.93%
Yelp	30,499	20,068	317,182	10.4	99.95%

Table 1: Statistics of datasets.

Experiments

Models		Amazon Sports				Amazon Toys				Yelp			
		HIT@5	HIT@10	N@5	N@10	HIT@5	HIT@10	N@5	N@10	HIT@5	HIT@10	N@5	N@10
RNN	GRU4Rec	0.0186	0.0300	0.0121	0.0158	0.0352	0.0519	0.0240	0.0294	0.0217	0.0360	0.0144	0.0189
	GRU4Rec_L	**0.0225**	**0.0353**	**0.0141**	**0.0182**	**0.0387**	**0.0557**	**0.0268**	**0.0323**	**0.0249**	**0.0433**	**0.0160**	**0.0220**
	Impro.	20.97%	17.67%	16.53%	15.19%	9.94%	7.32%	11.67 %	9.86%	14.74%	23.06%	11.11%	16.40%
CNN	Caser	0.0141	0.0226	0.0087	0.0115	0.0183	0.0302	0.0113	0.0151	0.0231	0.0351	0.0164	0.0202
	Caser_L	**0.0173**	**0.0283**	**0.0109**	**0.0144**	**0.0228**	**0.0390**	**0.0132**	**0.0185**	**0.0260**	**0.0372**	**0.0185**	**0.0221**
	Impro.	22.70%	25.22%	25.29%	25.27%	24.59%	22.56%	16.81%	22.52 %	12.55%	5.98%	12.80%	9.41%
Attention	SASRec	0.0317	0.0484	0.0172	0.0226	0.0630	0.0909	0.0354	0.0444	0.0422	0.0595	0.0322	0.0377
	SASRec_L	**0.0332**	**0.0515**	**0.0192**	**0.0252**	**0.0632**	**0.0919**	**0.0359**	**0.0452**	**0.0441**	**0.0627**	**0.0326**	**0.0386**
	Impro.	4.73%	6.40%	11.63%	11.50%	0.32 %	1.10%	1.41%	1.80%	4.50%	5.38%	1.24%	2.39%
DuoRec	DuoRec	0.0328	0.0505	0.0192	0.0249	0.0648	0.0929	**0.0388**	0.0479	0.0434	0.0618	0.0319	0.0378
	DuoRec_L	**0.0342**	**0.0522**	**0.0200**	**0.0257**	**0.0652**	**0.0946**	**0.0388**	**0.0484**	**0.0441**	**0.0624**	**0.0321**	**0.0380**
	Impro.	4.27%	3.37%	4.17%	3.21%	0.62%	1.83%	0%	1.04%	1.61%	0.97%	0.63%	0.53%
Logic	LINN	0.0151	0.0256	0.0101	0.0132	0.0196	0.0320	0.0133	0.0172	0.0215	0.0355	0.0146	0.0192
	NCR-I	0.0162	0.0263	0.0110	0.0146	0.0201	0.0322	0.0135	0.0176	0.0204	0.0354	0.0146	0.0191

Table 2: Overall performance on all datasets. 'XX_L' means the SR-PLR method that uses 'XX' as the backbone and the numbers in bold indicate the better results that are generated by the same feature encoder. 'N' denotes 'NDCG' and 'Impro.' denotes performance improvement compared with backbones.

Experiments

		r	0.1	0.2	0.3	0.4	0.5
Sports	H@5	SASRec	0.0296	0.0292	0.0241	0.0214	0.0197
		SASRec_L	**0.0334**	**0.0310**	**0.0262**	**0.0215**	**0.0198**
	H@10	SASRec	0.0473	0.0462	0.0402	**0.0361**	**0.0327**
		SASRec_L	**0.0499**	**0.0484**	**0.0417**	0.0357	0.0320
	N@5	SASRec	0.0162	0.0158	0.0135	0.0122	0.0118
		SASRec_L	**0.0189**	**0.0185**	**0.0160**	**0.0136**	**0.0124**
	N@10	SASRec	0.0219	0.0213	0.0187	0.0169	0.0160
		SASRec_L	**0.0242**	**0.0241**	**0.0210**	**0.0182**	**0.0163**
Toys	H@5	SASRec	0.0582	0.0581	0.0504	0.0457	0.0384
		SASRec_L	**0.0626**	**0.0600**	**0.0582**	**0.0551**	**0.0446**
	H@10	SASRec	0.0880	0.0872	0.0779	0.0700	0.0594
		SASRec_L	**0.0890**	**0.0888**	**0.0879**	**0.0821**	**0.0733**
	N@5	SASRec	0.0329	0.0327	0.0285	0.0265	0.0228
		SASRec_L	**0.0359**	**0.0343**	**0.0337**	**0.0324**	**0.0271**
	N@10	SASRec	0.0426	0.0421	0.0374	0.0343	0.0295
		SASRec_L	**0.0444**	**0.0436**	**0.0433**	**0.0411**	**0.0363**

Table 3: Robustness analysis on Sports and Toys.

Experiments

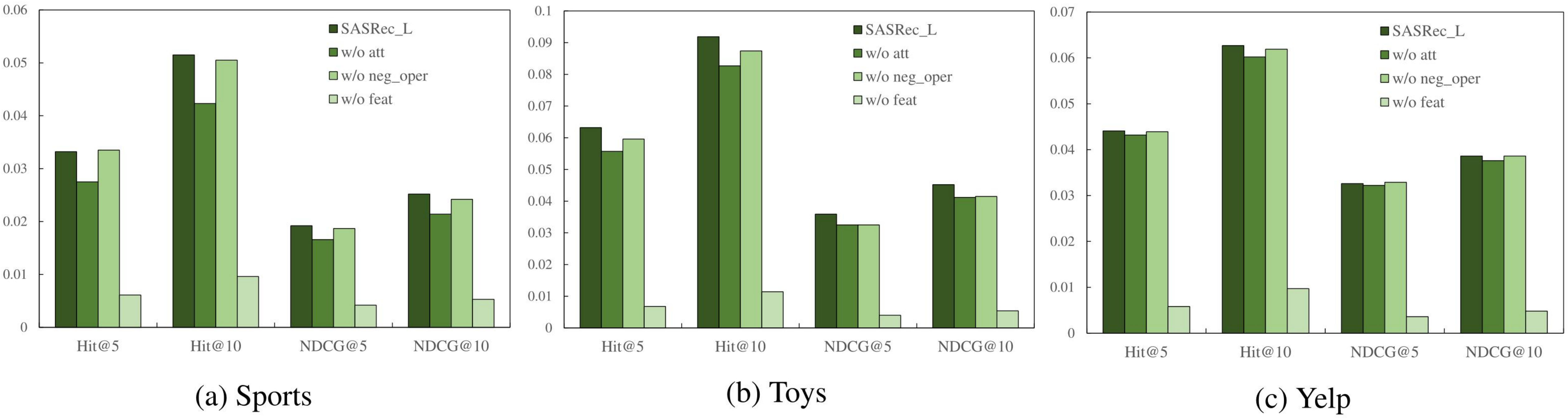


Figure 2: Ablation study of SR-PLR on three datasets.

Experiments

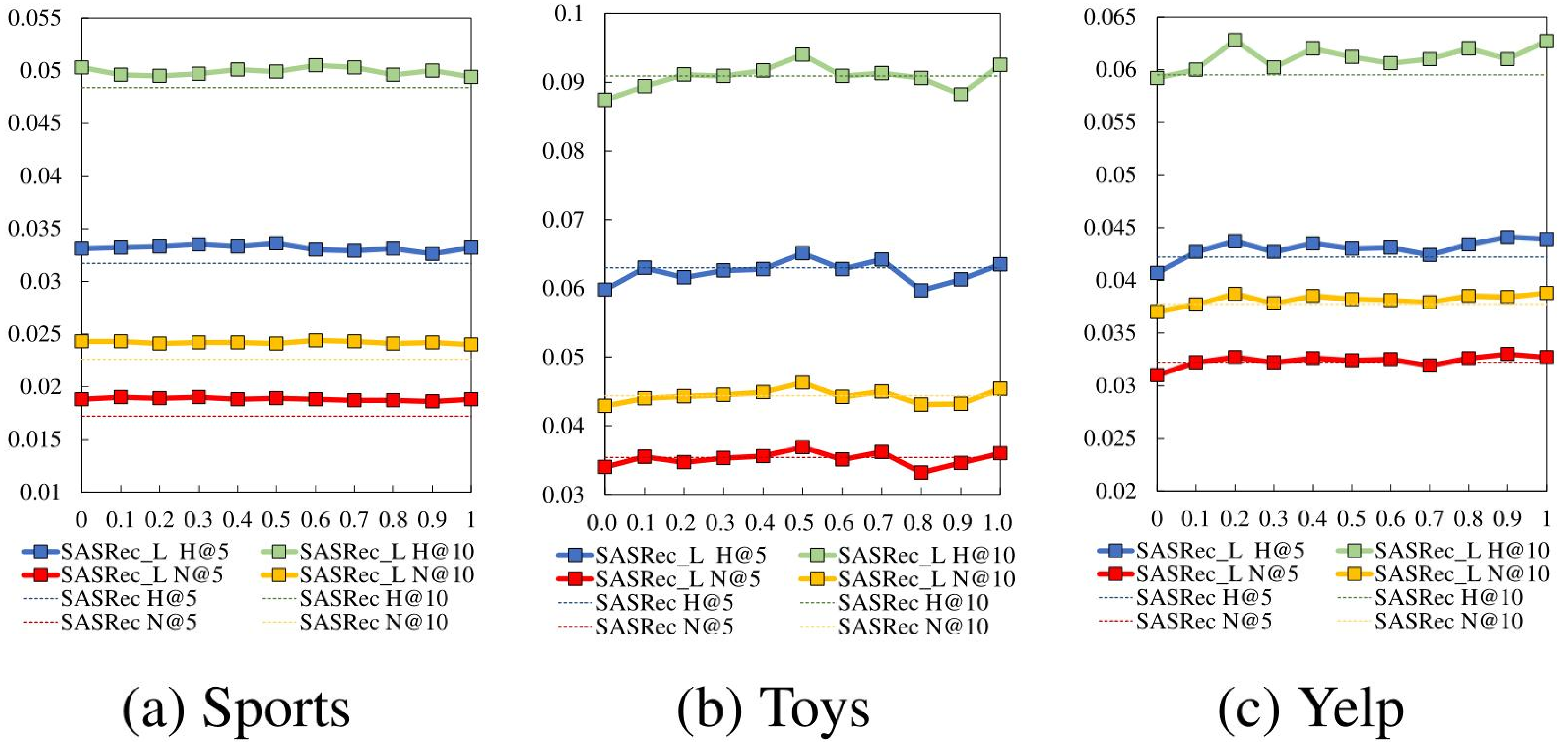
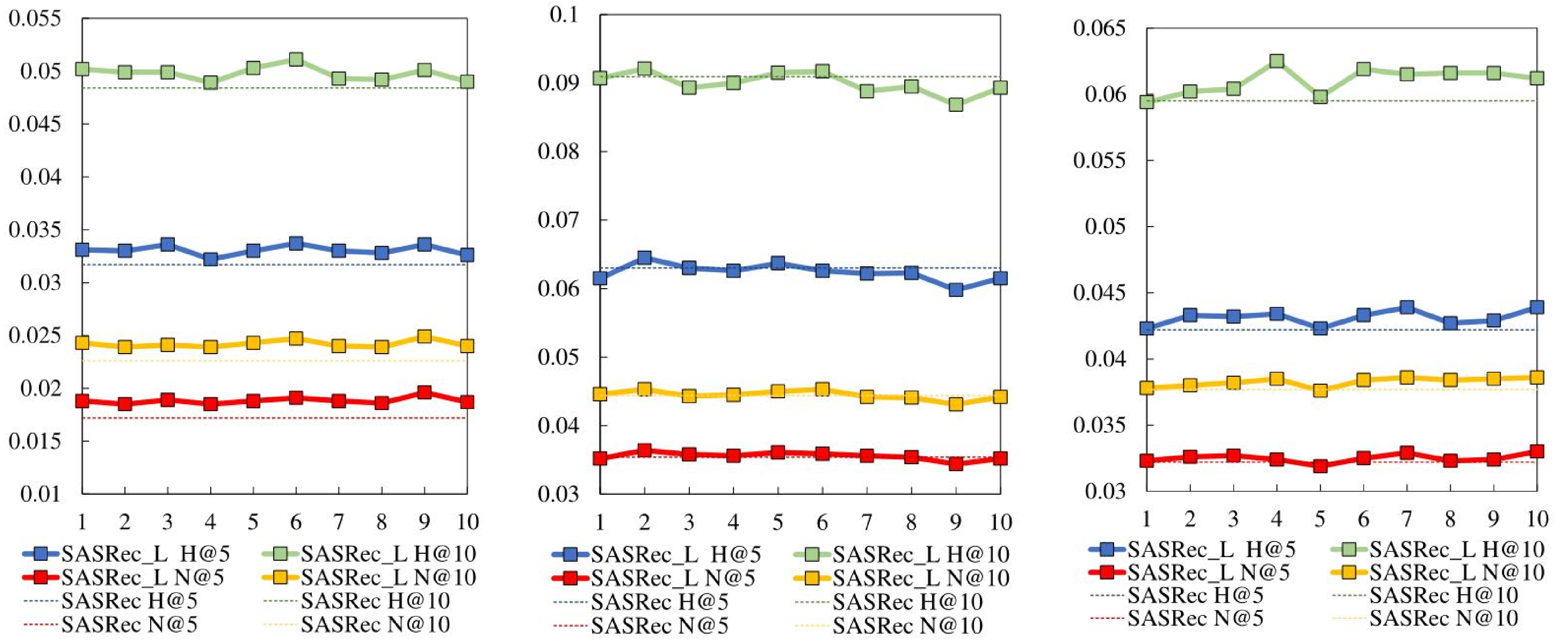


Figure 3: Sensitivity of λ on three datasets.

Experiments



(a) Sports (b) Toys (c) Yelp

Figure 4: Sensitivity of negative item number on three datasets.

Thanks