



(12) 发明专利

(10) 授权公告号 CN 113961816 B

(45) 授权公告日 2022.07.01

(21) 申请号 202111423525.2

G06N 3/04 (2006.01)

(22) 申请日 2021.11.26

G06N 3/08 (2006.01)

(65) 同一申请的已公布的文献号  
申请公布号 CN 113961816 A

(56) 对比文件

CN 113326425 A, 2021.08.31

US 2019340649 A1, 2019.11.07

(43) 申请公布日 2022.01.21

审查员 彭傲雪

(73) 专利权人 重庆理工大学  
地址 400054 重庆市巴南区李家沱红光大道69号

(72) 发明人 朱小飞 唐頔

(74) 专利代理机构 重庆博凯知识产权代理有限公司 50212  
专利代理师 黄河

(51) Int. Cl.

G06F 16/9535 (2019.01)

G06F 40/35 (2020.01)

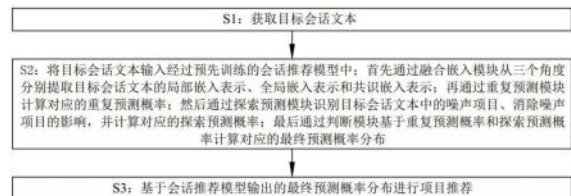
权利要求书4页 说明书16页 附图5页

(54) 发明名称

一种基于结构增强的图卷积神经网络会话推荐方法

(57) 摘要

本发明涉及会话推荐技术领域,具体涉及一种基于结构增强的图卷积神经网络会话推荐方法,首先获取目标会话文本的会话表示;然后基于会话表示生成对应的会话图,进而通过会话图识别目标会话文本中的噪声项目;再结合注意力机制重置噪声项目的注意力权重,以消除噪声项目的影响;最后计算最终预测概率分布,并基于最终预测概率分布进行项目推荐。本发明中基于结构增强的图卷积神经网络会话推荐方法能够识别噪声项目并解决噪声项目影响,从而能够提升会话项目推荐的准确性。



1. 一种基于结构增强的图卷积神经网络会话推荐方法,其特征在于:首先获取目标会话文本的会话表示;然后基于会话表示生成对应的会话图,进而通过会话图识别目标会话文本中的噪声项目;再结合注意力机制重置噪声项目的注意力权重,以消除噪声项目的影响;最后计算最终预测概率分布,并基于最终预测概率分布进行项目推荐;

具体包括以下步骤:

S1:获取目标会话文本;

S2:将目标会话文本输入经过预先训练的会话推荐模型中;

所述会话推荐模型首先通过融合嵌入模块从三个角度分别提取目标会话文本的局部嵌入表示、全局嵌入表示和共识嵌入表示;再通过重复预测模块计算对应的重复预测概率;然后通过探索预测模块识别目标会话文本中的噪声项目、消除噪声项目的影响,并计算对应的探索预测概率;最后通过判断模块基于重复预测概率和探索预测概率计算对应的最终预测概率分布;

S3:基于会话推荐模型输出的最终预测概率分布进行项目推荐;

其中,重复预测模块通过如下步骤计算重复预测概率:

S221:基于局部嵌入表示  $\tilde{H}_s^{(l)}$ 、全局嵌入表示  $\tilde{G}_s^{(l)}$ 、共识嵌入表示  $\tilde{M}_s^{(l)}$  和位置嵌入表示  $\tilde{P}_s^{(l)}$  计算对应的新会话表示  $X_s = (x_1^s, x_2^s, \dots, x_n^s)$ ;

其中,  $x_i^s = \sigma(W_x(\tilde{h}_i^{s,(L)}, \tilde{g}_i^{s,(L)}, \tilde{m}_i^{s,(L)}) + b_x) + \tilde{p}_i^{s,(L)}$ ;

S222:基于新会话表示  $X_s$  结合如下公式计算重复预测概率  $\hat{Y}_i^r$ ;

$$\hat{Y}_i^r = \begin{cases} z_i & \text{if: } v_i \notin S, \\ 0 & \text{if: } v_i \in S; \end{cases}$$

$$\text{其中, } z_i = \frac{\exp(W_r^T x_i^s)}{\sum_{j=1}^n \exp(W_r^T x_j^s)};$$

上述式中:  $W_x \in \mathbb{R}^{d \times 3d}$ 、 $b_x \in \mathbb{R}^d$ 、 $W_r \in \mathbb{R}^d$  表示可训练参数;  $x_j^s$  表示聚合了会话中第  $j$  个项目  $h_j^s$  的局部表示、全局表示、共识表示以及其位置嵌入后的综合表示;

探索预测模块包括用于消除噪声项目影响的图增强注意力网络;

图增强注意力网络通过如下步骤消除噪声项目的影响:

S231:基于局部嵌入表示  $\tilde{H}_s^{(l)}$  建立对应的错误点击集合  $S_{wrong}^L$ ,通过错误点击集合  $S_{wrong}^L$  结合位置嵌入表示  $\tilde{P}_s^{(l)}$  计算对应的长期偏好  $S_{long}^L$ ,结合长期偏好  $S_{long}^L$  和局部嵌入表示  $\tilde{H}_s^{(l)}$  中的当前偏好  $\tilde{h}_n^{s,(l)}$  计算对应的会话局部表示  $\hat{S}_{long}^L$ ;

$$\text{其中, } S_{long}^L = \sum_{i=1}^n \lambda_i \tilde{h}_i^{s,(L)}, \quad \lambda_i = \frac{\exp(\beta_i)}{\sum_{j=1}^n \exp(\beta_j)}, \quad \beta_i = \begin{cases} W_e^T (W_2 \tilde{h}_i^{s,(L)} + W_3 \tilde{p}_i^{s,(L)}), & \text{if: } \tilde{h}_i^{s,(L)} \notin (\tilde{H}_s^{(L)} - S_{wrong}^L) \\ -\infty, & \text{if: } \tilde{h}_i^{s,(L)} \in S_{wrong}^L, \end{cases};$$

S232:基于全局嵌入表示  $\tilde{G}_s^{(l)}$  建立对应的错误点击集合  $S_{wrong}^G$ ,通过错误点击集合  $S_{wrong}^G$  结合位置嵌入表示  $\tilde{P}_s^{(l)}$  计算对应的长期偏好  $S_{long}^G$ ,结合长期偏好  $S_{long}^G$  和全局嵌入表示  $\tilde{G}_s^{(l)}$  中的当前偏好  $\tilde{g}_n^{s,(l)}$  计算对应的会话全局表示  $\hat{S}_{long}^G$ ;

S233: 基于共识嵌入表示  $\tilde{M}_s^{(l)}$  建立对应的错误点击集合  $S_{wrong}^M$ , 通过错误点击集合  $S_{wrong}^M$  结合位置嵌入表示  $\tilde{P}_s^{(l)}$  计算对应的长期偏好  $S_{long}^M$ , 结合长期偏好  $S_{long}^M$  和共识嵌入表示  $\tilde{M}_s^{(l)}$  中的当前偏好  $\tilde{m}_n^{s,(l)}$  计算对应的会话共识表示  $\hat{S}_{long}^M$ ;

S234: 结合会话局部表示  $\hat{S}_{long}^L$ 、会话全局表示  $\hat{S}_{long}^G$  和会话共识表示  $\hat{S}_{long}^M$  计算对应的会话结构表示  $s_{long}^{str}$ ;

$$\text{其中, } s_{long}^{str} = \hat{S}_{long}^L + \hat{S}_{long}^G + \hat{S}_{long}^M;$$

上述式中:  $W_e \in \mathbb{R}^d$ 、 $W_2 \in \mathbb{R}^{d \times d}$ 、 $W_3 \in \mathbb{R}^d$  表示可训练参数;

探索预测模块通过如下步骤计算探索预测概率:

S241: 通过Bi-GRU模块结合局部嵌入表示  $\tilde{H}_s^{(l)}$ 、全局嵌入表示  $\tilde{G}_s^{(l)}$  和共识嵌入表示  $\tilde{M}_s^{(l)}$  计算对应的会话顺序表示  $s_{long}^{seq}$ ;

$$\text{其中, } s_{long}^{seq} = d_n^{GRU};$$

$$d_i^{GRU} = \text{LayerNorm}(W_{gru}([\tilde{d}_i^{GRU}; \tilde{d}_i^{GRU}]));$$

$$\tilde{d}_i^{GRU} = \text{GRU}_{forward}(\tilde{d}_{i-1}^{GRU}, d_i, \phi_{gru});$$

$$\tilde{d}_i^{GRU} = \text{GRU}_{backward}(\tilde{d}_{i+1}^{GRU}, d_i, \phi_{gru});$$

$$d_i = W_{seq}([\tilde{h}_i^{s,(l)}; \tilde{s}_i^{s,(l)}; \tilde{m}_i^{s,(l)}]) + b_{seq};$$

S242: 基于会话结构表示  $s_{long}^{str}$  和会话顺序表示  $s_{long}^{seq}$  计算对应的探索预测概率  $\hat{y}_i^e$ ;

$$\hat{y}_i^e = \frac{\exp(\rho_i)}{\sum_{i=1}^{|V|} \exp(\rho_i)};$$

$$\text{其中, } \rho_i = \begin{cases} s_e^T h_i, & \text{if } v_i \notin S \\ -\infty, & \text{if } v_i \in S \end{cases}, \quad s_e = s_{long}^{str} + s_{long}^{seq};$$

上述式中:  $W_{seq} \in \mathbb{R}^{d \times 3d}$ 、 $W_{gru} \in \mathbb{R}^{d \times 2d}$ 、 $b_{seq} \in \mathbb{R}^d$ 、 $\phi_{gru}$  表示可训练参数; LayerNorm表示归一化;  $h_i \in \mathbb{R}^d$  表示候选项目集合V中第i个项目的嵌入表示。

2. 如权利要求1所述的基于结构增强的图卷积神经网络会话推荐方法, 其特征在于: 所述融合嵌入模块包括用于基于目标会话文本生成全局表示的动态全局邻居注意力网络, 以及用于基于目标会话文本的局部表示和全局表示融合生成对应的局部嵌入表示、全局嵌入表示和共识嵌入表示的多通道图卷积神经网络。

3. 如权利要求2所述的基于结构增强的图卷积神经网络会话推荐方法, 其特征在于, 所述动态全局邻居注意力网络通过如下步骤生成全局表示:

S201: 对目标会话文本  $S = (v_1^s, v_2^s, \dots, v_n^s)$  进行编码, 得到对应的局部表示  $H_s = (h_1^s, h_2^s, \dots, h_n^s)$ ;

S202: 从候选项目集合  $V = \{v_1, v_1, \dots, v_{|V|}\}$  中提取K个候选项目  $v_j \in V$  作为目标项目  $v_i^s \in S$  的全局邻居  $N^i = (n_1^i, n_2^i, \dots, n_K^i)$ ;

S203: 基于全局邻居  $N^i = (n_1^i, n_2^i, \dots, n_K^i)$  生成对应的全局邻居嵌入表示  $C^i = (c_1^i, c_2^i, \dots, c_K^i)$ ;

S204: 通过注意力池化机制聚合目标项目  $v_i^s$  的全局邻居嵌入表示  $C^i = (c_1^i, c_2^i, \dots, c_K^i)$  得到

目标项目  $v_i^s$  的全局表示  $g_i^s$ ;

$$\text{其中, } g_i^s = \sum_{j=1}^K \partial_j^i c_j^i, \quad \partial_j^i = \frac{\exp(W_g^T [c_j^i; h_i^s])}{\sum_{j=1}^K \exp(W_g^T [c_j^i; h_i^s])};$$

S205: 计算各个目标项目  $v_i^s$  的会话全局表示  $g_i^s$ , 并生成目标会话文本S的全局表示  $G_s = (g_1^s, g_2^s, \dots, g_n^s)$ ;

上述式中:  $W_g \in \mathbb{R}^{2d}$  表示可训练参数;  $[\cdot]$  表示拼接操作。

4. 如权利要求3所述的基于结构增强的图卷积神经网络会话推荐方法, 其特征在于, 所述多通道图卷积神经网络通过如下步骤生成局部嵌入表示、全局嵌入表示和共识嵌入表示:

S211: 基于局部表示  $H_s = (h_1^s, h_2^s, \dots, h_n^s)$  构造对应的局部图  $A^{L, (1-1)}$ , 并通过局部图  $A^{L, (1-1)}$  计算对应的局部嵌入表示  $\tilde{H}_s^{(l)} = (\tilde{h}_1^{s, (l)}, \tilde{h}_2^{s, (l)}, \dots, \tilde{h}_n^{s, (l)})$ ;

$$\text{其中, } \tilde{H}_s^{(l)} = HWGCN(A^{L, (1-1)}, H_s^{(1-1)}), \quad A^{L, (1-1)} = BSG(\tilde{H}_s^{(l-1)});$$

S212: 基于全局表示  $G_s = (g_1^s, g_2^s, \dots, g_n^s)$  构造对应的全局图  $A^{G, (1-1)}$ , 并通过全局图  $A^{G, (1-1)}$  计算对应的全局嵌入表示  $\tilde{G}_s^{(l)} = (\tilde{g}_1^{s, (l)}, \tilde{g}_2^{s, (l)}, \dots, \tilde{g}_n^{s, (l)})$ ;

$$\text{其中, } \tilde{G}_s^{(l)} = HWGCN(A^{G, (1-1)}, G_s^{(1-1)}); \quad A^{G, (1-1)} = BSG(\tilde{G}_s^{(l-1)});$$

S213: 基于局部图  $A^{L, (1-1)}$  和全局图  $A^{G, (1-1)}$  融合生成对应的共识图  $A^{M, (1-1)}$ ;

$$\text{其中, } A^{M, (1-1)} = f_l \cdot A^{L, (1-1)} + (1 - f_l) \cdot A^{G, (1-1)}, \quad f_l = \sigma(W_{\omega}^T [h_i^{s, (l)}; g_i^{s, (l)}]);$$

S214: 通过共识图  $A^{M, (1-1)}$  计算对应的共识嵌入表示  $\tilde{M}_s^{(l)} = (\tilde{m}_1^{s, (l)}, \tilde{m}_2^{s, (l)}, \dots, \tilde{m}_n^{s, (l)})$ ;

$$\text{其中, } \tilde{M}_s^{(l)} = HWGCN(A^{M, (1-1)}, \tilde{M}_s^{(l-1)});$$

上述式中: HWGCN表示多通道图卷积神经网络;  $l$ 表示多通道图卷积神经网络的层数;  $\tilde{H}_s^{(l-1)}$ 表示第  $l-1$  层HWGCN输出的局部嵌入表示;  $\tilde{G}_s^{(l-1)}$ 表示第  $l-1$  层HWGCN输出的全局嵌入表示;  $\tilde{M}_s^{(l-1)}$ 表示第  $l-1$  层HWGCN输出的共识嵌入表示;  $\sigma$ 表示Sigmoid函数;  $W_{\omega} \in \mathbb{R}^{2d}$ 表示可训练参数。

5. 如权利要求4所述的基于结构增强的图卷积神经网络会话推荐方法, 其特征在于: 所述会话推荐模型还包括图增强的位置编码器;

所述位置编码器通过共识图  $A^{M, (1-1)}$  计算对应的位置嵌入表示  $\tilde{P}_s^{(l)} = (\tilde{p}_1^{s, (l)}, \tilde{p}_2^{s, (l)}, \dots, \tilde{p}_n^{s, (l)})$ ;

$$\text{其中, } \tilde{P}_s^{(l)} = HWGCN(A^{M, (1-1)}, \tilde{P}_s^{(l-1)});$$

位置编码器的结构表示为:

$$\tilde{p}_i^s = p_i^s + p_{len}^s;$$

$$p_{len}^s = Embed_{len}(S_{len});$$

上述式中:  $p_i^s \in \mathbb{R}^d$ 表示目标会话文本S中第  $i$  个位置的位置嵌入;  $S_{len}$ 表示目标会话文本S的会话长度;  $p_{len}^s \in \mathbb{R}^d$ 表示会话长度  $S_{len}$  的嵌入表示。

6. 如权利要求1所述的基于结构增强的图卷积神经网络会话推荐方法, 其特征在于, 判

断模块通过如下公式计算最终预测概率；

$$\hat{y}_i = p \cdot \hat{y}_i^r + (1-p) \cdot \hat{y}_i^e;$$

$$\text{其中, } p = \delta(W_p^T([\mathbf{o}_r; \mathbf{s}_e] \mathbf{D})), \quad \mathbf{o}_r = \sum_{i=1}^n z_i \mathbf{x}_i^s, \quad \mathbf{s}_e = \mathbf{s}_{long}^{str} + \mathbf{s}_{long}^{seq};$$

上述式中： $W_p \in \mathbb{R}^d$ 表示可训练参数； $\mathbf{o}_r$ 和 $\mathbf{s}_e$ 分别表示重复预测模块和探索预测模块的会话表示； $\delta$ 表示超参数。

## 一种基于结构增强的图卷积神经网络会话推荐方法

### 技术领域

[0001] 本发明涉及互联网大数据技术领域,具体涉及一种基于结构增强的图卷积神经网络会话推荐方法。

### 背景技术

[0002] 基于会话的推荐是一种针对匿名用户或未登录用户的一种推荐模式,其在如今的各大电商平台(淘宝、京东等)或流媒体平台(抖音,YouTobe等)发挥着重要作用。在实际场景中,某些时候只能获取到用户的短期历史交互,比如:新用户或未登录用户。此时,依赖于用户长期历史交互的推荐算法在会话推荐中的表现会收到限制,例如基于协同过滤或马尔可夫链的方法。因此,基于会话的推荐成为一个研究热点,其目标是根据用户在会话中的行为序列来推荐用户感兴趣的下一个项目(或商品)。

[0003] 针对现有会话推荐方法的项目推荐准确性不高的问题,公开号为CN112035746A的中国专利公开了《一种基于时空序列图卷积网络的会话推荐方法》,其包括:将所有会话序列建模为有向会话图;以会话中共有的商品为链接,构建全局图;将ARMA过滤器嵌入到门控图神经网络中,提取图模型中随时间变化的拓扑图信号,并得到会话图中涉及的各个节点的特征向量;采用注意力机制从用户历史会话中得到全局偏好信息;从用户点击的最后一个会话中获取用户的局部偏好信息,并结合全局偏好信息得到用户最终偏好信息;预测每个会话中下一点击商品可能出现的概率,并给出Top-K推荐商品。

[0004] 上述现有方案中的会话推荐方法从全局图中捕获丰富的会话表示(上下文关系),通过注意力机制学习用户的全局和局部偏好,进而提供准确的商品预测。但是,现有GNN(图卷积网络)中的图结构是由静态方式构造的,一旦用户的偏好发生转移,就很容易为图结构带来噪声项目(信息),导致图结构不准确,同时,项目转换关系中通常包含由用户意外或错误点击引起的噪声项目。然而,传统的注意力机制会为会话文本中的每一个项目分配注意力权重(重要性系数),即噪声项目也会分配到相应的注意力权重,这容易将噪声信息引入到会话表示中,导致模型学习不准确的用户偏好表示,进而导致推荐准确性偏低。因此,如何设计一种能够解决噪声影响的会话推荐方法是亟需解决的技术问题。

### 发明内容

[0005] 针对上述现有技术的不足,本发明所要解决的技术问题是:如何提供一种能够识别噪声项目并解决噪声项目影响的会话推荐方法,从而提升会话项目推荐的准确性。

[0006] 为了解决上述技术问题,本发明采用了如下的技术方案:

[0007] 一种基于结构增强的图卷积神经网络会话推荐方法:首先获取目标会话文本的会话表示;然后基于会话表示生成对应的会话图,进而通过会话图识别目标会话文本中的噪声项目;再结合注意力机制重置噪声项目的注意力权重,以消除噪声项目的影响;最后计算最终预测概率分布,并基于最终预测概率分布进行项目推荐。

[0008] 优选的,具体包括以下步骤:

[0009] S1:获取目标会话文本;

[0010] S2:将目标会话文本输入经过预先训练的会话推荐模型中;

[0011] 所述会话推荐模型首先通过融合嵌入模块从三个角度分别提取目标会话文本的局部嵌入表示、全局嵌入表示和共识嵌入表示;再通过重复预测模块计算对应的重复预测概率;然后通过探索预测模块识别目标会话文本中的噪声项目、消除噪声项目的影响,并计算对应的探索预测概率;最后通过判断模块基于重复预测概率和探索预测概率计算对应的最终预测概率分布;

[0012] S3:基于会话推荐模型输出的最终预测概率分布进行项目推荐。

[0013] 优选的,所述融合嵌入模块包括用于基于目标会话文本生成全局表示的动态全局邻居注意力网络,以及用于基于目标会话文本的局部表示和全局表示融合生成对应的局部嵌入表示、全局嵌入表示和共识嵌入表示的多通道图卷积神经网络。

[0014] 优选的,所述动态全局邻居注意力网络通过如下步骤生成全局表示:

[0015] S201:对目标会话文本  $S=(v_1^s, v_2^s, \dots, v_n^s)$  进行编码,得到对应的局部表示  $H_s=(h_1^s, h_2^s, \dots, h_n^s)$ ;

[0016] S202:从候选项目集合  $V=\{v_1, v_1, \dots, v_{|V|}\}$  中提取K个候选项目  $v_j \in V$  作为目标项目  $v_i^s \in S$  的全局邻居  $N^i=(n_1^i, n_2^i, \dots, n_k^i)$ ;

[0017] S203:基于全局邻居  $N^i=(n_1^i, n_2^i, \dots, n_k^i)$  生成对应的全局邻居嵌入表示  $C^i=(c_1^i, c_2^i, \dots, c_k^i)$ ;

[0018] S204:通过注意力池化机制聚合目标项目  $v_i^s$  的全局邻居嵌入表示  $C^i=(c_1^i, c_2^i, \dots, c_k^i)$  得到目标项目  $v_i^s$  的全局表示  $g_i^s$ ;

[0019] 其中,  $g_i^s = \sum_{j=1}^K \partial_j^i c_j^i$ ,  $\partial_j^i = \frac{\exp(W_g^T [c_j^i; h_i^s])}{\sum_{j=1}^K \exp(W_g^T [c_j^i; h_i^s])}$ ;

[0020] S205:计算各个目标项目  $v_i^s$  的会话全局表示  $g_i^s$ ,并生成目标会话文本S的全局表示  $G_s=(g_1^s, g_2^s, \dots, g_n^s)$ ;

[0021] 上述式中: $W_g \in \mathbb{R}^{2d}$  表示可训练参数;[;]表示拼接操作。

[0022] 优选的,所述多通道图卷积神经网络通过如下步骤生成局部嵌入表示、全局嵌入表示和共识嵌入表示:

[0023] S211:基于局部表示  $H_s=(h_1^s, h_2^s, \dots, h_n^s)$  构造对应的局部图  $A^{L, (1-1)}$ ,并通过局部图  $A^{L, (1-1)}$  计算对应的局部嵌入表示  $\tilde{H}_s^{(l)}=(\tilde{h}_1^{s,(l)}, \tilde{h}_2^{s,(l)}, \dots, \tilde{h}_n^{s,(l)})$ ;

[0024] 其中,  $\tilde{H}_s^{(l)} = HWGCN(A^{L,(l-1)}, \tilde{H}_s^{(l-1)})$ ,  $A^{L,(l-1)} = BSG(\tilde{H}_s^{(l-1)})$ ;

[0025] S212:基于全局表示  $G_s=(g_1^s, g_2^s, \dots, g_n^s)$  构造对应的全局图  $A^{G, (1-1)}$ ,并通过全局图  $A^{G, (1-1)}$  计算对应的全局嵌入表示  $\tilde{G}_s^{(l)}=(\tilde{g}_1^{s,(l)}, \tilde{g}_2^{s,(l)}, \dots, \tilde{g}_n^{s,(l)})$ ;

[0026] 其中,  $\tilde{G}_s^{(l)} = HWGCN(A^{G,(l-1)}, \tilde{G}_s^{(l-1)})$ ;  $A^{G,(l-1)} = BSG(\tilde{G}_s^{(l-1)})$ ;

[0027] S213:基于局部图 $A^{L, (1-1)}$ 和全局图 $A^{G, (1-1)}$ 融合生成对应的共识图 $A^{M, (1-1)}$ ;

[0028] 其中,  $A^{M, (l-1)} = f_i \cdot A_i^{L, (l-1)} + (1-f_i) \cdot A_i^{G, (l-1)}$ ,  $f_i = \sigma(W_{\omega}^T [\tilde{h}_i^{s, (l)}, \tilde{g}_i^{s, (l)}])$ ;

[0029] S214:通过共识图 $A^{M, (1-1)}$ 计算对应的共识嵌入表示 $\tilde{M}_s^{(l)} = (\tilde{m}_1^{s, (l)}, \tilde{m}_2^{s, (l)}, \dots, \tilde{m}_n^{s, (l)})$ ;

[0030] 其中,  $\tilde{M}_s^{(l)} = HWGCN(A^{M, (l-1)}, \tilde{M}_s^{(l-1)})$ ;

[0031] 上述式中:HWGCN表示多通道图卷积神经网络;1表示多通道图卷积神经网络的层数; $\tilde{H}_s^{(l-1)}$ 表示第1-1层HWGCN输出的局部嵌入表示; $\tilde{G}_s^{(l-1)}$ 表示第1-1层HWGCN输出的全局嵌入表示; $\tilde{M}_s^{(l-1)}$ 表示第1-1层HWGCN输出的共识嵌入表示; $\sigma$ 表示Sigmoid函数; $W_{\omega} \in \mathbb{R}^{2d}$ 表示可训练参数。

[0032] 优选的,所述会话推荐模型还包括图增强的位置编码器;

[0033] 所述位置编码器通过共识图 $A^{M, (1-1)}$ 计算对应的位置嵌入表示 $\tilde{P}_s^{(l)} = (\tilde{p}_1^{s, (l)}, \tilde{p}_2^{s, (l)}, \dots, \tilde{p}_n^{s, (l)})$ ;

[0034] 其中,  $\tilde{P}_s^{(l)} = HWGCN(A^{M, (l-1)}, \tilde{P}_s^{(l-1)})$ ;

[0035] 位置编码器的结构表示为:

[0036]  $\tilde{p}_i^s = p_i^s + p_{len}^s$ ;

[0037]  $p_{len}^s = Embed_{len}(S_{len})$ ;

[0038] 上述式中: $p_i^s \in \mathbb{R}^d$ 表示目标会话文本S中第i个位置的位置嵌入; $S_{len}$ 表示目标会话文本S的会话长度; $p_{len}^s \in \mathbb{R}^d$ 表示会话长度 $S_{len}$ 的嵌入表示。

[0039] 优选的,重复预测模块通过如下步骤计算重复预测概率:

[0040] S221:基于局部嵌入表示 $\tilde{H}_s^{(l)}$ 、全局嵌入表示 $\tilde{G}_s^{(l)}$ 、共识嵌入表示 $\tilde{M}_s^{(l)}$ 和位置嵌入表示 $\tilde{P}_s^{(l)}$ 计算对应的新会话表示 $X_s = (x_1^s, x_2^s, \dots, x_n^s)$ ;

[0041] 其中,  $x_i^s = \sigma(W_x (\tilde{h}_i^{s, (l)}, \tilde{g}_i^{s, (l)}, \tilde{m}_i^{s, (l)} + b_x)) + \tilde{p}_i^{s, (l)}$ ;

[0042] S222:基于新会话表示 $X_s$ 结合如下公式计算重复预测概率 $\hat{y}_i^r$ ;

[0043]  $\hat{y}_i^r = \begin{cases} z_i & \text{if } v_i \notin S, \\ 0 & \text{if } v_i \in S \end{cases}$ ;

[0044] 其中,  $z_i = \frac{\exp(W_r^T x_i^s)}{\sum_{j=1}^n \exp(W_r^T x_j^s)}$ ;

[0045] 上述式中: $W_x \in \mathbb{R}^{d \times 3d}$ 、 $b_x \in \mathbb{R}^d$ 、 $W_r \in \mathbb{R}^d$ 表示可训练参数; $x_j^s$ 表示聚合了会话中第j个项目 $h_j^s$ 的局部表示、全局表示、共识表示以及其位置嵌入后的综合表示。

[0046] 优选的,探索预测模块包括用于消除噪声项目影响的图增强注意力网络;图增强注意力网络通过如下步骤消除噪声项目的影响:

[0047] S231:基于局部嵌入表示 $\tilde{H}_s^{(l)}$ 建立对应的错误点击集合 $S_{wrong}^L$ ,通过错误点击集合 $S_{wrong}^L$ 结合位置嵌入表示 $\tilde{P}_s^{(l)}$ 计算对应的长期偏好 $S_{long}^L$ ,结合长期偏好 $S_{long}^L$ 和局部嵌入表示 $\tilde{H}_s^{(l)}$ 中的当前偏好 $\tilde{h}_n^{s, (l)}$ 计算对应的会话局部表示 $\hat{S}_{long}^L$ ;



[0048] 其中,  $s_{long}^L = \sum_{i=1}^n \lambda_i \tilde{h}_i^{s,(L)}$ ,  $\lambda_i = \frac{\exp(\beta_i)}{\sum_{j=1}^n \exp(\beta_j)}$ ,  $\beta_i = \begin{cases} W_e^T (W_2 \tilde{h}_i^{s,(L)} + W_3 \tilde{p}_i^{s,(L)}), & \text{if } \tilde{h}_i^{s,(L)} \notin (\tilde{H}_s^L - S_{wrong}^L) \\ -\infty, & \text{if } \tilde{h}_i^{s,(L)} \in S_{wrong}^L \end{cases}$ ;

[0049] S232: 基于全局嵌入表示  $\tilde{G}_s^{(l)}$  建立对应的错误点击集合  $S_{wrong}^G$ , 通过错误点击集合  $S_{wrong}^G$  结合位置嵌入表示  $\tilde{P}_s^{(l)}$  计算对应的长期偏好  $S_{long}^G$ , 结合长期偏好  $S_{long}^G$  和全局嵌入表示  $\tilde{G}_s^{(l)}$  中的当前偏好  $\tilde{g}_n^{s,(l)}$  计算对应的会话全局表示  $\hat{S}_{long}^G$ ;

[0050] S233: 基于共识嵌入表示  $\tilde{M}_s^{(l)}$  建立对应的错误点击集合  $S_{wrong}^M$ , 通过错误点击集合  $S_{wrong}^M$  结合位置嵌入表示  $\tilde{P}_s^{(l)}$  计算对应的长期偏好  $S_{long}^M$ , 结合长期偏好  $S_{long}^M$  和共识嵌入表示  $\tilde{M}_s^{(l)}$  中的当前偏好  $\tilde{m}_n^{s,(l)}$  计算对应的会话共识表示  $\hat{S}_{long}^M$ ;

[0051] S234: 结合会话局部表示  $\hat{S}_{long}^L$ 、会话全局表示  $\hat{S}_{long}^G$  和会话共识表示  $\hat{S}_{long}^M$  计算对应的会话结构表示  $s_{long}^{str}$ ;

[0052] 其中,  $s_{long}^{str} = \hat{S}_{long}^L + \hat{S}_{long}^G + \hat{S}_{long}^M$ ;

[0053] 上述式中:  $W_e \in \mathbb{R}^d$ 、 $W_2 \in \mathbb{R}^{d \times d}$ 、 $W_3 \in \mathbb{R}^d$  表示可训练参数。

[0054] 优选的, 探索预测模块通过如下步骤计算探索预测概率:

[0055] S241: 通过Bi-GRU模块结合局部嵌入表示  $\tilde{H}_s^{(l)}$ 、全局嵌入表示  $\tilde{G}_s^{(l)}$  和共识嵌入表示  $\tilde{M}_s^{(l)}$  计算对应的会话顺序表示  $s_{long}^{seq}$ ;

[0056] 其中,  $s_{long}^{seq} = d_n^{GRU}$ ;

[0057]  $d_i^{GRU} = \text{LayerNorm}(W_{gru}([\tilde{d}_i^{GRU}; \tilde{d}_i^{GRU}]))$ ;

[0058]  $\tilde{d}_i^{GRU} = \text{GRU}_{forward}(\tilde{d}_{i-1}^{GRU}, d_i, \phi_{gru})$ ;

[0059]  $\tilde{d}_i^{GRU} = \text{GRU}_{backward}(\tilde{d}_{i+1}^{GRU}, d_i, \phi_{gru})$ ;

[0060]  $d_i = W_{seq}([\tilde{h}_i^{s,(l)}; \tilde{g}_i^{s,(l)}; \tilde{m}_i^{s,(l)}]) + b_{seq}$ ;

[0061] S242: 基于会话结构表示  $s_{long}^{str}$  和会话顺序表示  $s_{long}^{seq}$  计算对应的探索预测概率  $\hat{y}_i^e$ ;

$$\hat{y}_i^e = \frac{\exp(\rho_i)}{\sum_{i=1}^{|V|} \exp(\rho_i)}$$

[0062] 其中,  $\rho_i = \begin{cases} s_e^T h_i, & \text{if } v_i \notin S \\ -\infty, & \text{if } v_i \in S \end{cases}$ ,  $s_e = s_{long}^{str} + s_{long}^{seq}$ ;

[0063] 上述式中:  $W_{seq} \in \mathbb{R}^{d \times 3d}$ 、 $W_{gru} \in \mathbb{R}^{d \times 2d}$ 、 $b_{seq} \in \mathbb{R}^d$ 、 $\phi_{gru}$  表示可训练参数; LayerNorm表示归一化;  $h_i \in \mathbb{R}^d$  表示候选项目集合V中第i个项目的嵌入表示。

[0064] 优选的, 判断模块通过如下公式计算最终预测概率;

[0065]  $\hat{y}_i = p \cdot \hat{y}_i^r + (1-p) \cdot \hat{y}_i^e$ ;

[0066] 其中,  $p = \delta(W_p^T([\mathbf{o}_r; \mathbf{s}_e]))$ ,  $\mathbf{o}_r = \sum_{i=1}^n z_i \mathbf{x}_i^s$ ,  $\mathbf{s}_e = s_{long}^{str} + s_{long}^{seq}$ ;

[0067] 上述式中:  $W_p \in \mathbb{R}^d$  表示可训练参数;  $\mathbf{o}_r$  和  $\mathbf{s}_e$  分别表示重复预测模块和探索预测模块

的会话表示; $\delta$ 表示超参数。

[0068] 本发明中的会话推荐方法与现有技术相比,具有如下有益效果:

[0069] 在本发明中,通过会话图的结构能够准确识别会话文本中的噪声项目,进而通过调整注意机制及其注意力权重的方式消除噪声项目的影响,使得能够得到会话文本中的长期结构表示,避免模型学习不准确的用户偏好表示,从而能够提升会话项目推荐的准确性。同时,本发明利用项目相似性(而不是项目转换模式)能够捕获项目之间的高质量关系,使得能够减少由用户意外或错误点击造成的噪音信息。此外,本发明从三个角度分别获取了局部嵌入表示、全局嵌入表示和共识嵌入表示,使得能够增强会话文本的表示性,从而能够更好的提升会话项目推荐的全面性。最后,本发明中通过“重复和探索”的机制能够很好的计算得到符合用户偏好的概率分布,即能够更精确的进行项目分布概率的预测,从而能够提升会话项目推荐预测的准确性。

## 附图说明

[0070] 为了使发明的目的、技术方案和优点更加清楚,下面将结合附图对本发明作进一步的详细描述,其中:

[0071] 图1为图卷积神经网络会话推荐方法的逻辑框图;

[0072] 图2为会话推荐模型的网络结构图;其中,(A)表示重复预测模块和探索预测模块,(B)表示动态全局邻居注意力网络,(C)表示多通道图卷积神经网络,(D)表示图增强的位置编码器;

[0073] 图3为DGS-MGNN、DGS-MGNN w/o GPE和DGS-MGNN-POS的性能示意图;

[0074] 图4为DGS-MGNN、DGS-MCGNN-STR和DGS-MCGNN-SEQ的性能示意图;

[0075] 图5为不同全局邻居数量时DGS-MGNN的性能示意图;

[0076] 图6为DGS-MGNN、DSAN和GCE-GNN的性能示意图。

## 具体实施方式

[0077] 下面通过具体实施方式进一步详细的说明:

[0078] 实施例:

[0079] 本实施例中公开了一种基于结构增强的图卷积神经网络会话推荐方法。

[0080] 基于结构增强的图卷积神经网络会话推荐方法:首先获取目标会话文本的会话表示;然后基于会话表示生成对应的会话图,进而通过会话图识别目标会话文本中的噪声项目;再结合注意力机制重置噪声项目的注意力权重,以消除噪声项目的影响;最后计算最终预测概率分布,并基于最终预测概率分布进行项目推荐。具体的,将噪声项目的注意力权重设置为0。

[0081] 结合图1所示,具体包括以下步骤:

[0082] S1:获取目标会话文本;

[0083] S2:将目标会话文本输入经过预先训练的会话推荐模型中;

[0084] 结合图2所示,会话推荐模型首先通过融合嵌入模块从三个角度分别提取目标会话文本的局部嵌入表示、全局嵌入表示和共识嵌入表示;再通过重复预测模块计算对应的重复预测概率;然后通过探索预测模块识别目标会话文本中的噪声项目、消除噪声项目的

影响,并计算对应的探索预测概率;最后通过判断模块基于重复预测概率和探索预测概率计算对应的最终预测概率分布;

[0085] S3:基于会话推荐模型输出的最终预测概率分布进行项目推荐。

[0086] 在本发明中,通过会话图的结构能够准确识别会话文本中的噪声项目,进而通过调整注意机制及其注意力权重的方式消除噪声项目的影响,使得能够得到会话文本中的长期结构表示,避免模型学习不准确的用户偏好表示,从而能够提升会话项目推荐的准确性。同时,本发明利用项目相似性(而不是项目转换模式)能够捕获项目之间的高质量关系,使得能够减少由用户意外或错误点击造成的噪音信息。此外,本发明从三个角度分别获取了局部嵌入表示、全局嵌入表示和共识嵌入表示,使得能够增强会话文本的表示性,从而能够更好的提升会话项目推荐的全面性。最后,本发明中通过“重复和探索”的机制能够很好的计算得到符合用户偏好的概率分布,即能够更精确的进行项目分布概率的预测,从而能够提升会话项目推荐预测的准确性。

[0087] 具体实施过程中,融合嵌入模块包括用于基于目标会话文本生成全局表示的动态全局邻居注意力网络(Dynamic Global Attention,DGNA),以及用于基于目标会话文本的局部表示和全局表示融合生成对应的局部嵌入表示、全局嵌入表示和共识嵌入表示的多通道图卷积神经网络(Multi-Channel Graph Neural Network,MC-GNN)。

[0088] 具体的,动态全局邻居注意力网络通过如下步骤生成全局表示:

[0089] S201:对目标会话文本  $S=(v_1^s, v_2^s, \dots, v_n^s)$  进行编码,得到对应的局部表示  $H_s=(h_1^s, h_2^s, \dots, h_n^s)$ ;

[0090] S202:从候选项目集合  $V=\{v_1, v_1, \dots, v_{|V|}\}$  中提取K个候选项目  $v_j \in V$  作为目标项目  $v_i^s \in S$  的全局邻居  $N^i=(n_1^i, n_2^i, \dots, n_k^i)$ ;选取全局邻居时,计算目标项目  $v_i^s$  与各个候选项目  $v_j$  的余弦相似度,然后选取余弦相似度最高的K个候选项目  $v_j$  来作为全局邻居;

[0091] 其中,  $r_{i,j} = \frac{h_i^{sT} \cdot h_j}{\|h_i^s\| \|h_j\|}$ ;  $r_{i,j}$  表示目标项目  $v_i^s$  与候选项目  $v_j$  间的余弦相似度。

[0092] S203:基于全局邻居  $N^i=(n_1^i, n_2^i, \dots, n_k^i)$  生成对应的全局邻居嵌入表示  $C^i=(c_1^i, c_2^i, \dots, c_k^i)$ ;

[0093] S204:通过注意力池化机制聚合目标项目  $v_i^s$  的全局邻居嵌入表示  $C^i=(c_1^i, c_2^i, \dots, c_k^i)$  得到目标项目  $v_i^s$  的全局表示  $g_i^s$ ;

[0094] 其中,  $g_i^s = \sum_{j=1}^K \partial_j^i c_j^i$ ,  $\partial_j^i = \frac{\exp(W_g^T [c_j^i; h_i^s])}{\sum_{j=1}^K \exp(W_g^T [c_j^i; h_i^s])}$ ;

[0095] S205:计算各个目标项目  $v_i^s$  的会话全局表示  $g_i^s$ ,并生成目标会话文本S的全局表示  $G_s=(g_1^s, g_2^s, \dots, g_n^s)$ ;

[0096] 上述式中:  $W_g \in \mathbb{R}^{2d}$  表示可训练参数;  $[\cdot]$  表示拼接操作。

[0097] 多通道图卷积神经网络通过如下步骤生成局部嵌入表示、全局嵌入表示和共识嵌入表示:

[0098] S211: 基于局部表示  $\mathbf{H}_s = (\mathbf{h}_1^s, \mathbf{h}_2^s, \dots, \mathbf{h}_n^s)$  构造对应的局部图  $A^{L, (1-1)}$ , 并通过局部图  $A^{L, (1-1)}$  计算对应的局部嵌入表示  $\tilde{\mathbf{H}}_s^{(l)} = (\tilde{h}_1^{s,(l)}, \tilde{h}_2^{s,(l)}, \dots, \tilde{h}_n^{s,(l)})$ ;

[0099] 其中,  $\tilde{\mathbf{H}}_s^{(l)} = \text{HWGCN}(A^{L,(l-1)}, \tilde{\mathbf{H}}_s^{(l-1)})$ ,  $A^{L,(l-1)} = \text{BSG}(\tilde{\mathbf{H}}_s^{(l-1)})$ ;

[0100] S212: 基于全局表示  $\mathbf{G}_s = (\mathbf{g}_1^s, \mathbf{g}_2^s, \dots, \mathbf{g}_n^s)$  构造对应的全局图  $A^{G, (1-1)}$ , 并通过全局图  $A^{G, (1-1)}$  计算对应的全局嵌入表示  $\tilde{\mathbf{G}}_s^{(l)} = (\tilde{g}_1^{s,(l)}, \tilde{g}_2^{s,(l)}, \dots, \tilde{g}_n^{s,(l)})$ ;

[0101] 其中,  $\tilde{\mathbf{G}}_s^{(l)} = \text{HWGCN}(A^{G,(l-1)}, \tilde{\mathbf{G}}_s^{(l-1)})$ ;  $A^{G,(l-1)} = \text{BSG}(\tilde{\mathbf{G}}_s^{(l-1)})$ ;

[0102] S213: 基于局部图  $A^{L, (1-1)}$  和全局图  $A^{G, (1-1)}$  融合生成对应的共识图  $A^M, (1-1)$ ;

[0103] 其中,  $A^{M,(l-1)} = f_i \cdot A_i^{L,(l-1)} + (1-f_i) \cdot A_i^{G,(l-1)}$ ,  $f_i = \sigma(\mathbf{W}_{\infty}^T [\tilde{h}_i^{s,(l)}, \tilde{g}_i^{s,(l)}])$ ;

[0104] S214: 通过共识图  $A^M, (1-1)$  计算对应的共识嵌入表示  $\tilde{\mathbf{M}}_s^{(l)} = (\tilde{m}_1^{s,(l)}, \tilde{m}_2^{s,(l)}, \dots, \tilde{m}_n^{s,(l)})$ ;

[0105] 其中,  $\tilde{\mathbf{M}}_s^{(l)} = \text{HWGCN}(A^{M,(l-1)}, \tilde{\mathbf{M}}_s^{(l-1)})$ ;

[0106] 上述式中: HWGCN表示多通道图卷积神经网络; 1表示多通道图卷积神经网络的层数;  $\tilde{\mathbf{H}}_s^{(l-1)}$ 表示第1-1层HWGCN输出的局部嵌入表示;  $\tilde{\mathbf{G}}_s^{(l-1)}$ 表示第1-1层HWGCN输出的全局嵌入表示;  $\tilde{\mathbf{M}}_s^{(l-1)}$ 表示第1-1层HWGCN输出的共识嵌入表示;  $\sigma$ 表示Sigmoid函数;  $\mathbf{W}_{\infty} \in \mathbb{R}^{2d}$ 表示可训练参数。具体的, 通过高速公路网络 (highway network) 来缓解堆叠图卷积神经网络带来过度平滑的问题, 将第1层的图传播过程定义如下:

[0107]  $\tilde{h}_i^{s,(l)} = g_i \cdot h_i^{s,(l)} + (1-g_i) \cdot h_i^{s,(0)}$

[0108]  $h_i^{s,(l)} = A_i^{L,(l-1)}(\tilde{h}_1^{s,(l-1)}, \tilde{h}_2^{s,(l-1)}, \dots, \tilde{h}_n^{s,(l-1)})$ ;

[0109]  $g_i = \sigma(\mathbf{W}_1^T [h_i^{s,(l)}, h_i^{s,(0)}])$ ;

[0110] 式中:  $A_i^{L,(l-1)}$ 表示  $A^{L, (1-1)}$  的第i行;  $\mathbf{W}_1 \in \mathbb{R}^{2d}$ 表示可训练参数。

[0111] BSG表示构建相似度图的过程 (Build Similar Graph), 以局部表示  $\mathbf{H}_s = (\mathbf{h}_1^s, \mathbf{h}_2^s, \dots, \mathbf{h}_n^s)$  为例, 构建其对应相似度图的过程定义如下:

[0112] 
$$\tau_{i,j} = \frac{\mathbf{h}_i^{s,T} \cdot \mathbf{h}_j^s}{\|\mathbf{h}_i^s\| \|\mathbf{h}_j^{s,T}\|}$$
;

[0113] 
$$e_{i,j} = \begin{cases} 0, & \text{if } \tau_{i,j} \leq 0, \\ \frac{\exp(\tau_{i,j})}{\sum_{j=1}^n \exp(\tau_{i,j})}, & \text{if } \tau_{i,j} > 0; \end{cases}$$

[0114] 其中  $\mathbf{H}_s = (\mathbf{h}_1^s, \mathbf{h}_2^s, \dots, \mathbf{h}_n^s)$  中每个项目表示  $h_i^s$  表示图中的一个节点,  $e_{i,j}$  表示节点  $h_i^s$  和节点  $h_j^s$  的连接强度。同时对图中的噪音边进行了过滤操作 (即:  $\tau_{i,j} \leq 0$  则将节点  $h_i^s$  和节点  $h_j^s$  之间的连接强度置为0)。

[0115] 在本发明中, 基于局部表示和全局表示构造了对应的局部图和全局图, 进而基于局部图和全局图融合生成了共识图, 即能够以动态的方式构造会话图, 使得能够通过共识图维护局部图和全局图间的一致性。同时, 本发明从三个角度分别获取了局部嵌入表示、全

局嵌入表示和共识嵌入表示,使得能够增强会话文本的表示性,从而能够更好的提升会话项目推荐的全面性。

[0116] 具体实施过程中,会话推荐模型还包括图增强的位置编码器(Graph Position Encoder,GPE);

[0117] 位置编码器通过共识图 $A^{M,(l-1)}$ 计算对应的位置嵌入表示 $\tilde{P}_s^{(l)} = (\tilde{p}_1^{s,(l)}, \tilde{p}_2^{s,(l)}, \dots, \tilde{p}_n^{s,(l)})$ ;

[0118] 其中, $\tilde{P}_s^{(l)} = HWGCN(A^{M,(l-1)}, \tilde{P}_s^{(l-1)})$ ;

[0119] 位置编码器的结构表示为:

[0120]  $\tilde{P}_i^s = P_i^s + P_{len}^s$ ;

[0121]  $P_{len}^s = Emebed_{len}(S_{len})$ ;

[0122] 上述式中: $P_i^s \in \mathbb{R}^d$ 表示目标会话文本S中第i个位置的位置嵌入; $S_{len}$ 表示目标会话文本S的会话长度; $P_{len}^s \in \mathbb{R}^d$ 表示会话长度 $S_{len}$ 的嵌入表示。 $Emebed_{len}$ 表示一种针对会话长度的向量嵌入操作。例如会话长度为 $S_{len} = 6$ 的会话将其长度6映射到一个向量 $P_{len}^s \in \mathbb{R}^d$ ,该向量是可训练参数。

[0123] 传统建模位置信息的方法通常采用绝对位置编码的方式,其难以很好的适应会话推荐任务。例如:会话 $S_1 = (\text{苹果手机}, \text{耳机}, \text{短袖}, \text{短裙}, \text{毛衣})$ ,会话 $S_2 = (\text{小米手机}, \text{耳机})$ 。会话 $S_1$ 和会话 $S_2$ 的第二个项目都是耳机,然而 $S_2$ 中的耳机更能体现用户的当前偏好,这是由于耳机分别位于 $S_2$ 的最后一个位置和 $S_1$ 的倒数第四个位置,而现有绝对位置编码会忽略掉这种会话的个性化位置信息。

[0124] 为此,本发明通过整合会话长度信息以及捕获会话固有的拓扑位置结构来增强会话中每个项目的位置嵌入,使得模型能够更好的关注到会话的个性化位置信息,从而能够辅助提升会话项目的推荐准确性和全面性。

[0125] 具体实施过程中,重复预测模块通过如下步骤计算重复预测概率:

[0126] S221:基于局部嵌入表示 $\tilde{H}_s^{(l)}$ 、全局嵌入表示 $\tilde{G}_s^{(l)}$ 、共识嵌入表示 $\tilde{M}_s^{(l)}$ 和位置嵌入表示 $\tilde{P}_s^{(l)}$ 计算对应的新会话表示 $X_s = (x_1^s, x_2^s, \dots, x_n^s)$ ;

[0127] 其中, $x_i^s = \sigma(W_x(\tilde{h}_i^{s,(L)}, \tilde{g}_i^{s,(L)}, \tilde{m}_i^{s,(L)}) + b_x) + \tilde{p}_i^{s,(L)}$ ;

[0128] S222:基于新会话表示 $X_s$ 结合如下公式计算重复预测概率 $\hat{y}_i^r$ ;

[0129] 
$$\hat{y}_i^r = \begin{cases} z_i & \text{if: } v_i \notin S, \\ 0 & \text{if: } v_i \in S \end{cases}$$
;

[0130] 其中, $z_i = \frac{\exp(W_r^T x_i^s)}{\sum_{j=1}^n \exp(W_r^T x_j^s)}$ ;

[0131] 上述式中: $W_x \in \mathbb{R}^{d \times 3d}$ 、 $b_x \in \mathbb{R}^d$ 、 $W_r \in \mathbb{R}^d$ 表示可训练参数; $x_j^s$ 表示聚合了会话中第j个项目 $h_j^s$ 的局部表示、全局表示、共识表示以及其位置嵌入后的综合表示。

[0132] 在本发明中,通过重复预测模块能够基于局部嵌入表示、全局嵌入表示、共识嵌入表示和位置嵌入表示计算当前会话中出现过的项目的概率分布,即重复预测概率,这能够很好的辅助计算最终预测概率分布。

[0133] 具体实施过程中,探索预测模块包括用于消除噪声项目影响的图增强注意力网络(Graph-Enhanced Attention Network,GEA);图增强注意力网络通过如下步骤消除噪声项目的影响:

[0134] S231:基于局部嵌入表示 $\tilde{H}_s^{(l)}$ 建立对应的错误点击集合 $S_{wrong}^L$ ,通过错误点击集合 $S_{wrong}^L$ 结合位置嵌入表示 $\tilde{P}_s^{(l)}$ 计算对应的长期偏好 $S_{long}^L$ ,结合长期偏好 $S_{long}^L$ 和局部嵌入表示 $\tilde{H}_s^{(l)}$ 中的当前偏好 $\tilde{h}_n^{s,(l)}$ 计算对应的会话局部表示 $\hat{S}_{long}^L$ ;

[0135] 其中, $s_{long}^L = \sum_{i=1}^n \lambda_i \tilde{h}_i^{s,(L)}$ ,  $\lambda_i = \frac{\exp(\beta_i)}{\sum_{j=1}^n \exp(\beta_j)}$ ,  $\beta_i = \begin{cases} W_e^T (W_2 \tilde{h}_i^{s,(L)} + W_3 \tilde{p}_i^{s,(L)}), & \text{if: } \tilde{h}_i^{s,(L)} \notin (\tilde{H}_s^{(L)} - S_{wrong}^L) \\ -\infty, & \text{if: } \tilde{h}_i^{s,(L)} \in S_{wrong}^L \end{cases}$ ;

[0136] S232:基于全局嵌入表示 $\tilde{G}_s^{(l)}$ 建立对应的错误点击集合 $S_{wrong}^G$ ,通过错误点击集合 $S_{wrong}^G$ 结合位置嵌入表示 $\tilde{P}_s^{(l)}$ 计算对应的长期偏好 $S_{long}^G$ ,结合长期偏好 $S_{long}^G$ 和全局嵌入表示 $\tilde{G}_s^{(l)}$ 中的当前偏好 $\tilde{g}_n^{s,(l)}$ 计算对应的会话全局表示 $\hat{S}_{long}^G$ ;

[0137] S233:基于共识嵌入表示 $\tilde{M}_s^{(l)}$ 建立对应的错误点击集合 $S_{wrong}^M$ ,通过错误点击集合 $S_{wrong}^M$ 结合位置嵌入表示 $\tilde{P}_s^{(l)}$ 计算对应的长期偏好 $S_{long}^M$ ,结合长期偏好 $S_{long}^M$ 和共识嵌入表示 $\tilde{M}_s^{(l)}$ 中的当前偏好 $\tilde{m}_n^{s,(l)}$ 计算对应的会话共识表示 $\hat{S}_{long}^M$ ;

[0138] S234:结合会话局部表示 $\hat{S}_{long}^L$ 、会话全局表示 $\hat{S}_{long}^G$ 和会话共识表示 $\hat{S}_{long}^M$ 计算对应的会话结构表示 $s_{long}^{str}$ ;

[0139] 其中, $s_{long}^{str} = \hat{S}_{long}^L + \hat{S}_{long}^G + \hat{S}_{long}^M$ ;

[0140] 上述式中: $W_e \in \mathbb{R}^d$ 、 $W_2 \in \mathbb{R}^{d \times d}$ 、 $W_3 \in \mathbb{R}^d$ 表示可训练参数。

[0141] 探索预测模块通过如下步骤计算探索预测概率:

[0142] S241:通过Bi-GRU模块结合局部嵌入表示 $\tilde{H}_s^{(l)}$ 、全局嵌入表示 $\tilde{G}_s^{(l)}$ 和共识嵌入表示 $\tilde{M}_s^{(l)}$ 计算对应的会话顺序表示 $s_{long}^{seq}$ ;

[0143] 其中, $s_{long}^{seq} = d_n^{GRU}$ ;

[0144]  $d_i^{GRU} = LayerNorm(W_{gru}([\bar{d}_i^{GRU}; \bar{d}_i^{GRU}]))$ ;

[0145]  $\bar{d}_i^{GRU} = GRU_{forward}(\bar{d}_{i-1}^{GRU}, d_i, \phi_{gru})$ ;

[0146]  $\bar{d}_i^{GRU} = GRU_{forward}(\bar{d}_{i+1}^{GRU}, d_i, \phi_{gru})$ ;

[0147]  $d_i = W_{seq}([\tilde{h}_i^{s,(l)}; \tilde{g}_i^{s,(l)}; \tilde{m}_i^{s,(l)}]) + b_{seq}$ ;

[0148] S242:基于会话结构表示 $s_{long}^{str}$ 和会话顺序表示 $s_{long}^{seq}$ 计算对应的探索预测概率 $\hat{y}_i^e$ ;

$$\hat{y}_i^e = \frac{\exp(\rho_i)}{\sum_{i=1}^{|\mathcal{V}|} \exp(\rho_i)}$$

[0149] 其中, $\rho_i = \begin{cases} s_e^T h_i, & \text{if: } v_i \notin S \\ -\infty, & \text{if: } v_i \in S \end{cases}$ ,  $s_e = s_{long}^{str} + s_{long}^{seq}$ ;

[0150] 上述式中： $W_{seq} \in \mathbb{R}^{d \times 3d}$ 、 $W_{gru} \in \mathbb{R}^{d \times 2d}$ 、 $b_{seq} \in \mathbb{R}^d$ 、 $\phi_{gru}$  表示可训练参数；LayerNorm表示归一化； $h_i \in \mathbb{R}^d$ 表示候选项目集合V中第i个项目的嵌入表示。forward表示GRU的前向计算，backward表示GRU的反向计算。

[0151] 在本发明中,通过探索预测模块能够基于局部嵌入表示、全局嵌入表示、共识嵌入表示和位置嵌入表示分别计算得到会话结构表示和会话顺序表示,进而能够从会话结构和会话顺序两个方面去预测概率分布,这同样能够很好的辅助计算最终预测概率分布。

[0152] 具体实施过程中,判断模块通过如下公式计算最终预测概率;

$$[0153] \quad \hat{y}_i = p \cdot \hat{y}_i^r + (1-p) \cdot \hat{y}_i^e;$$

$$[0154] \quad \text{其中, } p = \delta(W_p^T([\mathbf{o}_r; \mathbf{s}_e])), \quad \mathbf{o}_r = \sum_{i=1}^n z_i \mathbf{x}_i^s, \quad \mathbf{s}_e = \mathbf{s}_{long}^{str} + \mathbf{s}_{long}^{seq};$$

[0155] 上述式中： $W_p \in \mathbb{R}^d$  表示可训练参数； $\mathbf{o}_r$ 和 $\mathbf{s}_e$ 分别表示重复预测模块和探索预测模块的会话表示； $\delta$ 表示超参数。

[0156] 为了更好的说明本发明中会话推荐方法的优势,本实施例中公开了如下实验。

[0157] 一、数据集

[0158] 本实验在三个广泛使用的基准数据集 (Diginetica, Yoochoose, Retailrocket) 上对DGS-MGNN和一系列基线模型的性能进行了测试。

[0159] Diginetica:来自于2016年CIKM Cup挑战赛。由于其包含商品交易类型的数据,所以经常被用于会话推荐任务。本实验提取其最后一周的数据作为测试数据。

[0160] Yoochoose:来自于2015年的ResSys挑战赛,其包含了6个月内电子商务网站的点击流数据。由于该数据量过大,提取按时间排序的最近的1/4和1/64的数据作为训练集,它们分别被命名未“Yoochoose1/4”和“Yoochoose1/64”,最后一周的数据作为测试集。

[0161] Retailrocket:来自于2016年的Kaggle竞赛,其包含了用户在电子商务网站上4.5个月的行为数据。本实验提取最近的1/4数据作为训练集,最后15天的数据作为测试集。

[0162] 在三个数据集中,将会话长度小于2的会话和项目出现次数小于5的项目进行过滤。

[0163] 二、基线模型

[0164] 为了全面的检验DGS-MGNN(本发明提出的会话推荐模型)的性能,本实验将其与三种类型的基线模型进行了性能对比(传统推荐方法、基于循环设计网络和注意力机制的方法、基于图神经网络的方法),详细的介绍如下:

[0165] 1) 传统推荐方法:

[0166] POP:这是推荐系统中常用的一种基线方法,它推荐训练集中出现频率最高的前N个项目。

[0167] Item-KNN:是一种基于协同过滤的方法,其通过向用户推荐与当前会话最相似的项目。

[0168] FPMC:将矩阵分解和马尔科夫链结合起来,其中序列数据由转移矩阵建模,所有转移矩阵都是用户特定的。它引入了一个因子分解模型,该模型给出了转换立方体的低秩近似值,其中每一个部分都是用户历史点击在马尔科夫链下的转移矩阵。

[0169] 2) 基于循环神经网络和注意力机制的方法:

[0170] GRU4REC:利用门控循环神经网络GRU模拟用户的顺序行为并采用并行小批次训练方案进行模型训练。

[0171] NARM:使用循环神经网络RNN来建模用户的顺序行为并结合注意力机制来捕获用户的主要偏好。同时,它结合双线性匹配机制为每个候选项目生成推荐概率。

[0172] STAMP:通过捕获用户的长期偏好和短期兴趣来缓解用户的偏好转移的问题。

[0173] CSR:提出利用协作邻域信息进行基于会话的推荐。它利用内部编码器捕获当前会话的信息,同时它也利用外部编码器捕获邻域会话的协作信息。

[0174] DSAN:将双稀疏注意力网络应用于会话推荐任务。它首先探究了会话内每个项目的间的关系并结合自注意力网络目标项目表示,然后利用普通的注意力网络来捕获项目的重要性,最后利用神经网络将目标表示和会话表示结合起来从而得到最终的表示。

[0175] 3) 基于图神经网络的方法:

[0176] SR-GNN:通过将会话序列建模为会话图捕获项目在会话中复杂的转换关系。同时,它还结合门控神经网络和自注意力机制来生成会话表示。

[0177] TAGNN:通过会话序列建模为会话图并通过图神经网络获取项目的嵌入表示,它还引入了目标感知模块,以揭示给定目标项目与所有候选项目的相关性,从而提升会话表示质量。

[0178] GCE-GNN:是目前性能最好的模型,它通过2个不同的视角学习项目的表示,例如:会话视角和全局视角。会话视角旨在通过会话内项目的转换关系学习项目的表示,全局视角旨在通过项目在所有会话中的转换关系学习项目的表示。

[0179] 三、实验参数设置

[0180] 本实验在DGS-MGNN的所有实验中,设置训练批次大小为256,项目嵌入的向量维度为256。全局邻居的数量K设置为10,HWGCN的层数为1。根据基线模型的原始论文设置其实验参数。在实验中的模型参数初始化按照均值为0,方差为0.1进行初始化。使用Adam优化器并配备0.001的学习率,该学习率会每训练3轮衰减为之前的0.1倍。 $\delta$ 设置为0.4,同时dropout也在模型中使用,其值设定为0.4。另一方面,利用 $L_2$ 正则化来避免过拟合,其值设定为 $1e-4$ 。

[0181] 四、评测指标

[0182] 本实验使用2个广泛使用的评测指标P@20和MRR@20来评估所有模型的性能,P@K和MRR@K的值越高,代表模型性能越好。

[0183] P@K (Precision):它衡量目标项目在top-K推荐中排名时的数量比例,是评估未排名结果的指标。

[0184] 其中, $P@K = \frac{n_{hit}}{N}$ ;

[0185] 式中:N表示测试集数量; $n_{hit}$ 表示目标项目在预测的top-K列表中的样本数量。

[0186] MRR@K (Mean Reciprocal Rank):它是目标项在推荐列表中的倒数排名的平均值。此指标考虑正确推荐项目在排名列表中的位置。

[0187] 其中, $MRR@K = \frac{1}{N} \sum_{i=1}^N \frac{1}{rank_i}$ ;

[0188] 式中:N表示测试集数量; $rank_i$ 表示第i个目标项目在推荐列表中的位置。若目标



项目未在top-K推荐列表中,则MRR@K为0。

[0189] 五、整体实验

[0190] 表1展示了DGS-MGNN和基线模型在三个基准数据集上的性能测试实验结果。

[0191] 表1

[0192]

Dataset Methods	Diginetica		Yoochoose1/64		Yoochoose1/4		Retailrocket	
	P@20	MRR@20	P@20	MRR@20	P@20	MRR@20	P@20	MRR@20
POP	1.18	0.28	6.71	1.65	1.37	0.31	1.97	0.75
Item-KNN	35.75	11.57	51.60	21.81	52.31	21.70	10.23	3.96
FPMC	22.14	6.66	45.62	15.01	51.86	17.50	9.65	4.32
GRU4REC	30.79	8.22	60.64	22.89	59.53	22.60	41.35	25.54
NARM	48.32	16.00	68.32	28.63	69.73	29.23	59.46	41.48
STAMP	46.62	15.13	68.74	29.67	70.44	30.00	58.48	38.96
CSRM	50.55	16.38	69.85	29.71	70.63	29.68	61.09	40.28
DSAN	53.76	18.99	69.68	<u>31.23</u>	<u>72.06</u>	<u>32.22</u>	62.56	<u>42.39</u>
SR-GNN	51.26	17.78	70.57	30.94	71.36	31.89	60.19	39.64
TAGNN	51.31	18.03	<u>71.02</u>	31.12	71.32	32.11	59.31	39.65
GCE-GNN	<u>54.22</u>	<u>19.04</u>	70.91	30.63	71.40	31.49	<u>63.29</u>	40.35
<b>DGS-MGNN</b>	<b>55.54*</b>	<b>19.68*</b>	<b>72.62*</b>	<b>32.49*</b>	<b>73.10*</b>	<b>33.55*</b>	<b>66.61*</b>	<b>45.78*</b>

[0193] 从表1中可以看出,本发明提出的DGS-MGNN始终优于所有最先进的基线方法。具体来说,DGS-MGNN相比较最优基线GCE-GNN表现出了显著的改善( $t$ 检验, $p \leq 0.01$ )并在Diginetica、Yoochoose和Retailrocket的MRR@20上高出了3.36%,6.07%和13.46%。同样的性能优势在P@20上也能得到体现。DGS-MGNN取得性能优势的主要原因是其可以动态的聚合全局信息。

[0194] 本发明提出了一个MC-GNN(多通道图卷积神经网络)来学习项目不同视角的丰富的信息。此外,DGS-MGNN结合图曾庆的GPE(图增强的位置编码器)和GEA(图增强的注意力网络)来分别增强项目的位置嵌入和过滤会话中的噪音项目。

[0195] 六、DGNA和MC-GNN的影响

[0196] 为了验证动态全局邻居注意力网络(DGNA)和多通道图卷积神经网络(MC-GNN)的有效性,本实验设计了相关的变体模型:

[0197] DGS-MGNN w/o DGNA:从DGS-MGNN中将动态全局邻居注意力网络(DGNA)移除。

[0198] DGS-MGNN-MLP:将DGS-MGNN中的多通道图神经网络(MC-GNN)替换为多层感知机(MLP)。

[0199] DGS-MGNN-GGNN:将DGS-MGNN中的多通道图神经网络(MC-GNN)替换为SR-GNN中所使用的门控图神经网络(GGNN),同时其图的构造方法于SR-GNN中一致。

[0200] DGS-MGNN-GAT:将DGS-MGNN中的多通道图神经网络(MC-GNN)替换为GCE-GNN中所使用的图注意力网络(GAT),同时其图的构造方法于GCE-GNN中一致。

[0201] 表2展示了所有变体模型的性能实验结果。

[0202] 表2

[0203]	Datasets	Diginetica		Yoochoose1/64		Yoochoose1/4		Retailrocket	
	Methods	P@20	M@20	P@20	M@20	P@20	M@20	P@20	M@20
	DGS-MGNN w/o DGNA	53.95	19.42	71.88	32.15	72.46	32.98	63.92	45.21
	DGS-MGNN-MLP	53.51	18.33	71.12	31.10	71.72	32.01	62.85	44.82
	DGS-MGNN-GGNN	53.68	17.92	68.85	30.52	69.45	31.23	62.52	45.11
	DGS-MGNN-GAT	54.22	18.74	71.35	31.36	72.13	32.33	63.75	44.86
	<b>DGS-MGNN</b>	<b>55.54</b>	<b>19.68</b>	<b>72.62</b>	<b>32.49</b>	<b>73.10</b>	<b>33.55</b>	<b>66.61</b>	<b>45.78</b>

[0204] 从表2中可以看出,从DGS-MGNN中移除DGNA会使DGS-MGNN的性能出现明显的下降,这验证了DGNA的有效性。

[0205] 从表2中可以看到配备有MC-GNN的DGS-MGNN的性能始终优于其他变体模型,例如将MC-GNN替换为MLP、GGNN与GAT,在所有三种变体中,我们可以观察到DGS-MGNN-GGNN与DGS-MGNN-MLP相比几乎没有改善,原因是DGS-MGNN-GGNN没有区分项目之间的连接强度。此外,它还受到用户意外或错误点击的影响,这可能会给用户表示带来噪声信息。DGS-MGNN-GAT在所有比较的变体中实现了最佳性能,因为它通过引入注意机制生成的权重作为项目之间的连接强度来缓解连接强度问题。然而,与DGS-MGNN-GGNN类似,DGS-MGNN-GAT仍然会遇到由用户意外或错误点击引起的噪声信息的影响。与DGS-MGNN-GAT不同,我们提出的带有MC-GNN的DGS-MGNN模型通过动态调整图的结构和项目之间的连接强度,可以有效地处理连接强度和噪声信息问题。

[0206] 七、图增强的位置编码器的影响

[0207] 本实验对图增强的位置编码器(GPE)的有效性进行了验证。位置编码已经广泛的应用在了会话推荐任务中,然而应用在SASRec和Star-GNN中的绝对位置嵌入在会话推荐任务中并没有达到理想的效果。为了验证图增强的位置编码器的有效性,我们设置了2个相关的变体模型:

[0208] DGS-MGNN w/o GPE:将DGS-MGNN中的图增强的位置编码器(GPE)移除。

[0209] DGS-MGNN-POS:将DGS-MGNN中的图增强的位置编码器(GPE)替换为应用在SASRec和Star-GNN中的绝对位置嵌入。

[0210] 图3显示了各个对比模型的性能。实验结果揭示了配备了图增强的位置编码器(GPE)的DGS-MGNN的性能明显优于2个变体模型。更准确地说,带有图增强的位置编码器的模型(即本发明提出的DGS-MGNN和变体DGS-MGNN-POS)始终优于不考虑位置嵌入的变体模型(即DGS-MGNN w/o GPE)。此外,本发明提出的DGS-MGNN比采用传统图增强的位置编码器的DGS-MGNN-POS变体性能更好。图增强的位置编码器将会话长度信息引入到位置嵌入中以缓解上述问题。同时,图增强的位置编码器还能捕获会话固有的拓扑位置结构。

[0211] 八、图增强的注意力网络的影响

[0212] 为了验证探索图增强的注意力网络(GEA)对推荐性能的影响,我们设计了以下三种变体模型:

[0213] DGS-MGNN-S:用加合池化代替DGS-MGNN中的图增强的注意力网络(GEA)。

[0214] DGS-MGNN-M:用平均池化代替DGS-MGNN中的图增强的注意力网络(GEA)。

[0215] DGS-MGNN-SA:用普通注意力机制代替DGS-MGNN中的图增强的注意力网络(GEA)。

[0216] 表3

	Datasets	Diginetica		Yoochoose1/64		Yoochoose1/4		Retailrocket	
	Methods	P@20	M@20	P@20	M@20	P@20	M@20	P@20	M@20
[0217]	DGS-MGNN-S	53.07	18.91	70.02	30.81	70.87	31.68	63.82	44.48
	DGS-MGNN-M	53.68	19.42	71.95	32.21	72.55	33.12	62.79	45.12
	DGS-MGNN-SA	54.35	19.44	72.31	32.38	72.81	33.34	66.45	45.64
	<b>DGS-MGNN</b>	<b>55.54</b>	<b>19.68</b>	<b>72.62</b>	<b>32.49</b>	<b>73.10</b>	<b>33.55</b>	<b>66.61</b>	<b>45.78</b>

[0218] 如表3所示,基于加合池化的变体模型DGS-MGNN-S的性能最差。相反,基于平均池化的变体模型DGS-MGNN-M的性能优于DGS-MGNN-S。在所有三种变体中,基于自注意网络的变体模型DGS-MGNN-SA的性能最好,这表明了引入注意机制根据项目的重要性系数动态聚合信息的有效性。

[0219] 与所有三种变体相比,我们提出的带有GEA的DGS-MGNN模型在所有数据集上都表现出最佳性能。这主要是因为,在会话中,用户可能会有一些意外或错误的点击,传统的注意机制会为这些意外或错误的点击分配注意权重,从而给会话的表示带来噪音。相反,图增强的注意力网络(GEA)应用会话图的结构来识别用户在会话中的意外或错误点击,然后通过设置噪音项目的注意权重为0过滤掉噪声项目。

[0220] 九、顺序信息和结构信息的影响

[0221] 为了探索会话结构表示 $s_{long}^{str}$ 和会话顺序表示 $s_{long}^{seq}$ 对DGS-MCGNN性能的影响,我们将我们的模型与下列变体模型相比较:

[0222] DGS-MCGNN-STR: DGS-MGNN的一种变体,它只保留当前会话的长期结构表示,而忽略了DGS-MGNN中的长期序列信息。

[0223] DGS-MCGNN-SEQ: DGS-MGNN的一种变体,它只保留当前会话的长期顺序表示,而丢弃DGS-MGNN中的长期结构信息。

[0224] 图4展示了实验结果。可以看出,我们提出的DGS-MGNN在所有数据集上都获得了最佳性能,该模型既捕获了当前会话的长期结构表示,也捕获了当前会话的长期顺序表示。当DGS-MGNN仅配备当前会话的结构信息或序列信息时,其性能将显著降低。此外,在这两种变体中,DGS-MGNN-STR优于DGS-MGNN-SEQ,这表明:1) 当前会话的长期结构表示比当前会话的长期顺序表示更重要;2) 这两种长期表示是相辅相成的,它们的组合将极大提高性能。

[0225] 十、全局邻居数量对模型性能的影响

[0226] 为了探索全局邻居数量(K)对模型性能的影响,我们分析了不同全局邻居数量(0-50)下模型的性能变化。实验结果如图5所示,可以看出,在Diginetica数据上,DGS-MCGNN的性能随着全局邻居数量的增多而提升,当全局邻居数量超过10的时候,模型性能逐渐趋于稳定或轻微下降。在Yoochoose1/64数据上,当全局邻居数据小于20时,模型性能随着全局邻居数量的增多而提升,当超过20时,模型性能逐渐趋于稳定,相同的实验结论也在其他两个数据上得到了体现。结果表明,我们提出的方法DGS-MGNN可以在全局邻居数量相对较少的情况下获得令人满意的性能。

[0227] 十一、模型在不同长度的会话下的性能差异

[0228] 为了验证不同长度会话的模型性能,我们将会话分为两组(即长会话和短会话),其中长度大于5的会话被视为长会话,剩余的会话被视为短会话。我们将我们提出的方法DGS-MGNN与两个最具竞争力的基线模型,即DSAN和GCE-GNN分别在长会话和短会话上进行

了性能比较。从图6中可以看出,三种方法在短会话上的性能都优于它们在长会话上的性能,这可归因于长会话通常包含更复杂的用户兴趣(例如,兴趣转移),难以捕获,而短会话中的用户兴趣通常很简单。本发明提出的模型DGS-MGNN在所有数据集的两个指标上,同时在短会话和长会话上都始终优于两种最先进的基线方法。

[0229] 此外,与DSAN和GCE-GNN相比,DGS-MGNN在长会话上的性能改进明显大于在短会话上的性能改进。例如,在Diginetica数据集上,DGS-MGNN的P@20 (MRR@20) 在短会话上相对于DSAN和GCE-GNN的性能改进分别为1.58% (2.96%) 和1.78% (2.60%)。而长会话的相应改善率分别为5.23% (8.22%) 和5.61% (7.45%)。在其他3个数据集研究中也观察到类似的结果。

[0230] 十二、模型计算复杂度

[0231] 将DGS-MGNN的计算复杂度与其他四种最具竞争力的基线方法(包括SR-GNN、TAGNN、DSAN和GCE-GNN)进行了比较。在表4中展示了这些方法的理论计算复杂性以及实际训练时间和内存成本。

[0232] 从表4中可以看到SR-GNN的计算复杂度是 $O(s(nd^2+n^3)+d^2)$ ,其中n是会话长度,d是项目embedding的维度。为了简单起见,我们使用s表示不同图形神经网络的层数,即SR-GNN和TAGNN中的门控图神经网络(GGNN),GCE-GNN中的图注意力网络(GAT),以及我们提出的DGS-MGNN中的多通道图神经网络(MC-GNN)。与SR-GNN一样,TAGNN同样应用GGNN来学习节点的表示,此外它还开发了一种局部目标主义理论模块来测量当前会话中的每个项目与所有候选项目的注意力得分。因此,该方法具有更高的时间计算复杂度 $O(s(nd^2+n^3)+n|V|d^2+d^2)$ ,其中|V|是所有候选项目的数量。DSAN和GCE-GNN它们各自的计算复杂度分别是 $O(n^2d+d^2)$ 和 $O(sn^2d+nKd+d^2)$ ,我们所提出的模型DGS-MGNN的计算复杂度为 $O(sn^2d+n|V|d+d^2)$ ,其中主要的计算开销来自于MC-GNN模块中计算全局邻居的部分。从结果可以看出,DGS-MGNN的计算复杂度远低于TAGNN,略高于SR-GNN、DSAN和GCE-GNN。

[0233] 表4

Method	Datasets	Diginetica		Yoochose1/64		Yoochose1/4		Retailrocket	
	Complexity	Time	Memory	Time	Memory	Time	Memory	Time	Memory
SR-GNN	$O(s(nd^2+n^3)+d^2)$	80s	1359M	72s	1311M	1145s	1311M	69s	1435M
TAGNN	$O(s(nd^2+n^3)+n V d^2+d^2)$	827s	9745M	331s	8757M	5275s	8757M	250s	10621M
DSAN	$O(n^2d+d^2)$	128s	1523M	64s	1589M	878s	1589M	18s	1437M
GCE-GNN	$O(sn^2d+nKd+d^2)$	152s	2066M	76s	1699M	906s	1699M	20s	2208M
DGS-MGNN	$O(sn^2d+n V d+d^2)$	236s	3719M	126s	3371M	2058s	3371M	45s	2479M

[0235] 对于训练时间的消耗,我们可以观察到基线DSAN在所有数据集上比其他方法花费更少的训练时间,但Diginetica数据集上除外。基线TAGNN在所有数据集上具有最高的训练时间消耗。对于我们提出的方法DGS-MGNN,其训练时间成本远低于TAGNN,略高于SR-GNN、DSAN和GCE-GNN,这与理论计算复杂性的结果一致。就模型内存成本而言,基线SR-GNN和TAGNN分别显示了所有数据集的最低和最高内存成本。而其他两个基线模型DSAN和GCE-GNN显示了与SR-GNN相当或略高的内存成本。DGS-MGNN的存储成本远低于TAGNN,切与GCE-GNN具有的内存开销相比具有竞争性。基于上述分析,DGS-MGNN的计算复杂度适中,因此在潜在的应用中是可行的。

[0236] 需要说明的是,以上实施例仅用以说明本发明的技术方案而非限制,尽管通过参照本发明的优选实施例已经对本发明进行了描述,但本领域的普通技术人员应当理解,可

以在形式上和细节上对其作出各种各样的改变,而不偏离所附权利要求书所限定的本发明的精神和范围。同时,实施例中公知的具体结构及特性等常识在此未作过多描述。最后,本发明要求的保护范围应当以其权利要求的内容为准,说明书中的具体实施方式等记载可以用于解释权利要求的内容。



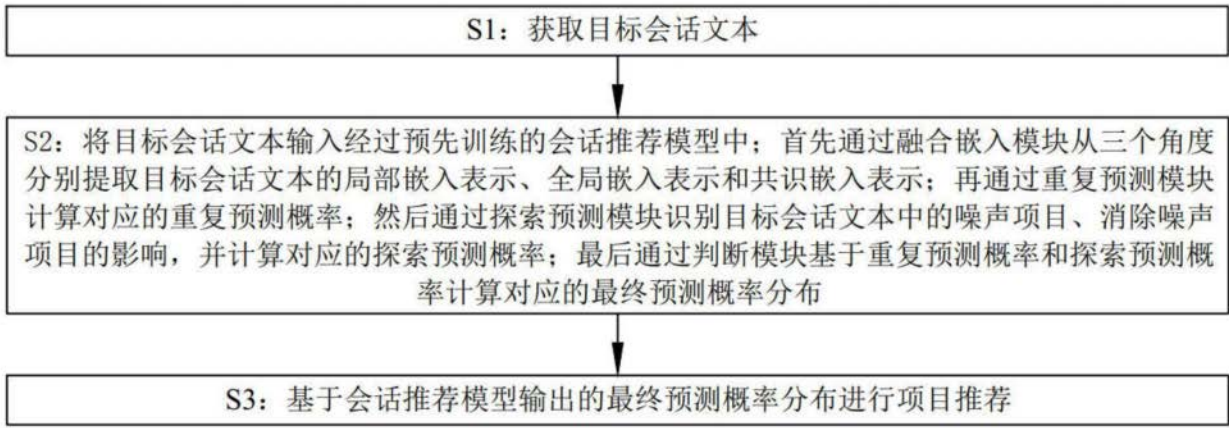


图1

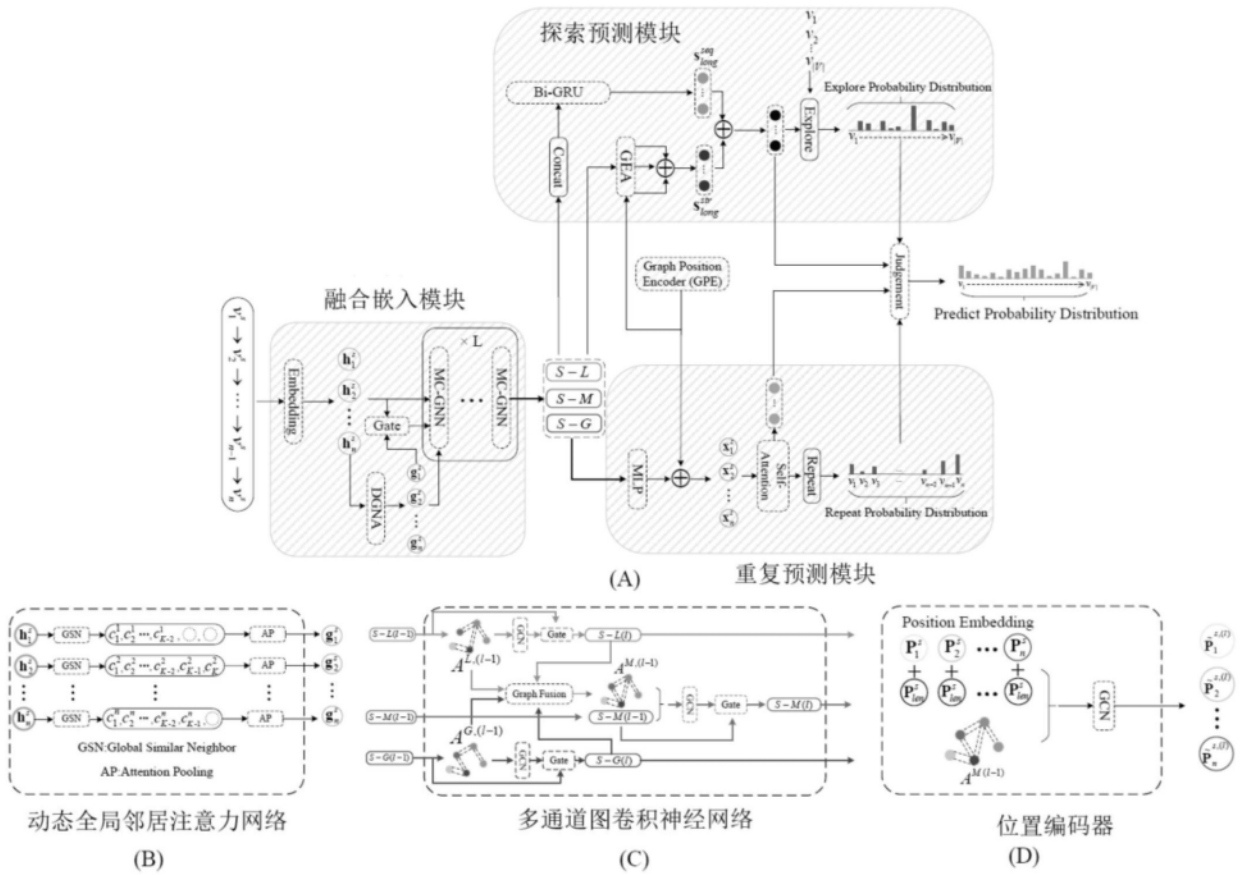


图2

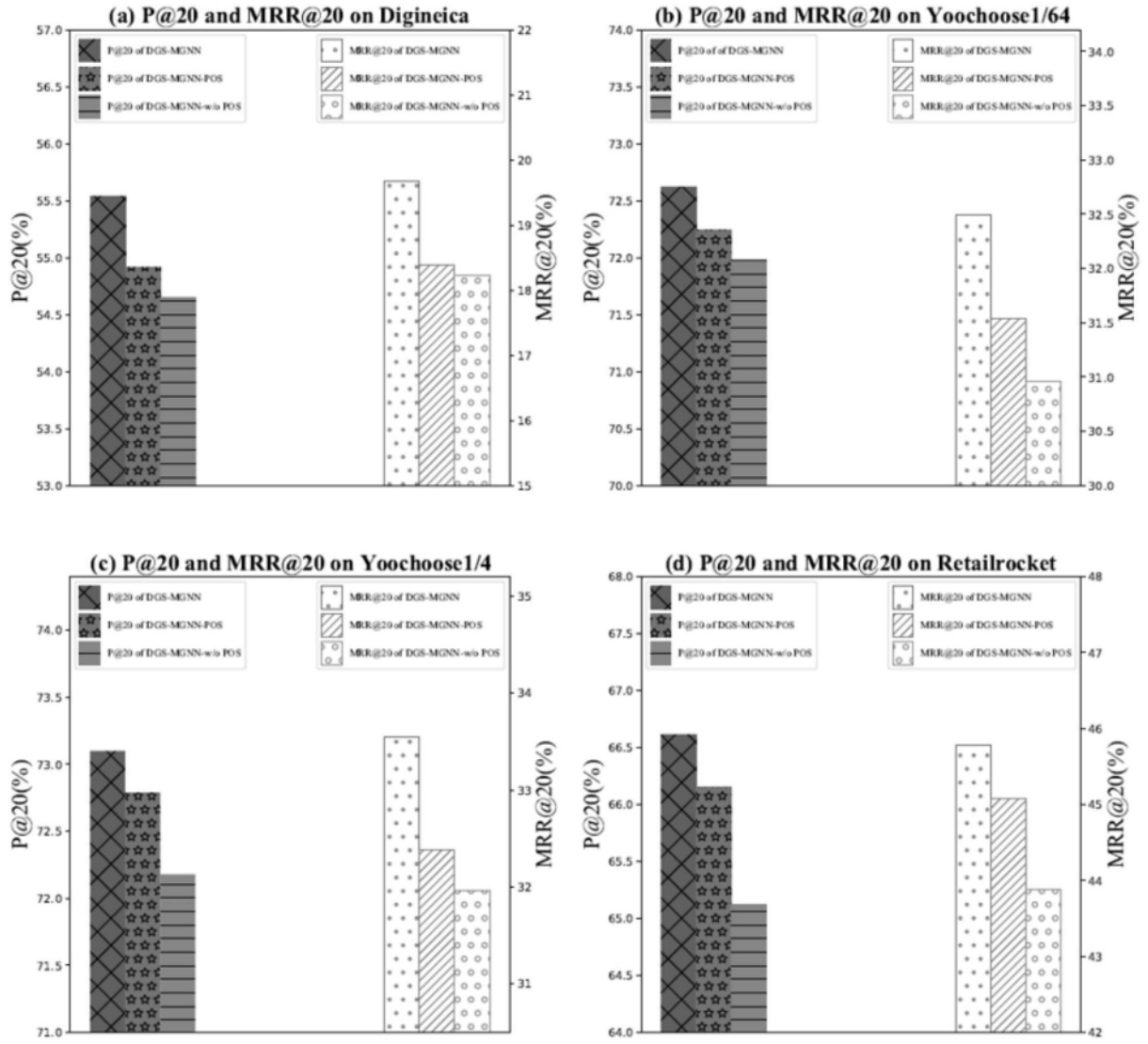


图3

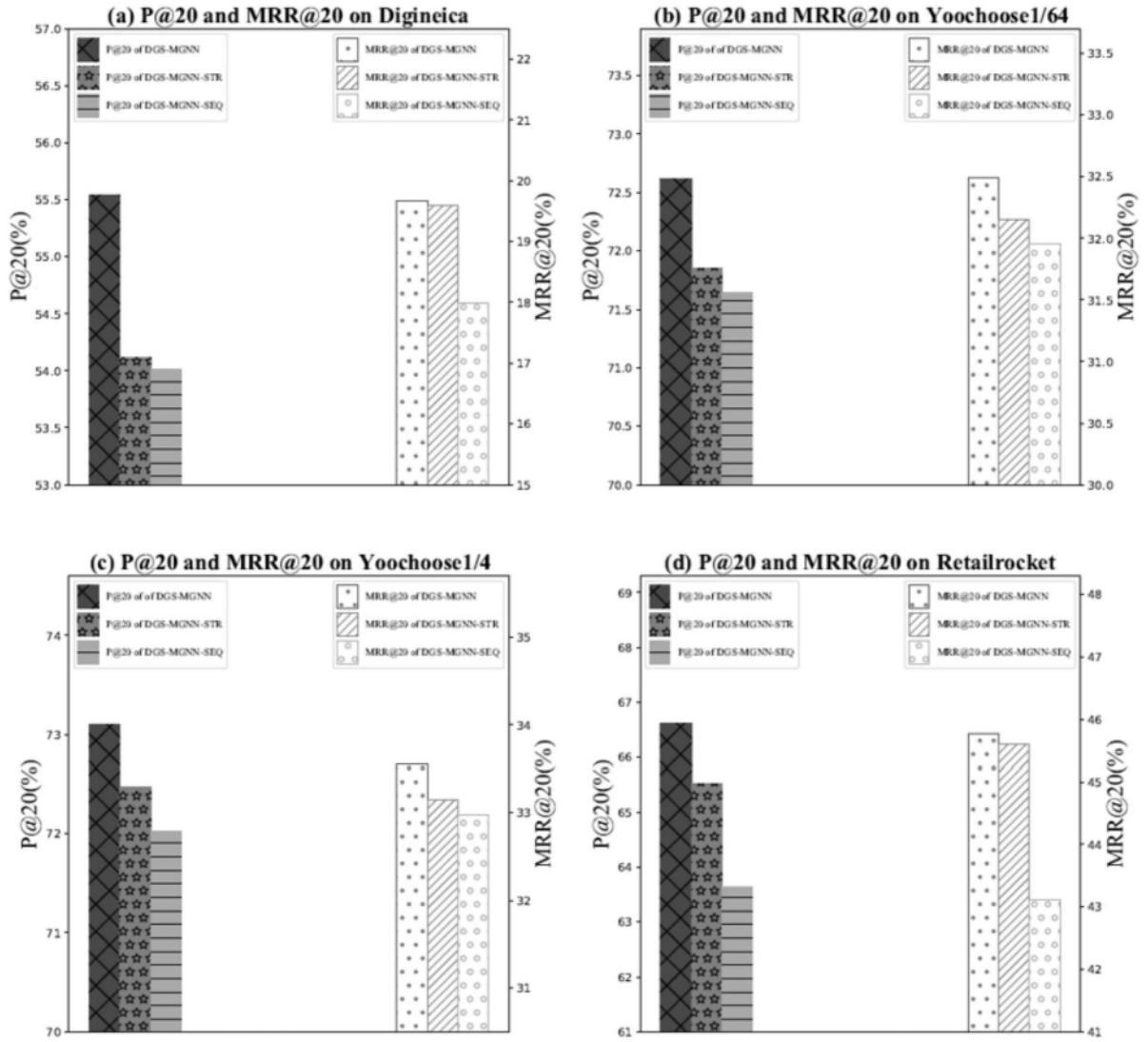


图4



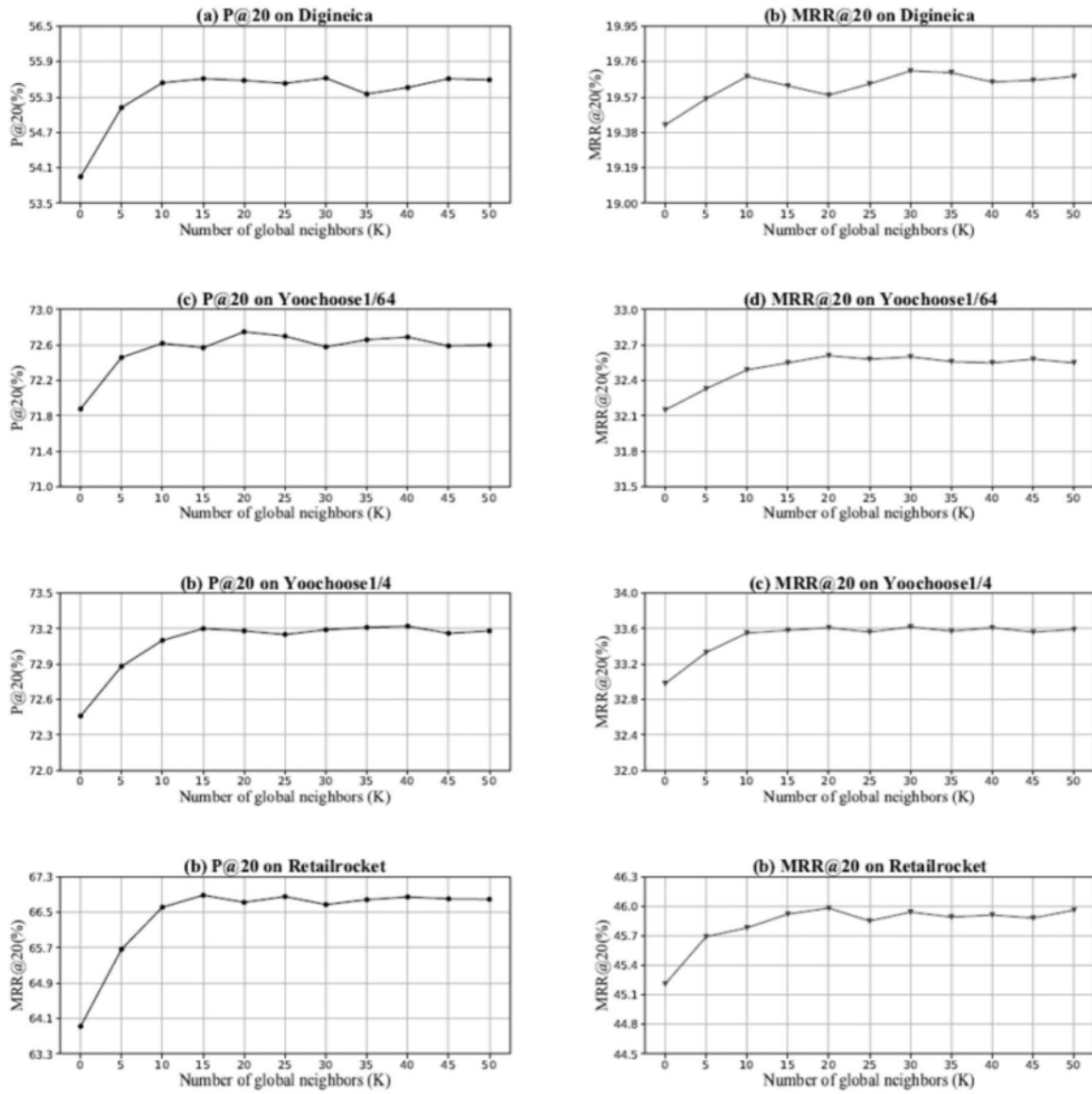


图5

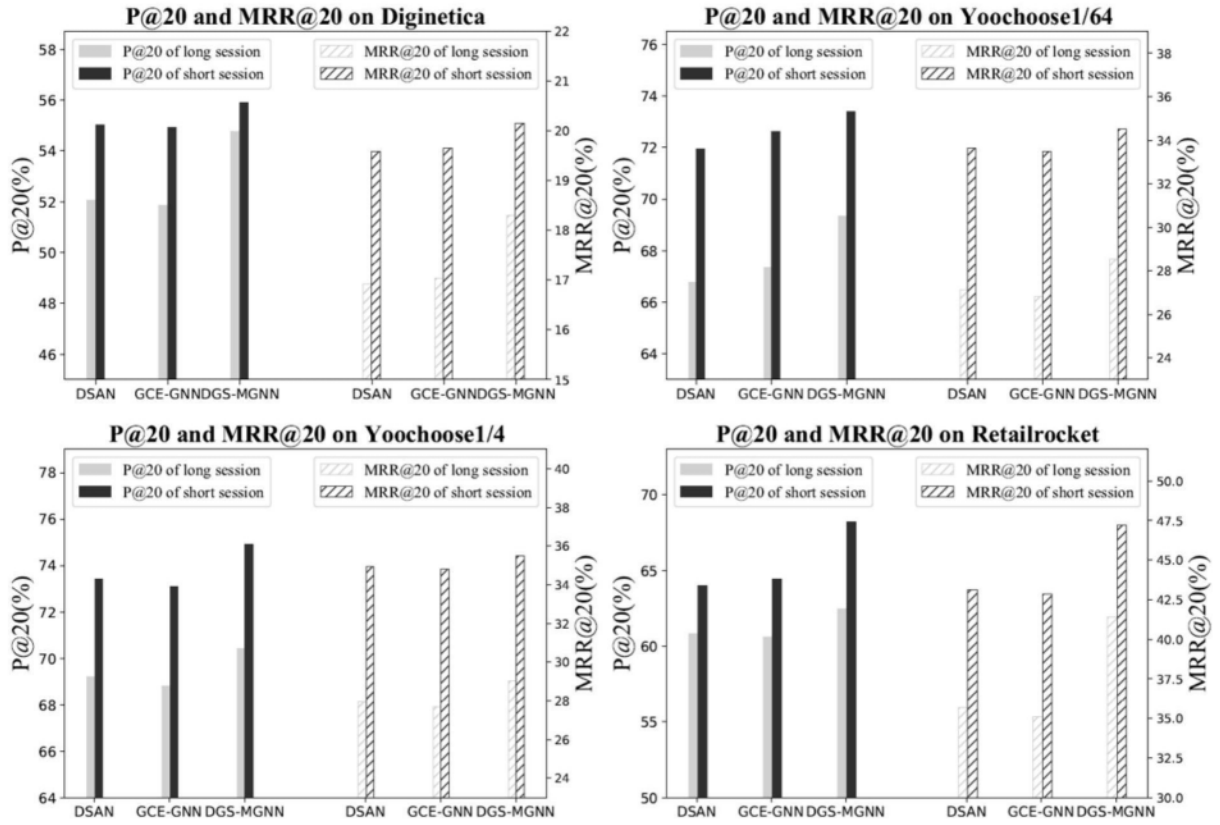


图6