Self-supervised Graph Learning for Recommendation

Jiancan Wu, Xiang Wang, Fuli Feng, Xiangnan He, Liang Chen, Jianxun Lian, Xing Xie

SIGIR 2021

目录

背景

传统基于GCN推荐模型的局限性

- 长尾问题：高度的结点对表征学习起了主导作用，导致对低度（长尾）项目的推荐变得困难
- 鲁棒性问题：节点表示容易受到噪声交互的影响

解决方案

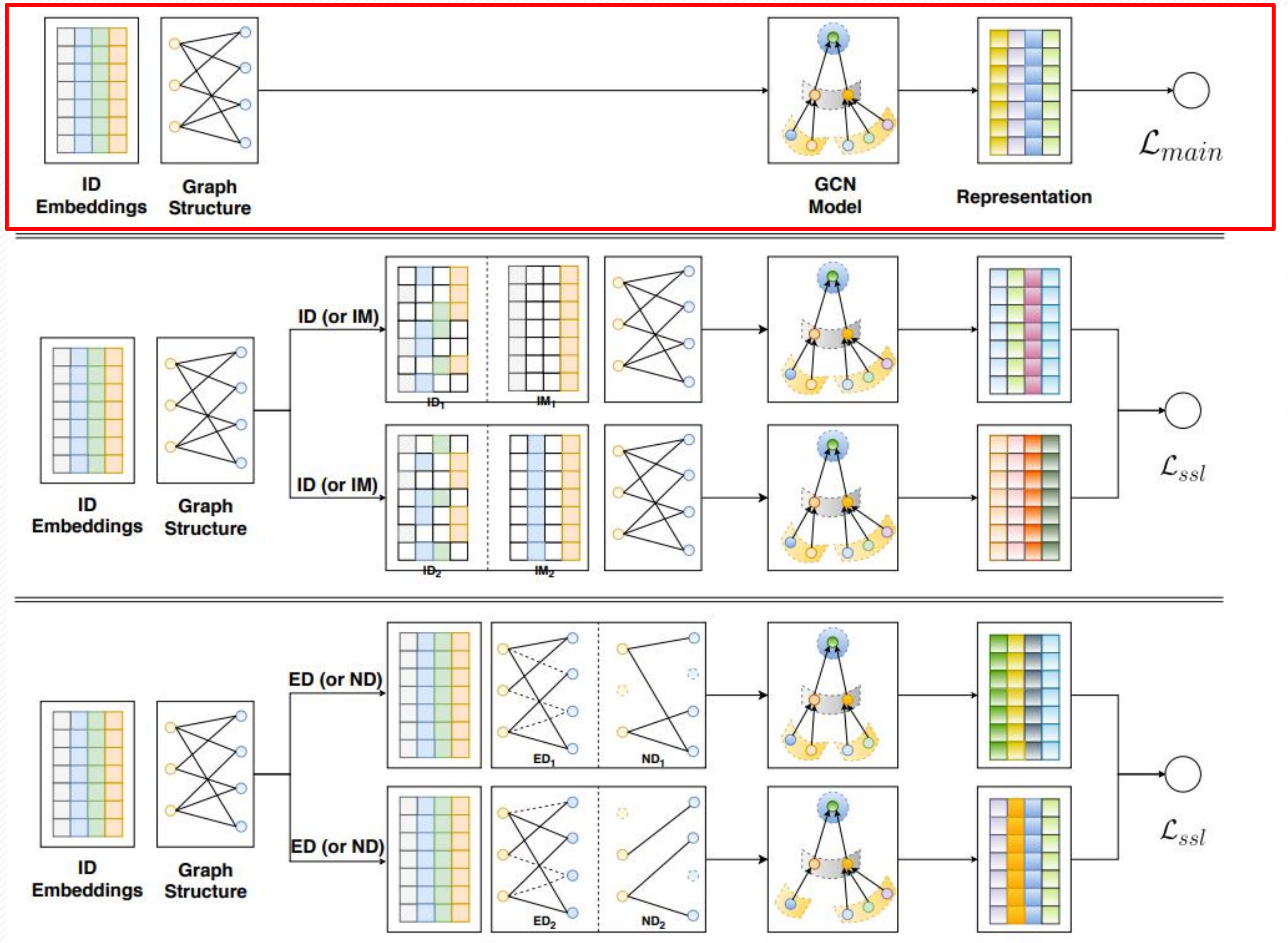
- 在传统监督任务的基础上，增加辅助的自监督学习任务，提高二分图推荐的准确性和鲁棒性

背景

本文贡献

- 第一个为基于图的推荐任务开发自监督学习的工作
- 设计了一个新的学习范式SGL，将节点自识别作为自监督任务，为节点表征学习提供辅助监督信号
- 在三个数据集上进行大量实验，证明了SGL的优越性

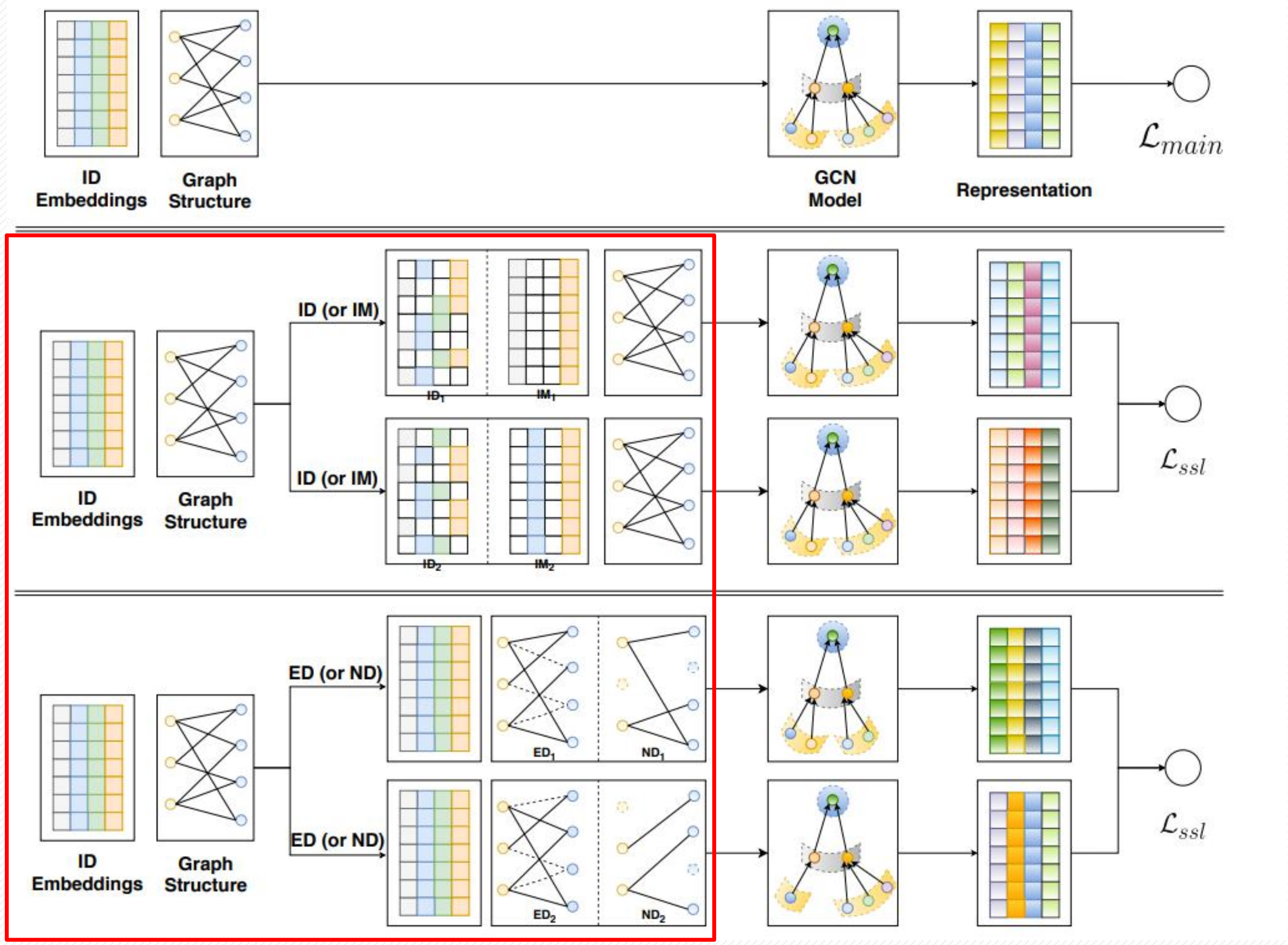
模型


ID Embeddings
Graph Structure
GCN Model
Representation
$\mathcal{L}_{main}$
ID (or IM)
ID_1
IM_1
ID_2
IM_2
$\mathcal{L}_{ssl}$
ED (or ND)
ED_1
ND_1
ED_2
ND_2

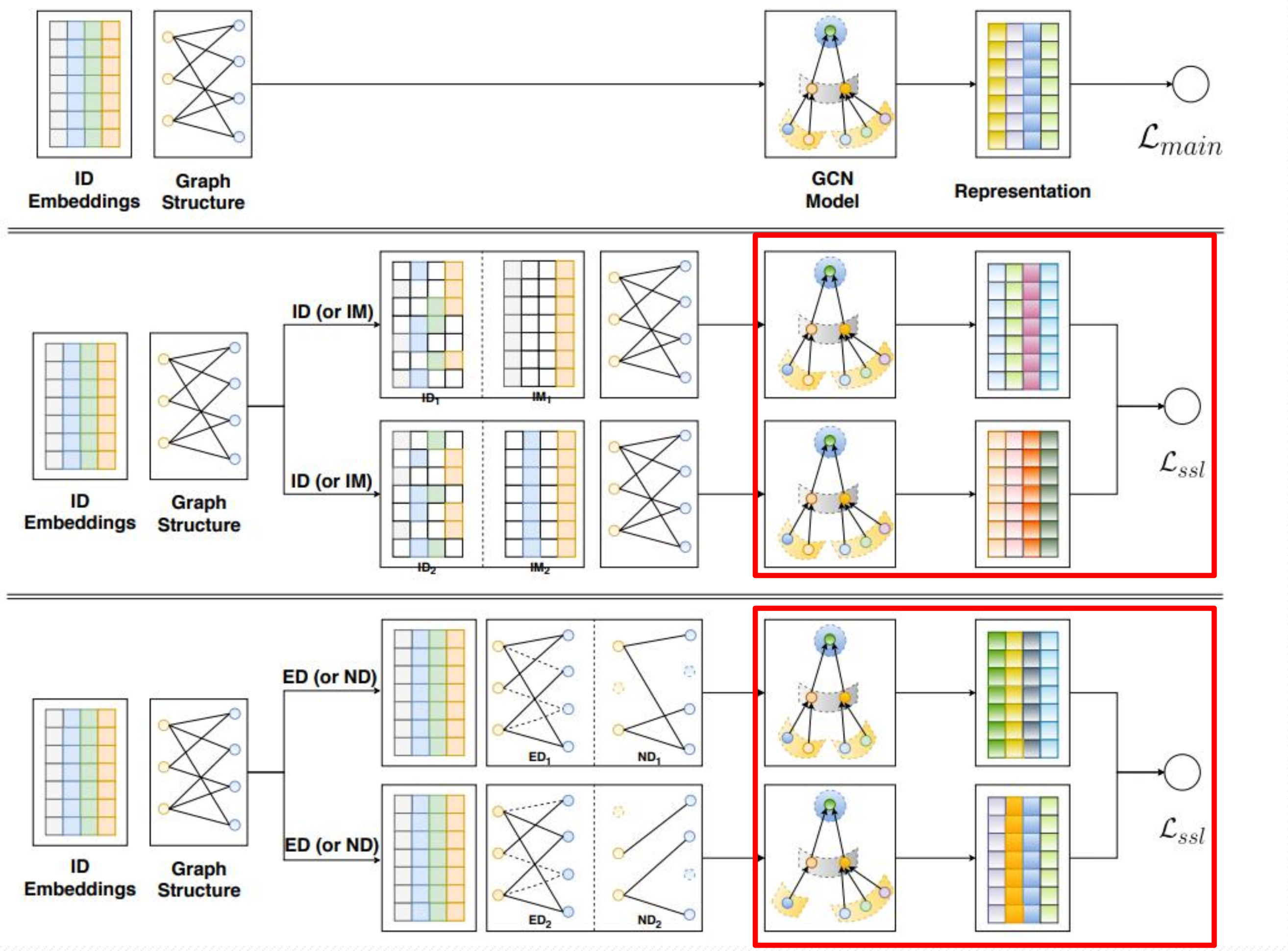
有监督任务流程

模型

一、数据增强，产生多个视图

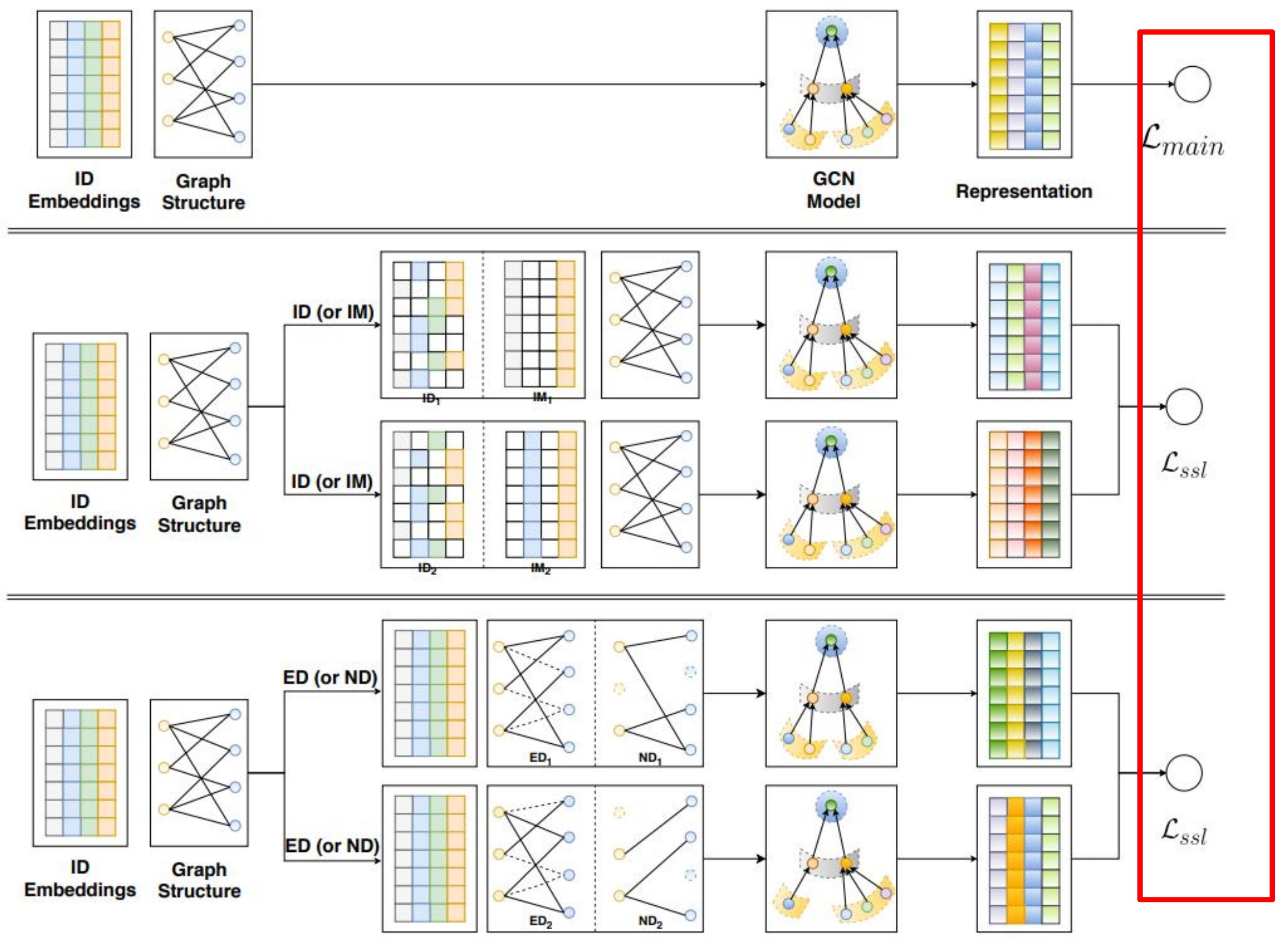


模型



二、学习不同视图上的节点表示，并在这些表示上做对比学习

模型



三、融合有监督任务和自监督任务，利用多目标学习框架优化

模型：数据增强

GCN节点编码范式：$Z^l = H(Z^{l-1}, G)$

输入：上一层的结点表征向量和原始图

输出：该层的结点表征向量

模型：数据增强

增强ID嵌入

- embedding masking（IM）
- embedding dropout（ID）

$$\mathbf{Z}' = H(t'(\mathbf{E}), \mathcal{G}), \quad \mathbf{Z}'' = H(t''(\mathbf{E}), \mathcal{G}), \quad t', t'' \sim \mathcal{T},$$

$$t'(\mathbf{E}) = \mathbf{M}' \odot \mathbf{E}, \quad t''(\mathbf{E}) = \mathbf{M}'' \odot \mathbf{E}, \quad M', M'' \in \{0, 1\}^{|\mathcal{V}|\times d}$$

作用：降低对某些信息的依赖性，提高模型的健壮性

M', M''通过伯努利分布$m \sim Bernoulli(\rho)$随机生成，ρ为dropout概率， M', M''完全独立

模型：数据增强

增强图结构

- Node Dropout（ND）
- Edge Dropout （ED）

$$\mathbf{Z}' = H(\mathbf{E}, s'(\mathcal{G})), \quad \mathbf{Z}'' = H(\mathbf{E}, s''(\mathcal{G})), \quad s', s'' \sim \mathcal{S},$$

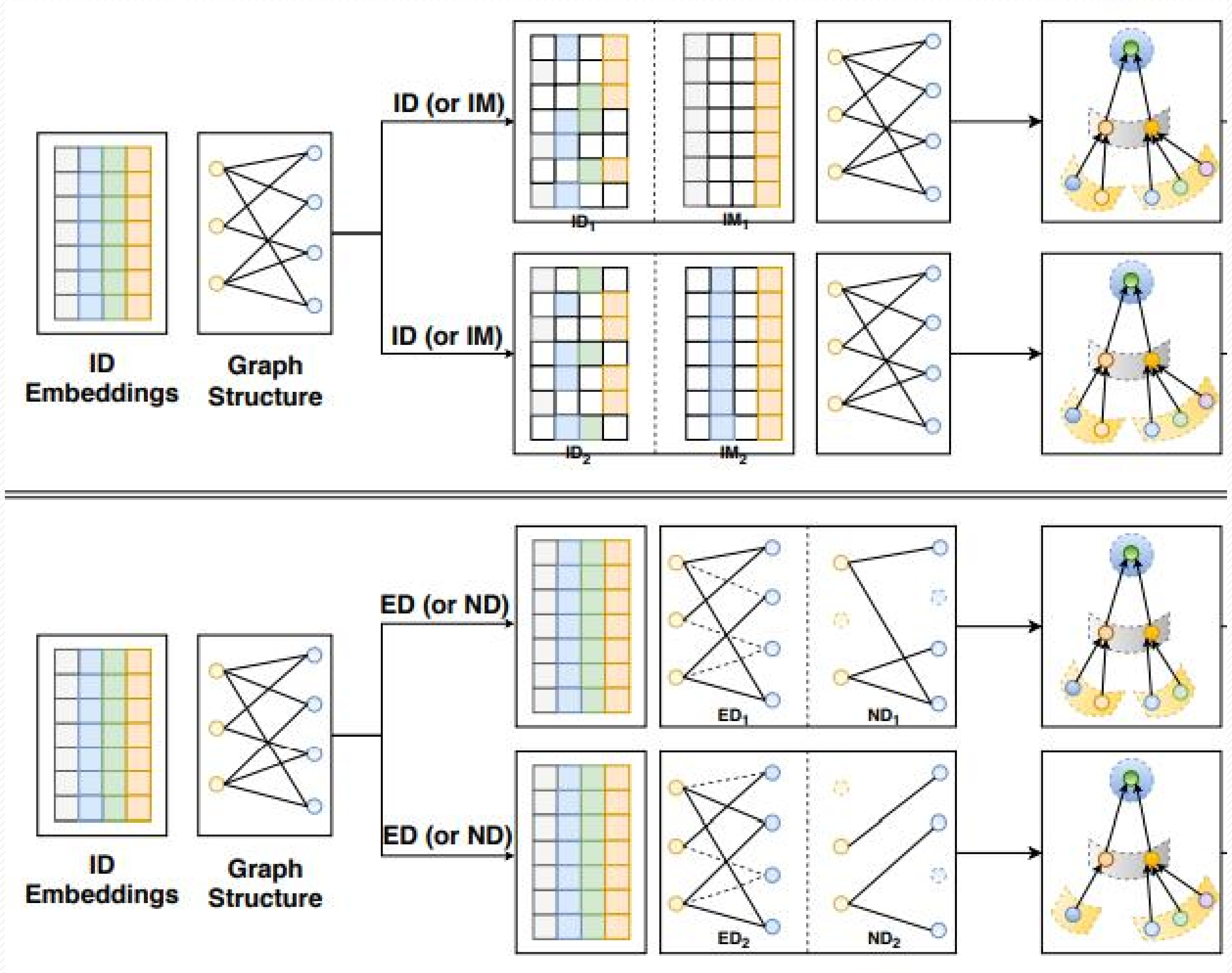
$$s'(\mathcal{G}) = (\mathbf{M}' \odot \mathcal{V}, \mathcal{E}), \quad s''(\mathcal{G}) = (\mathbf{M}'' \odot \mathcal{V}, \mathcal{E}), \qquad (ND)$$

$$s'(\mathcal{G}) = (\mathcal{V}, \mathbf{M}' \odot \mathcal{E}), \quad s''(\mathcal{G}) = (\mathcal{V}, \mathbf{M}'' \odot \mathcal{E}), \qquad (ED)$$

$$M',\, M'' \in \{0,\ 1\}^{|\mathcal{V}| \times d}$$

作用：从不同的视图中识别出有影响的节点，并降低表示学习对结构变化敏感性

模型：数据增强



ID (or IM)
ID (or IM)
ID Embeddings
Graph Structure
ED (or ND)
ED (or ND)
ID Embeddings
Graph Structure

模型：对比学习

同一节点在视图下可以产生不同的表示向量

- 正样本：$\{(Z'_u, Z''_u) | u \in \mathcal{U}\}$
- 负样本：$\{(Z'_u, Z''_v) | u, v \in \mathcal{U},\ u \neq v\}$

目标：最大化同一结点不同视图表征向量之间的相似性，最小化不同结点表征之间的相似性

$$\mathcal{L}_{ssl}^{user} = \sum_{u \in \mathcal{U}} -\log \frac{\exp(s(\mathbf{z}'_u, \mathbf{z}''_u)/\tau)}{\sum_{v \in \mathcal{U}} \exp(s(\mathbf{z}'_u, \mathbf{z}''_v)/\tau)},$$

$$\mathcal{L}_{ssl} = \mathcal{L}_{ssl}^{user} + \mathcal{L}_{ssl}^{item}$$

模型：多任务训练

本文采用多任务学习的方式训练模型

$$\mathcal{L} = \mathcal{L}_{main} + \lambda_1 \mathcal{L}_{ssl} + \lambda_2 \|\Theta\|_2^2$$

其中，θ表示图卷积神经网络的参数

Algorithm 1: Learning algorithm for SGL-ED

Input: Adjacency matrix of user-item graph, $\lambda_1, \lambda_2, \tau, \rho$

while *not converge* **do**
    foreach *epoch* **do**
        Perform Eq. (8) for data augmentation
        foreach *batch* **do**
            Evaluate $\mathcal{L}_{main}$ according to Eq. (5)
            Evaluate $\mathcal{L}_{self}$ according to Eq. (11)
            Evaluate $\mathcal{L}$ according to Eq. (12)
            Update the parameters by gradient descent
        end
    end
end

模型：实验

Table 2: Statistics of the datasets.

Dataset	#Users	#Items	#Interactions	Density
Yelp2018	31,668	38,048	1,561,406	0.00130
Amazon-Book	52,643	91,599	2,984,108	0.00062
Alibaba-iFashion	300,000	81,614	1,607,813	0.00007

Table 4: Overall Performance Comparison.

Dataset	Yelp2018		Amazon-Book		Alibaba-iFashion	
Method	Recall	NDCG	Recall	NDCG	Recall	NDCG
NGCF	0.0579	0.0477	0.0344	0.0263	0.1043	0.0486
LightGCN	0.0639	0.0525	0.0411	0.0315	0.1078	0.0507
Mult-VAE	0.0584	0.0450	0.0407	0.0315	0.1041	0.0497
SGL-ED	**0.0675**	**0.0555**	**0.0478**	**0.0379**	**0.1126**	**0.0538**

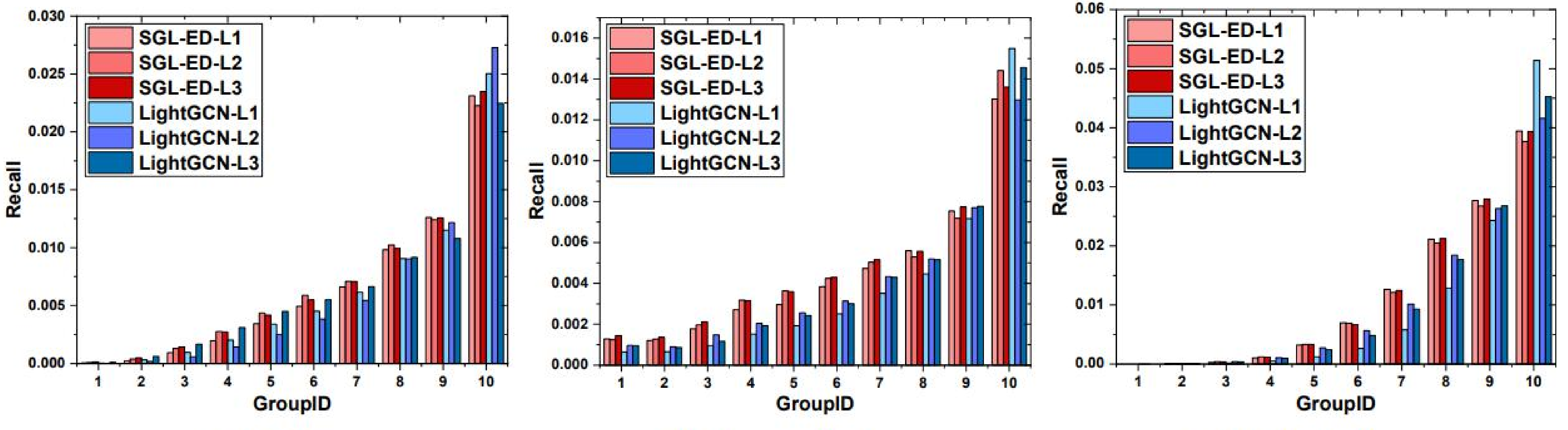
模型：实验

消融分析

Dataset		Yelp2018		Amazon-Book		Alibaba-iFashion	
#Layer	Method	Recall	NDCG	Recall	NDCG	Recall	NDCG
1 Layer	LightGCN	0.0631	0.0515	0.0384	0.0298	0.0990	0.0454
	SGL-ID	0.0634(+0.5%)	0.0518(+0.6%)	0.0417(+8.6%)	0.0322(+8.1%)	**0.1141(+15.3%)**	**0.0540(+18.9%)**
	SGL-IM	0.0631(+0%)	0.0513(-0.4%)	0.0429(11.7%)	0.0331(+11.1%)	0.1116(+12.7%)	0.0530(+16.7%)
	SGL-ND	**0.0643(+1.9%)**	**0.0529(+2.7%)**	0.0432(+12.5%)	0.0334(+12.1%)	0.1133(+14.4%)	0.0539(+18.7%)
	SGL-ED	0.0637(+1.0%)	0.0526(+2.1%)	**0.0451(+17.4%)**	**0.0353(+18.5%)**	0.1132(+14.3%)	0.0539(+18.7%)
2 Layers	LightGCN	0.0622	0.0504	0.0411	0.0315	0.1066	0.0505
	SGL-ID	0.0659(+5.9%)	0.0539(+6.9%)	0.0436(+6.1%)	0.0342(+8.6%)	0.1089(+2.2%)	0.0517(+2.4%)
	SGL-IM	0.0657(+5.6%)	0.0538(+6.7%)	0.0434(+5.6%)	0.0338(+7.3%)	0.1085(+1.8%)	0.0517(+2.4%)
	SGL-ND	0.0658(+5.8%)	0.0538(+6.7%)	0.0427(+3.9%)	0.0335(+6.3%)	**0.1106(+3.8%)**	**0.0526(+4.2%)**
	SGL-ED	**0.0668(+7.4%)**	**0.0549(+8.9%)**	**0.0468(+13.9%)**	**0.0371(+17.8%)**	0.1091(+2.3%)	0.0520(+3.0%)
3 Layers	LightGCN	0.0639	0.0525	0.0410	0.0318	0.1078	0.0507
	SGL-ID	0.0649(+1.6%)	0.0533(+1.5%)	0.0450(+9.8%)	0.0353(+11.0%)	0.1119(+3.8%)	0.0529(+4.3%)
	SGL-IM	0.0652(+2.0%)	0.0536(+2.1%)	0.0449(+9.5%)	0.0353(+11.0%)	0.1121(+4.0%)	0.0537(+5.9%)
	SGL-ND	0.0644(+0.8%)	0.0528(0.6%)	0.0440(+7.3%)	0.0346(+8.8%)	0.1126(4.5%)	0.0536(+5.7%)
	SGL-ED	**0.0675(+5.6%)**	**0.0555(+5.7%)**	**0.0478(+16.6%)**	**0.0379(+19.2%)**	**0.1126(+4.5%)**	**0.0538(+6.1%)**

模型：实验

长尾推荐

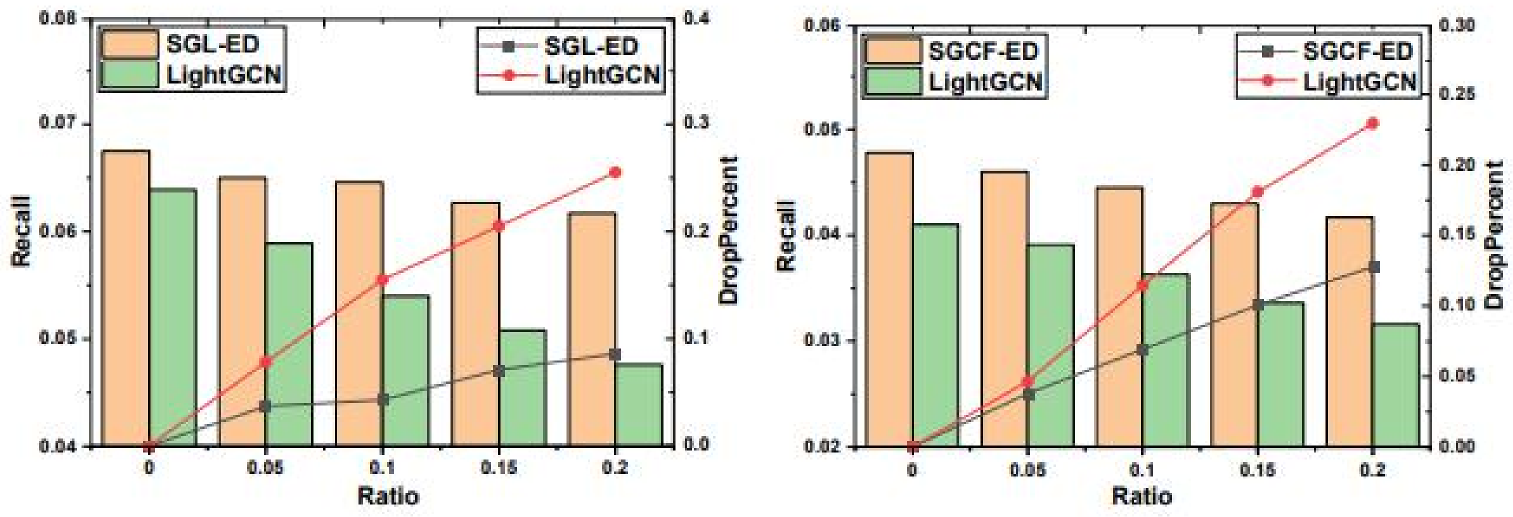


(a) Yelp2018 (b) Amazon-Book (c) Alibaba-iFashion

分组ID越小，度数越小，越长尾

模型：实验

鲁棒性分析



(a) Yelp2018 (b) Amazon-Book

在训练集中加入对抗样本