Explanation Guided Contrastive Learning for Sequential

Lei Wang, Ee-Peng Lim*
Singapore Management University
Singapore
{lei.wang.2019@phdcs.,eplim@}smu.edu.sg

Zhiwei Liu
Salesforce
USA
zhiweiliu@salesforce.com

Tianxiang Zhao
Penn State University
USA
tkz5084@psu.edu

code:https://github.com/demoleiwang/EC4SRec.

CIKM 2022

Reported by Minqin Li

Introduction

Figure 1: Motivation example: (a) A given user sequence with seven items and a red hair-dryer as the next item; (b) Two positive views generated by random mask operations on the given sequence and a negative view which is the sequence of another user. [M] represents a masked item.

When the positive and negative sequences turn out to be false positive and false negative respectively, it may lead to degraded recommendation performance. In this work, we address the above problem by proposing Explanation Guided Augmentations (EGA) and Explanation Guided Contrastive Learning for Sequential Recommendation (EC4SRec) model framework.

EC4SRec then combines both self-supervised and supervised contrastive learning over the positive and negative sequences generated by EGA operations to improve sequence representation learning for more accurate recommendation results.

Method

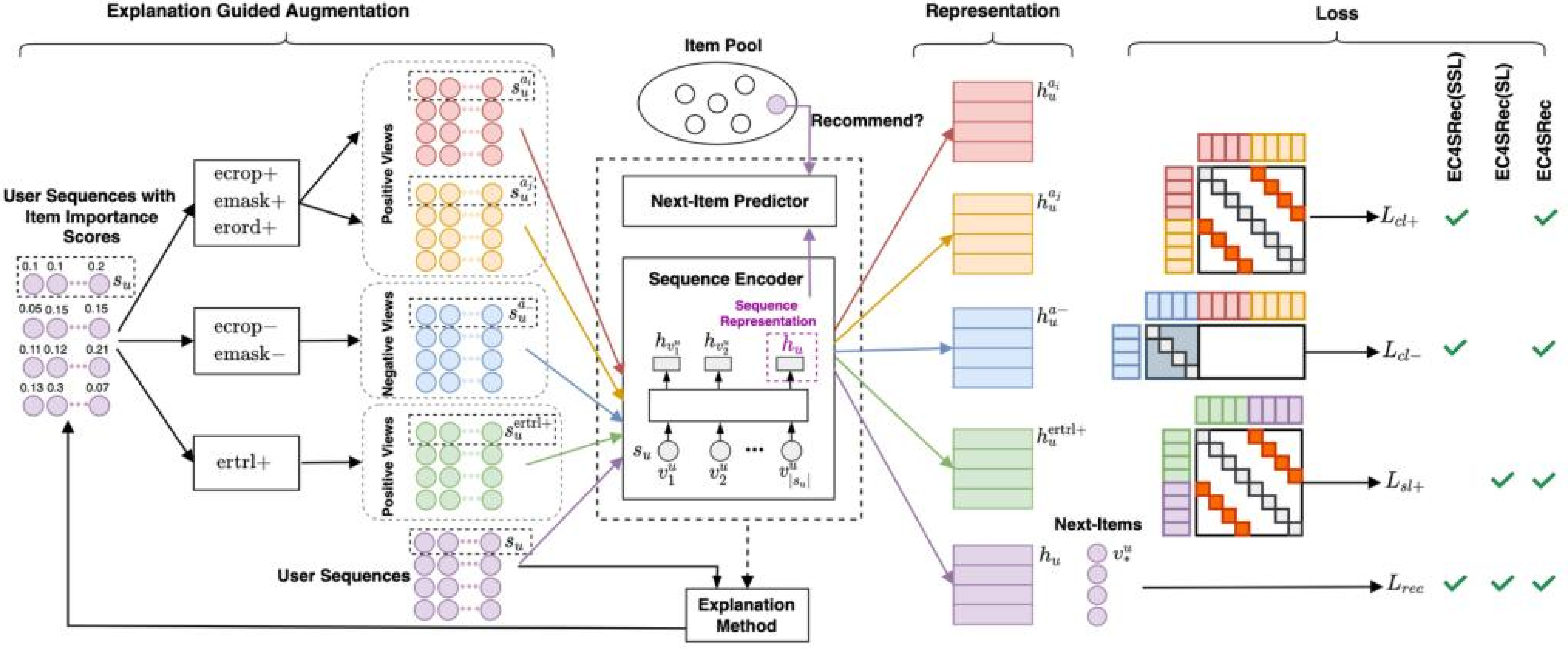

Figure 2: Proposed EC4SRec Framework.

Method

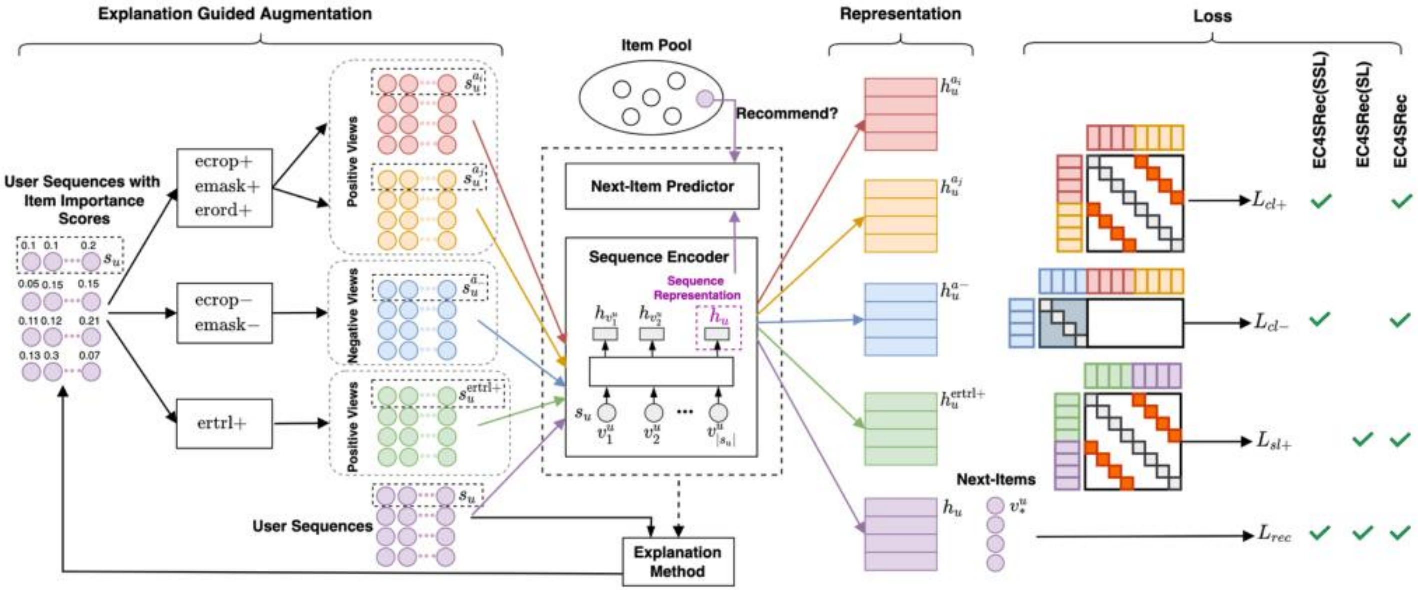

Figure 2: Proposed EC4SRec Framework.

$$\mathcal{L}_{rec}(s_u) = -\log \frac{\exp(sim(h_u, h_{v_*^u}))}{\exp(sim(h_u, h_{v_*^u})) + \sum_{v^- \in V^-} \exp(sim(h_u, h_{v^-}))} \quad (2)$$

$$\mathcal{L}_{CL4SRec} = \sum_{u \in U_B} \mathcal{L}_{rec}(s_u) + \lambda \mathcal{L}_{cl}(s_u^{a_i}, s_u^{a_j}) \quad (1)$$

$$\mathcal{L}_{cl}(s_u^{a_i}, s_u^{a_j}) = -\log \frac{\exp(sim(h_u^{a_i}, h_u^{a_j}))}{\exp(sim(h_u^{a_i}, h_u^{a_j})) + \sum_{s^- \in S_u^-} \exp(sim(h_u^{a_i}, h^-))} \quad (3)$$

Method

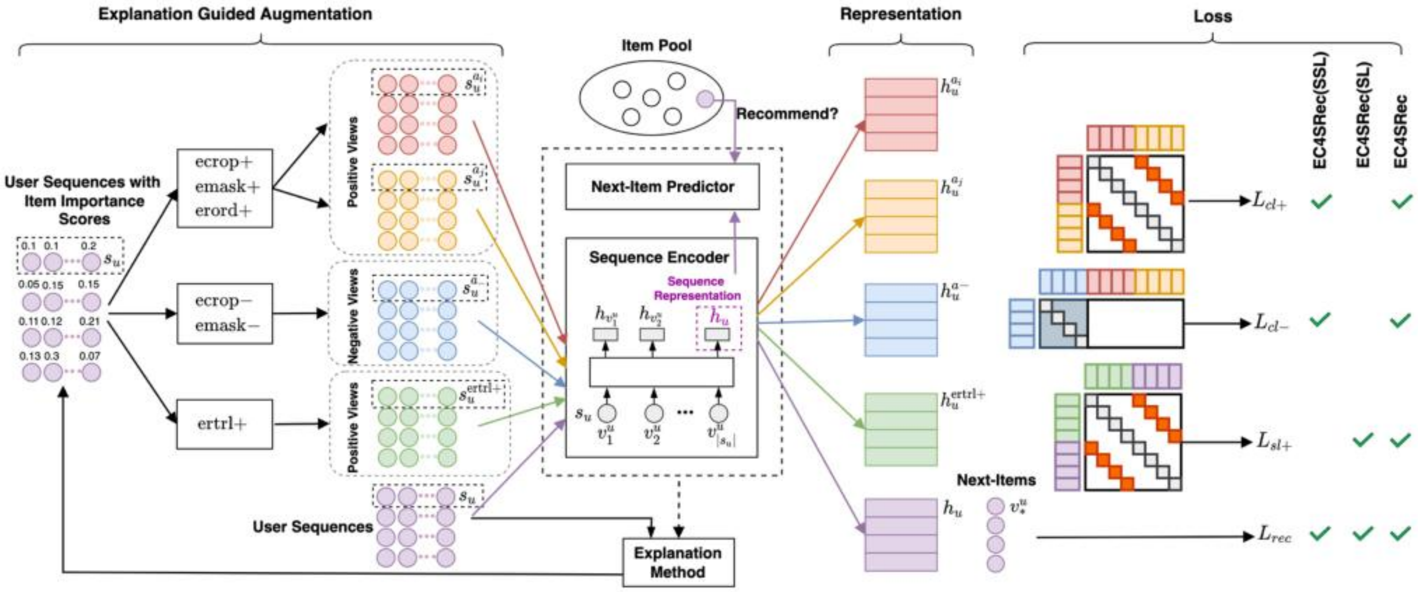

Figure 2: Proposed EC4SRec Framework.

$$\mathcal{L}_{DuoRec} = \sum_{u \in U_B} \mathcal{L}_{rec}(s_u) + \lambda \mathcal{L}_{sl}(s_u) \tag{4}$$

$$\mathcal{L}_{sl}(s_u) = -\Bigg(\log \frac{\exp(sim(h_u, h_u^{\mathrm{rtrl}})/\tau)}{\exp(sim(h_u, h_u^{\mathrm{rtrl}})/\tau) + \sum_{s^- \in S_u^-} \exp(sim(h_u, h^-)/\tau)} + \log \frac{\exp(sim(h_u^{\mathrm{rtrl}}, h_u)/\tau)}{\exp(sim(h_u^{\mathrm{rtrl}}, h_u)/\tau) + \sum_{s^- \in S_u^-} \exp(sim(h_u^{\mathrm{rtrl}}, h^-)/\tau)}\Bigg)$$

Method

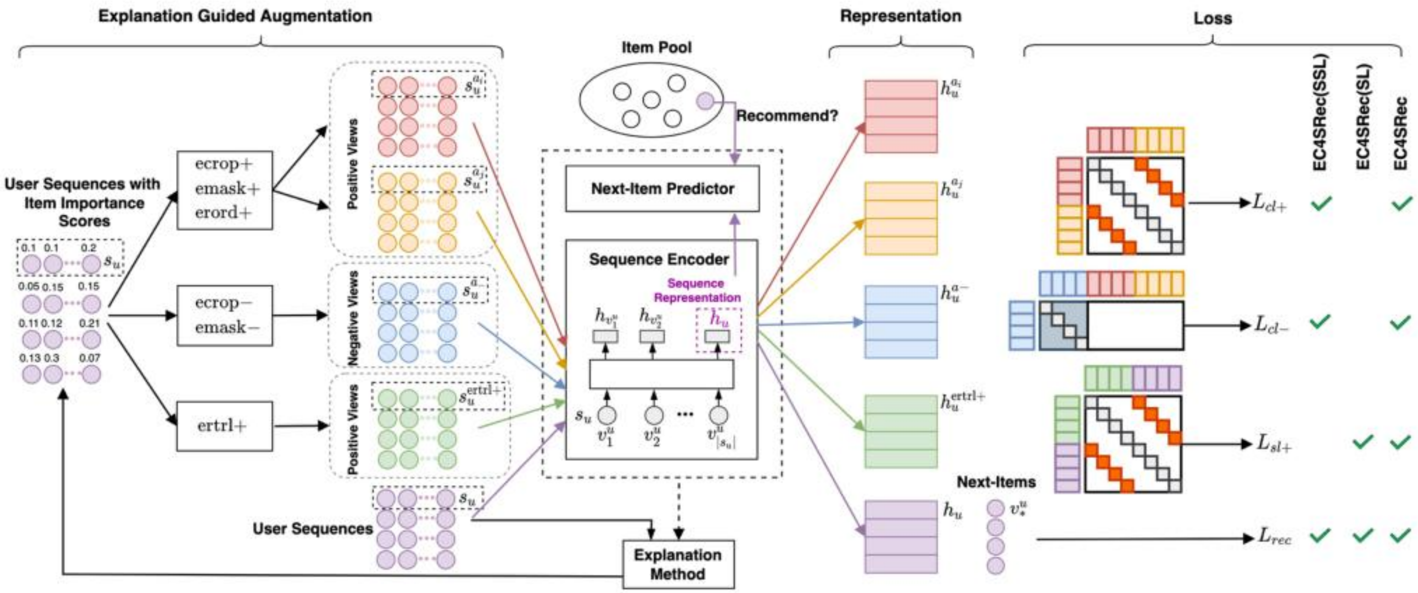

Figure 2: Proposed EC4SRec Framework.

$$score(v_i^u) = \frac{\sum_{j=1}^{d} score(e_{v_{i,j}^u})}{\sum_{i'=1}^{|s_u|} \sum_{j=1}^{d} score(e_{v_{i',j}^u})} \quad (5)$$

$$\mathcal{L}_{EC4SRec(SSL)} = \sum_{u \in U_B} \mathcal{L}_{rec}(s_u) + \lambda_{cl+}(\mathcal{L}_{cl+}(s_u) + \lambda_{cl-}\mathcal{L}_{cl-}(s_u)) \quad (6)$$

$$score(e_{v_{i,j}^u}) = \| \frac{\partial y_u}{\partial e_{v_{i,j}^u}} \|$$

Method

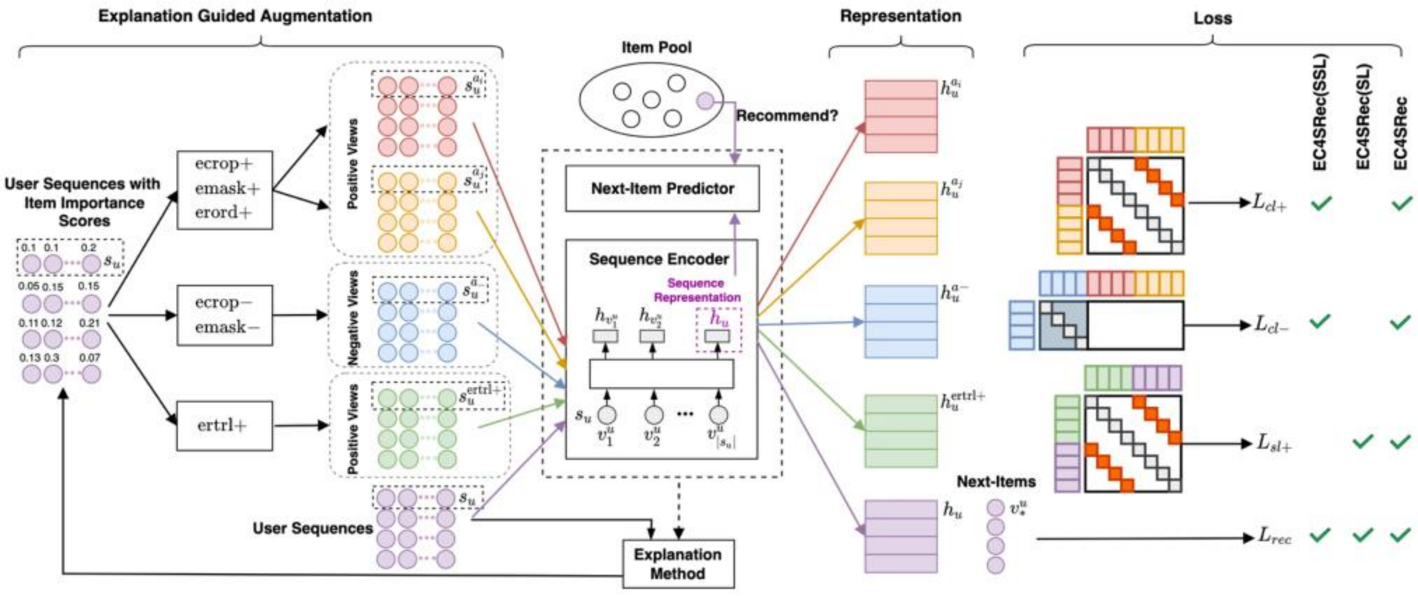

Figure 2: Proposed EC4SRec Framework.

$$\mathcal{L}_{EC4SRec(SL)} = \sum_{u \in U_B} \mathcal{L}_{rec}(s_u) + \lambda \mathcal{L}_{sl+}(s_u) \tag{7}$$

$$\mathcal{L}_{sl+}(s_u) = -\Bigg(\log \frac{\exp(sim(h_u, h_u^{ertrl+})/\tau)}{\exp(sim(h_u, h_u^{ertrl+})/\tau) + \sum_{s^- \in S_u^-} \exp(sim(h_u, h^-)/\tau)} + \log \frac{\exp(sim(h_u^{ertrl+}, h_u)/\tau)}{\exp(sim(h_u^{ertrl+}, h_u)/\tau) + \sum_{s^- \in S_u^-} \exp(sim(h_u^{ertrl+}, h^-)/\tau)} \Bigg) \tag{8}$$

$$\mathcal{L}_{EC4SRec} = \sum_{u \in U_B} \mathcal{L}_{rec}(s_u) + \lambda_{cl+}\mathcal{L}_{cl+}(s_u) + \lambda_{cl-}\mathcal{L}_{cl-}(s_u) + \lambda_{sl+}\mathcal{L}_{sl+}(s_u) \tag{9}$$

Experiments

Table 1: Results of CL4SRec on synthetic dataset with ground truth important items.

Masking Op.	Random	Oracle-based
HR@3	0.3560	0.5180
NDCG@3	0.2573	0.3645

Table 2: Dataset Statistics After Preprocessing.

Dataset	Beauty	Clothing	Sports	ML-1M
Users	22,363	39,387	35,598	6,041
Items	12,101	23,033	18,357	3,417
User-item Interactions	198,502	278,677	296,337	999,611
Avg Sequence Length	8.9	7.1	8.3	165.5
Sparsity	99.93%	99.97%	99.95%	95.16%

Experiments

Table 3: Overall Results. (The best and second best results are boldfaced and underlined. *: significant improvement of EC4SRec(SSL) over CL4SRec with p-value= 0.05. **: significant improvement of EC4SRec(SL) over DuoRec with p-value= 0.01.)

Dataset	Metric	Non-Seq. BPR-MF	Seq. Rec. w/o Contrastive Learning GRU4Rec	Caser	SASRec	BERT4Rec	S^3Rec_{MIP}	Seq. Rec. with Contrastive Learning CL4SRec	EC4SRec(SSL)**	DuoRec	EC4SRec(SL)*	**EC4SRec**
Beauty	HR@5	0.0120	0.0164	0.0191	0.0365	0.0193	0.0327	0.0495	0.0569	0.0548	**0.0585**	0.0569
	HR@10	0.0299	0.0365	0.0335	0.0627	0.0401	0.0591	0.0810	0.0853	0.0832	**0.0867**	0.0862
	NDCG@5	0.0065	0.0086	0.0114	0.0236	0.0187	0.0175	0.0299	0.0358	0.0345	0.0361	**0.0364**
	NDCG@10	0.0122	0.0142	0.0160	0.0281	0.0254	0.0268	0.0401	0.0450	0.0436	0.0455	**0.0458**
Clothing	HR@5	0.0067	0.0095	0.0049	0.0168	0.0125	0.0163	0.0187	0.0201	0.0196	0.0205	**0.0209**
	HR@10	0.0094	0.0165	0.0092	0.0272	0.0208	0.0237	0.0305	0.0314	0.0296	0.0311	**0.0320**
	NDCG@5	0.0052	0.0061	0.0029	0.0091	0.0075	0.0101	0.0104	0.0113	0.0112	0.0115	**0.0119**
	NDCG@10	0.0069	0.0083	0.0043	0.0124	0.0102	0.0132	0.0142	0.0149	0.0144	0.0149	**0.0155**
Sports	HR@5	0.0092	0.0137	0.0121	0.0218	0.0176	0.0157	0.0277	0.0323	0.0310	0.0317	**0.0331**
	HR@10	0.0188	0.0274	0.0204	0.0336	0.0326	0.0265	0.0455	0.0497	0.0480	0.0491	**0.0514**
	NDCG@5	0.0053	0.0096	0.0076	0.0127	0.0105	0.0098	0.0167	0.0201	0.0191	0.0194	**0.0203**
	NDCG@10	0.0085	0.0137	0.0103	0.0169	0.0153	0.0135	0.0224	0.0256	0.0246	0.0249	**0.0262**
ML-1M	HR@5	0.0164	0.0763	0.0816	0.1087	0.0733	0.1078	0.1583	**0.1699**	0.1672	0.1682	0.1672
	HR@10	0.0354	0.1658	0.1593	0.1904	0.1323	0.1952	0.2423	**0.2543**	0.2507	0.2526	0.2533
	NDCG@5	0.0097	0.0385	0.0372	0.0638	0.0432	0.0616	0.0996	0.1095	0.1076	**0.1104**	0.1102
	NDCG@10	0.0158	0.0671	0.0624	0.0910	0.0619	0.0917	0.1267	0.1368	0.1345	0.1375	**0.1380**

Experiments

Table 4: Results of EC4SRec with different Sequential Recommendation Backbones.

Backbone		Beauty				Clothing				Sports			
		HR@5	HR@10	NDCG@5	NDCG@10	HR@5	HR@10	NDCG@5	NDCG@10	HR@5	HR@10	NDCG@5	NDCG@10
GRU4Rec	CL4SRec	0.0420	0.0640	0.0270	0.0341	0.0104	0.0180	0.0065	0.0089	0.0244	0.0389	0.0154	0.0200
	EC4SRec(SSL)	0.0461	0.0674	0.0314	0.0382	0.0128	0.0213	0.0082	0.0109	0.0253	0.0396	0.0167	0.0213
	DuoRec	0.0471	0.0689	0.0318	0.0388	0.0118	0.0193	0.0078	0.0102	0.0259	0.0396	0.0163	0.0207
	EC4SRec(SL)	0.0490	0.0717	0.0327	0.0401	0.0130	0.0201	0.0086	0.0108	0.0273	0.0414	0.0173	0.0218
	EC4SRec	**0.0495**	**0.0745**	**0.0332**	**0.0412**	**0.0138**	**0.0218**	**0.0089**	**0.0115**	**0.0276**	**0.0437**	**0.0182**	**0.0233**
Caser	CL4SRec	0.0185	0.0335	0.0108	0.0157	0.0058	0.0100	0.0036	0.0049	0.0113	0.0191	0.0071	0.0096
	EC4SRec(SSL)	0.0228	0.0390	0.0137	0.0189	0.0064	0.0113	0.0039	0.0055	0.014	0.0244	0.0088	0.0121
	DuoRec	0.0207	0.0375	0.0129	0.0183	0.0053	0.0100	0.0031	0.0046	0.0127	0.0215	0.0082	0.0110
	EC4SRec(SL)	0.0262	0.0439	0.0161	0.0218	0.0064	0.0117	0.0039	0.0056	0.0146	0.0240	0.0097	0.0127
	EC4SRec	**0.0269**	**0.0456**	**0.0172**	**0.0232**	**0.0065**	**0.0124**	**0.0041**	**0.0060**	**0.0152**	**0.0266**	**0.0105**	**0.0139**

Experiments

Table 5: NDCG@5 Results of EC4SRec(SSL), abbreviated by E(SSL), with the removal of augmentation operation on Beauty, Clothing and Sports.

	Beauty		Clothing		Sports	
	CL4SRec	E(SSL)	CL4SRec	E(SSL)	CL4SRec	E(SSL)
None	0.0299	0.0358	0.0104	0.0113	0.0167	0.0201
−rord	0.0307	0.0344	0.0103	0.0110	0.0169	0.0181
−mask	0.0311	0.0350	0.0101	0.0116	0.0169	0.0200
−crop	0.0282	0.0353	0.0086	0.0116	0.0147	0.0200

Experiments

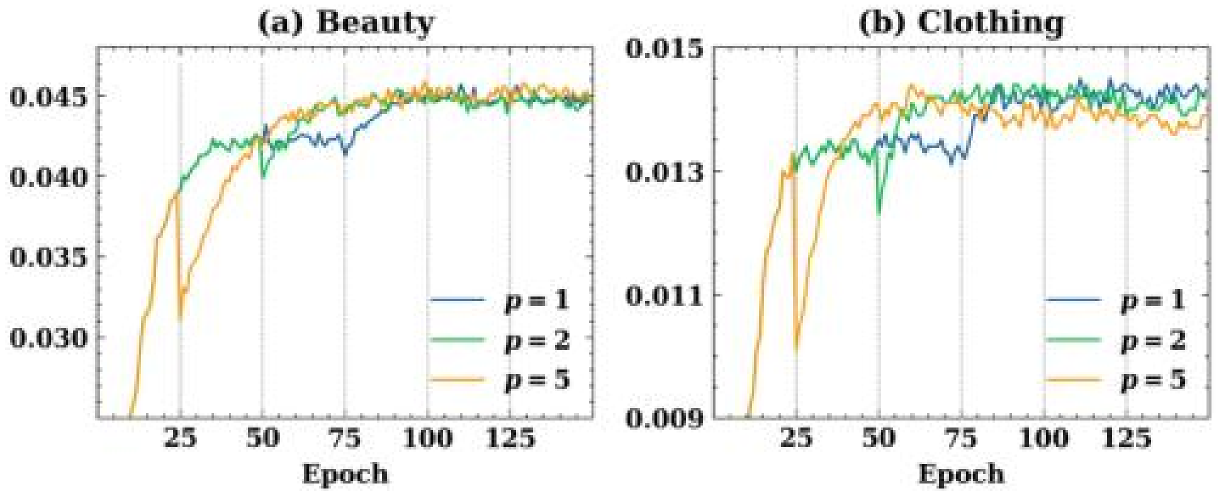

Figure 3: Changes of NDCG@5 for EC4SRec using different update schedules over 150 training epoches (p: number of importance score updates in training)

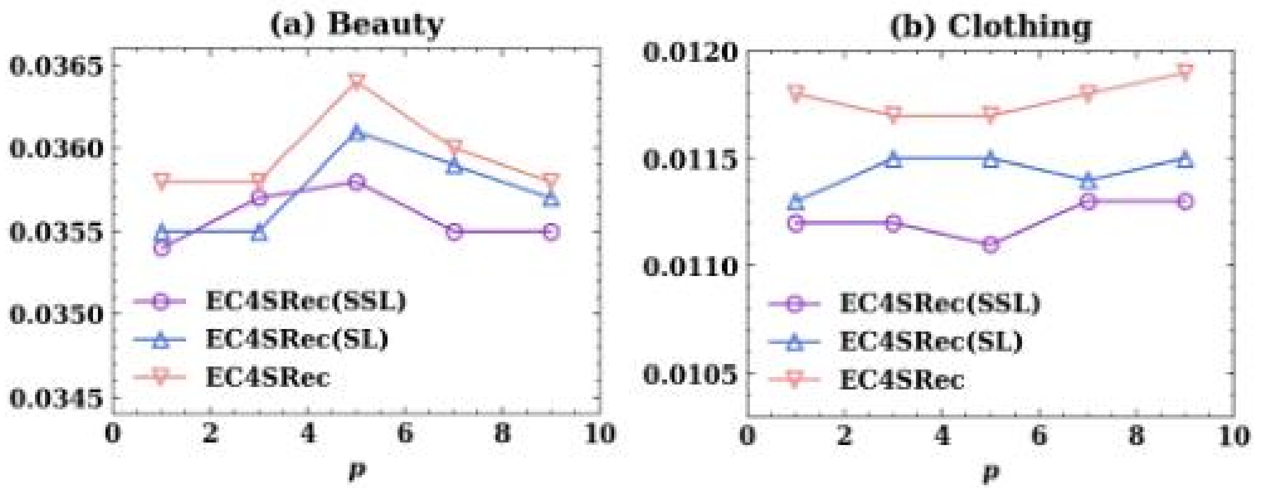

Figure 4: NDCG@5 Results with different update schedule settings (p: number of updates in model training).

Experiments

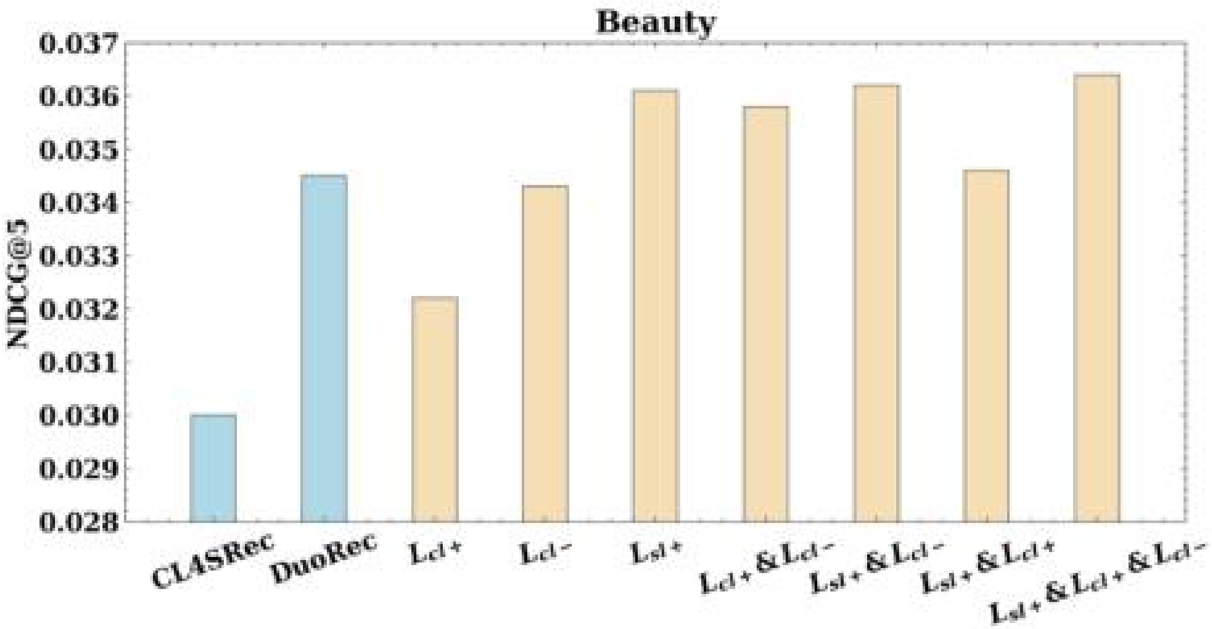

Figure 5: Ablation study of EC4SRec with different combinations of loss functions on Beauty dataset. (EC4SRec results are shown in yellow bars. As $\mathcal{L}_{rec}$ is included by default, EC4SRec(SSL) = EC4SRec with $\mathcal{L}_{cl+} + \mathcal{L}_{cl-}$; EC4SRec(SL) = EC4SRec with $\mathcal{L}_{sl+}$, and EC4SRec = one with all three losses.)

Experiments

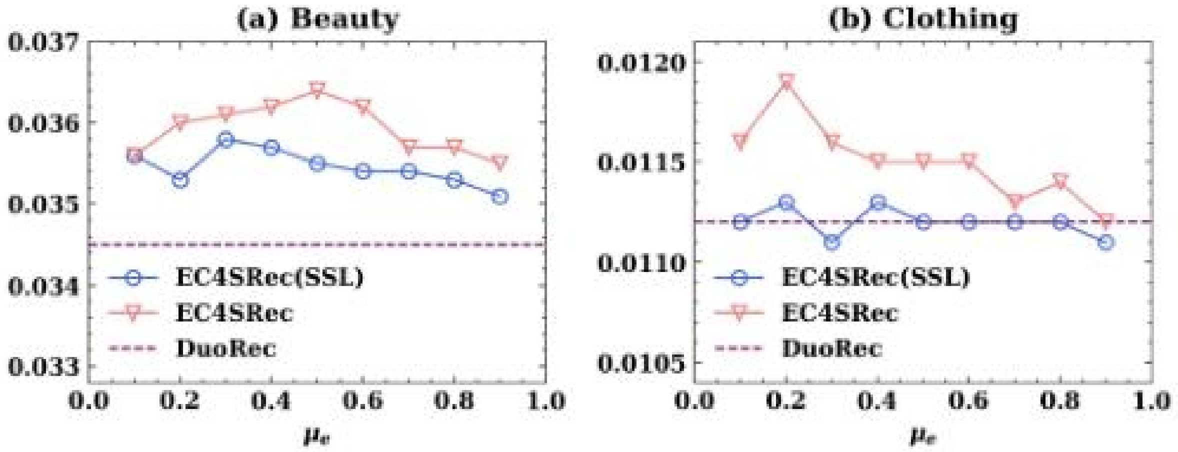

Figure 6: NDCG@5 of EC4SRec with different μ_e settings.

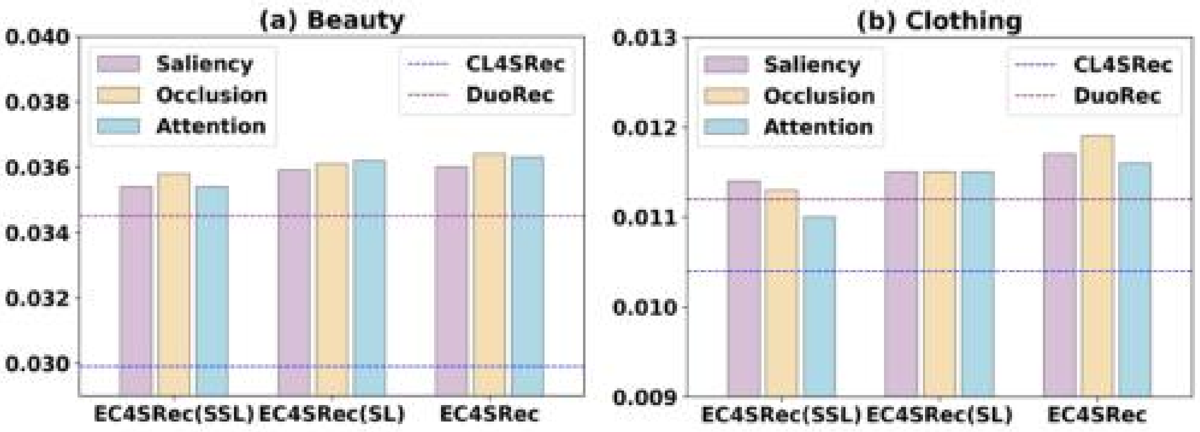

Figure 7: NDCG@5 using different explanation methods.

Thanks