

## 融合协同过滤的自组织神经网络多样化产品推荐

张秉楠<sup>1</sup>,李德玉<sup>1,2\*</sup>

(1.山西大学 计算机与信息技术学院,山西 太原 030006;  
2.山西大学 计算智能与中文信息处理教育部重点实验室,山西 太原 030006)

**摘要:**针对个性化推荐算法推荐结果容易存在冗余的问题,提出了一种融合协同过滤的自组织神经网络的多样化产品推荐方法。首先通过用户对产品的评分构建用户-产品、用户产品类别评分表,进而采用协同过滤算法得到基于评分相似用户的产品推荐列表;其次,将用户向量输入到自组织神经网络中聚类相似用户,利用相似用户查找目标用户可能感兴趣的产品类别,形成多样化推荐列表;最后,融合两个推荐列表形成满足多样化和准确性的产品推荐结果。通过亚马逊数据集上的实验,验证了所提方法在多样化推荐指标类别覆盖率(Category Coverage,CC)和项目层面的多样性(Item-Level Diversity,ILD)能够取得较好的结果,能有效地进行多样化推荐。

**关键词:**协同过滤;推荐系统;自组织映射神经网络;多样化

中图分类号:TP301 文献标志码:A 文章编号:0253-2395(2024)05-0954-10

## Diversified Product Recommendation by Integrating Collaborative Filtering with Self-organizing Neural Network

ZHANG Bingnan<sup>1</sup>, LI Deyu<sup>1,2\*</sup>

(1. School of Computer and Information Technology, Shanxi University, Taiyuan 030006, China;  
2. Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, Shanxi University, Taiyuan 030006, China)

**Abstract:** Aiming at the redundancy problem of the recommendation results for personalized recommendation algorithms, a diversified product recommendation method, integrating collaborative filtering with self-organizing neural network, is proposed. Firstly, the user-product rating table and user-product category rating table are constructed through the user's rating of the product. Furtherly, we adopt the collaborative filtering algorithm to get the product recommendation list based on similar users with similar ratings. Secondly, the user vectors are input into the self-organizing neural network to cluster similar users, and the similar users are used to help select the product categories that the target user may be interested in. A diversified recommendation list is formed. Finally, the two recommendation lists are fused to construct the diversified and accurate product recommendation results. The experiments on the Amazon datasets verified that the proposed method reaches better results on category coverage (CC) and item-level diversity (ILD) indexes and can effectively carry out diversified recommendations.

**Key words:** collaborative filtering; recommendation system; self-organizing neural network; diversified

收稿日期:2023-03-07;接受日期:2023-04-23

基金项目:国家自然科学基金(62072294)

作者简介:张秉楠(1995-),男,山西晋中人,研究生,研究方向为推荐系统。E-mail:474825505@qq.com

\* 通信作者:李德玉(LI Deyu),E-mail:lidysxu@163.com

引文格式:张秉楠,李德玉.融合协同过滤的自组织神经网络多样化产品推荐[J].山西大学学报(自然科学版),2024,47(5):954-963. DOI:10.13451/j.sxu.ns.2023068

## 0 引言

互联网个性化推荐主要根据用户的历史行为为用户推荐感兴趣的产品,能够提高用户的满意度。个性化推荐在推荐过程中,使用用户的显式、隐式偏好生成对用户的推荐<sup>[1]</sup>,推荐的结果可能冗余,一定程度降低了用户的服务体验质量。多样化产品推荐能够为用户提供不同类型的推荐结果,可以缓解个性化推荐的不足。多样化推荐可以满足用户不同的兴趣偏好,表现出项目类别属性的不同,提高推荐结果的多样化质量<sup>[2-5]</sup>。

在目前的推荐系统算法中,矩阵分解是最经典的推荐算法,大量多样化推荐算法的改进都以此为基础,Li等<sup>[6]</sup>使用核密度估计找到有共同兴趣的用户,使用兴趣用户扩充评分矩阵,达到增加多样化的目的。Gogna等<sup>[7]</sup>在低秩MF算法中引入电影的类别数据,添加正则项作为约束从而增加推荐多样性。有研究者将用户-商品图与多样化推荐结合,其中,用户与商品点之间使用连线表示存在关系,进而完成资源分配。An等<sup>[8]</sup>使用用户的活跃度得到虚拟专家,专家会分配到更多资源,将专家的资源换成更大范围的商品用以多样化推荐。Yu等<sup>[9]</sup>定义了信任度,使用新颖需求及多样化需求量化用户需求,控制信任度转移资源从而进行多样化推荐。随着深度学习技术的发展,深度网络融入了多样化推荐的研究,比如基于卷积神经网络(Convolutional Neural Networks, CNN)、循环神经网络(Recurrent Neural Networks, RNN)、生成对抗网络(Generative Adversarial Networks, GAN)等<sup>[10-11]</sup>。Fu等<sup>[12]</sup>使用长短期记忆网路(Long Short-Term Memory, LSTM)表示用户正反馈和负反馈改进网络,确保稳定推荐。Qin等<sup>[13]</sup>等对长尾物品使用CNN模型,从而实现多样化推荐。Gan等<sup>[14]</sup>、Liu等<sup>[15]</sup>和Wu等<sup>[16]</sup>采用行列式点过程(Determinantal Point Process, DPP)来改善不同推荐任务的多样性。Wu等使用DPP概率获得商品的体现,之后利用GAN由物品得出用户的多样化推荐。Liang等<sup>[17]</sup>受到视觉领域双边分支网络的启发。两个独立分支的体系结构自然使模型具备权衡准确性和多样性的能力,而无须广泛

调整权衡参数。上述工作对于多样化的研究还存在以下问题:(1)用户兴趣构建过程中,缺乏对用户兴趣类别多层次的研究(2)推荐结果的项目或多或少会有相似的情况,多样化程度较低,难以让用户满意。

本文在传统协同过滤推荐的基础上,结合自组织映射神经网络的聚类功能,将相似评分用户进行聚类;基于相似用户的多样化兴趣,提出融合协同过滤的自组织映射神经网络的多样化产品推荐方法,实现了多样化推荐,结果显示本文模型在推荐质量和推荐多样化上表现得更好,在Music数据集下多样化推荐指标类别覆盖率(Category Coverage, CC)和项目层面的多样性(Item-Level Diversity, ILD)上分别达到了0.73和0.79,在Beauty数据集下多样化推荐指标CC和ILD上分别达到了0.61和0.82。本文的主要贡献有:

(1)基于用户产品评分矩阵构建了用户产品类别评分矩阵,采用协同过滤算法实现了基于评分相似用户的产品排序推荐;进而融合用户产品类别评分矩阵到自组织映射神经网络,对用户进行聚类,发现相似用户。

(2)基于相似用户的多样化兴趣为目标用户进行推荐,克服传统个性化推荐的缺陷,提高多样化推荐的质量。与几种常见推荐方法相比较,本文提出的多样化推荐模型在Music数据集下多样化评价指标CC和ILD上分别提升了1.6%和1.9%,在Beauty数据集下多样化评价指标CC和ILD上分别提升了1.3%和1.4%,能够更好地为用户提供多样化产品。

## 1 融合协同过滤的自组织神经网络的多样化产品推荐方法

### 1.1 模型框架

本文将协同过滤的思想和自组织映射网络(Self-Organizing Maps, SOM)进行结合实现多样化推荐。图1给出了融合协同过滤的自组织神经网络多样化产品推荐框架。框架包括三部分,评分表的构建,多样化产品推荐列表构建,以及评分预测建模。

在评分表构建阶段,依据用户-产品评分信息和产品类别信息,构建用户-产品类别评分

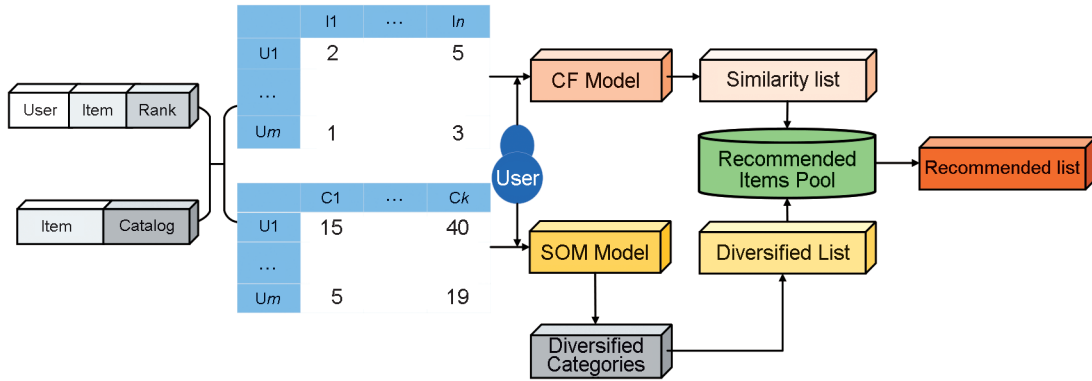


图1 融合协同过滤的自组织神经网络的多样化产品推荐框架

Fig. 1 Diversified recommendation framework, integrating collaborative filtering with self-organizing neural networks

表;进而计算所有用户对于每一个产品的平均评分,结合产品所属的类别,计算基于产品类别的产品评分表;最后根据每个类别的产品评分表计算得到每个产品类别的总评分,表示产品类别的热度。

在多样化产品推荐列表构建阶段,根据用户-产品评分表利用协同过滤算法,实现基于评分相似用户的产品推荐列表;进而将用户-产品类别评分矩阵输入到SOM神经网络,对用户进行聚类,基于聚类的相似用户查找目标用户可能感兴趣的产品类别,依据产品类别评分表对产品类别进行排序,并且对每个类别中产品评分进行排序,得到多样化产品推荐集合。

在评分预测建模阶段,通过融合基于协同过滤的相似推荐列表结果,以及基于SOM网络的多样化推荐列表结果,得到要推荐的产品集合;进而对集合中的产品进行热度计算,基于产品的热度对产品进行排序,为用户提供多样化的产品推荐排序列表。

1.2 融合协同过滤的自组织神经网络多样化产品推荐建模

1.2.1 评分表构建

依据数据集中已有的数据,构建用户-产品评分表,如表1所示。根据产品的类别,获取用户对于产品类别的评分构建用户-产品类别评分表。如表2所示。

表1 用户-产品评分表

Table 1 Rating table of user-item

	产品1	产品2	...	产品n
用户1	5	4	...	1
用户2	3	5	...	4

表2 用户-产品类别评分表

Table 2 Category rating table of user-item

	产品类别1	产品类别2	...	产品类别n
用户1	10	5	...	3
用户2	12	17	...	20

类别评分是用户对产品所属类别中所有产品进行的评分总和。根据用户在某类别上的评价,计算根据用户对产品类别的评分 $r_{iq}$ ,如公式(1)所示:

$$r_{iQ} = \sum_{q \in Q} r_{iq}, \tag{1}$$

其中Q为产品类别。 $r_{iq}$ 为用户i对产品q的评分。进而,计算某一产品类别中每个产品对于所有用户的平均评分,结合产品所属类别Q,获取某类产品中产品 $q_h$ 的评分 $r_{q_h}$ ,如公式(2)所示:

$$r_{q_h} = \frac{\sum_{i \in u} r_{iq_h}}{|u|}, \tag{2}$$

其中 $|u|$ 表示全部用户的数量, $\sum_{i \in u} r_{iq_h}$ 是所有用户对于类别Q下 $q_h$ 产品的总评分。

最后依据用户对产品类别的评分获取所有用户对于产品类别Q的总评分 $r_Q$ ,以此表示产品类别的总热度,如公式(3)所示:

$$r_Q = \sum_{i \in u} r_{iQ}. \tag{3}$$

1.2.2 多样化产品推荐列表构建

(1) 基于评分相似用户的协同推荐列表

对目标用户依据表1构建的用户物品评分矩阵使用协同过滤思想,得到基于评分相似用户的推荐列表。根据用户对于产品的评分,利用相似用户对目标用户未评分的物品进行评

分,来预测目标用户对物品的评分,按照评分降序排列,得出 Top-N 推荐。

具体步骤如下:首先构建用户-产品评分矩阵,这部分在表1已完成;计算用户间相似度,使用修正余弦相似度进行计算,如公式(4)所示:

$$Sim(i, k) = \frac{\sum_{q \in I_{i,k}} (r_{i,q} - \bar{r}_i)(r_{k,q} - \bar{r}_k)}{\sqrt{\sum_{q \in I_i} (r_{i,q} - \bar{r}_i)^2} \sqrt{\sum_{q \in I_k} (r_{k,q} - \bar{r}_k)^2}}, \quad (4)$$

其中  $Sim(i, k)$  表示用户  $i$  和用户  $k$  间的相似度,  $I_i$ 、 $I_k$  和  $I_{i,k}$  分别表示被用户  $i$ 、用户  $k$  和被用户  $i$ 、 $k$  共同评分过的物品集合,  $r_{i,q}$  和  $r_{k,q}$  分别表示用户  $i$  和用户  $k$  对物品  $c$  的评分,  $\bar{r}_i$  和  $\bar{r}_k$  则分别表示用户  $i$  和用户  $k$  的评分均值;根据与目标用户的相似度为其选择邻居用户集;在邻居用户集中地相似用户对某件待推荐产品的评分和用户间相似度计算目标用户对待推荐产品的预测评分。则用户  $i$  对产品  $q$  的预测评分  $r_{i,q}$  计算公式如公式(5)所示:

$$r_{i,q} = \frac{\sum_{v \in T} (S_{i,k} \cdot r_{k,q})}{\sum_{v \in T} |S_{i,k}|}, \quad (5)$$

其中  $S_{i,k}$  表示用户  $i$  和  $k$  之间的相似度,  $r_{k,q}$  表示邻居用户  $k$  对待推荐物品  $q$  的评分,  $T$  表示邻居用户集合。基于以上步骤,我们按照评分高低可得到一个基于评分相似用户的 Top-N 推荐列表。

(2) 基于自组织神经网络的多样化推荐列表

用户-产品类别矩阵中每一行向量表示当前用户的一条对于产品类别的兴趣度评分,使用自组织映射神经网络(SOM)对用户的兴趣进行聚类,基于聚类的相似用户查找目标用户可能感兴趣的产品类别,依据产品类别评分表对产品类别进行排序,并且对每个类别中产品评分进行排序,得到多样化产品推荐集合。SOM神经网络是一种基于竞争的无监督神经网络,可以将高维输入数据映射到低维空间,同时又保证网络的拓扑结构,而最终将输入映射到某一个类型的分类中去。其网络结构如图2所示。

SOM网络中竞争层采用竞争学习的方式来获取用户兴趣的分布,每个用户向量对应一个聚类类别。也即,每个用户向量的输入都只会一个神经元获胜,下面介绍SOM的聚类过程:

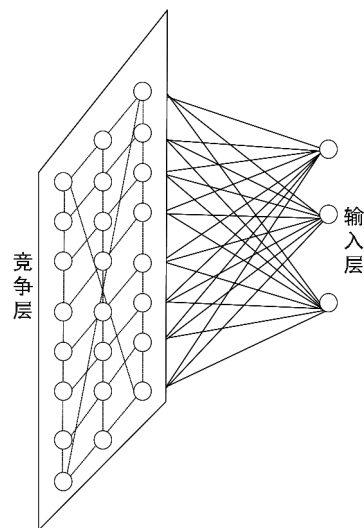


图2 SOM网络结构图

Fig. 2 Structure diagram of SOM network

竞争过程:

将用户向量表示为  $m$  维向量  $O = [o_1, o_2, \dots, o_m]^T$ , 假设输入层神经元个数为  $m$ , 即产品总类别是  $m$  个, 竞争层神经元个数为  $N$ , 即用户分为  $N$  类, 输入层与竞争层连接的权值表示为  $W = [w_1, w_2, \dots, w_N]$ , 其中  $w_i^T = [w_{i1}, w_{i2}, \dots, w_{im}]$ ,  $i \in (1, 2, \dots, N)$ 。

首先,随机初始化竞争层权值参数  $W$ , 计算输入向量  $O$  与竞争层连接权值的内积向量  $n$ , 如公式(6)所示:

$$n = [n_1, n_2, \dots, n_N] = WO = [w_1^T o, w_2^T o, \dots, w_N^T o]^T, \quad (6)$$

其中  $w_i^T o$  为输入向量  $O$  与神经元  $i$  之间连接权值的内积。(计算输入向量  $O$  与  $N$  个神经元之间的距离,选距离最大的为获胜神经元)。进一步,选取竞争过程中的获胜神经元,通过最大值来选择用户的粗粒度类别兴趣。一个获胜神经元所聚类得到的用户,具有相似的粗粒度类别兴趣。将内积向量  $n$  使用获胜公式计算,如公式(7)所示:

$$a = \text{compet}(n) = \begin{cases} 1, & i = i^* \\ 0, & i \neq i^* \end{cases}, \quad (7)$$

其中当  $n_{i^*} \geq n_i, \forall i, i^*$  为获胜神经元标号,并且  $i^* \leq i, \forall n_i = n_{i^*}$ 。获胜的原则是与输入向量内积越大说明神经元与向量越靠近,该神经元能表明输入向量的特征,最后只有一个神经元获胜,类似结果为  $a = [0, 0, \dots, 1, 0, \dots, 0]$ 。

学习过程:

SOM神经网络采用的权值调整为Kohonen学习规则。主要思想为让获胜的神经元以及邻近的神经元向当前的输入样本靠近,修改神经元权值向输入样本的方向移动,以达到神经元表示输入的目的,具体过程如下:

首先,以 $i^*$ 为中心确定迭代次数为 $z$ 时获胜神经元周围可做权值调整的神经元领域为 $D(z)$ ,在训练开始时领域最大,尽可能让多一些神经元进行学习,随训练次数增多,领域随之减小。如公式(8)所示:

$$D(z) = D(0) \cdot \left(1 - \frac{z}{T}\right), \quad (8)$$

其中 $T$ 为总迭代次数, $D(0)$ 表示初始领域。进而,计算迭代次数为 $z$ 时 $D(z)$ 神经元领域内权值向量 $W_i(z)$ 调整公式如公式(9)所示:

$$W_i(z) = W_i(z-1) + \alpha(z)(O_j(z) - W_i(z-1)), \quad (9)$$

其中 $W_i(z)$ 、 $W_i(z-1)$ 分别表示神经元 $i$ 在迭代次数 $z$ 和 $z-1$ 的权值向量, $O_j(z)$ 表示在时刻 $z$ 的输入向量, $\alpha(z)$ 表示 $z$ 时刻的学习率,如公式(10)所示:

$$\alpha(z) = \alpha(0) \cdot \left(1 - \frac{z}{T}\right). \quad (10)$$

其次,将所有用户向量输入到SOM神经网络中,经过以上步骤若干次后,竞争层的神经元可表示相似用户。根据相似用户感兴趣的产品类别集合,对目标用户感兴趣的产品类别求补集,求得目标用户可能感兴趣的类别集合,如公式(11)所示:

$$Q_{1,2,\dots,x} = Q_1 \cup Q_2 \cup \dots \cup Q_x, \quad (11)$$

其中 $Q_x$ 代表用户 $x$ 感兴趣的产品类别集。根据用户 $x$ 在 $Q_{1,2,\dots,x}$ 的产品类别,计算用户可能感兴趣的产品类别集合 $C_{Q_{1,2,\dots,x}} Q_x$ ,如公式(12)所示:

$$C_{Q_{1,2,\dots,x}} Q_x = Q_{1,2,\dots,x} - Q_x \quad (12)$$

最后,若要对用户 $x$ 推荐产品,将 $C_{Q_{1,2,\dots,x}} Q_x$ 集合中产品类别与产品类别总热度 $r_Q$ 进行排序,进而对产品类别中产品的热度进行排序,获得多样化的产品推荐集合。推荐流程如图3所示。

### 1.2.3 评分预测建模

给定目标用户 $x$ ,对两个推荐列表中的推荐产品依据评分 $r_{x,q}$ 和 $r_{q_i}$ 做归一化处理,得到的数值即为该产品的推荐评分。根据两个推荐列表得到推荐产品集合,使用逻辑回归来学习不同推荐列表集合的融合权重,并依据产品推荐评分计算产品的综合评分,根据产品的综合评分排序生成最终的融合推荐列表。综合评分 $s_q^{merge}$ 的计算公式如下式(13)所示:

$$s_q^{merge} = \omega_0 + \omega_1 s_q^{similarity} + \omega_2 s_q^{diversity}, \quad (13)$$

其中 $\omega = \{\omega_0, \omega_1, \omega_2\}$ 为两种列表融合的权重,对于某个产品 $q$ 来说,在相似性列表的推荐分数为 $s_q^{similarity}$ ,在多样化列表的推荐分数为 $s_q^{diversity}$ 。

进而,计算用户对产品感兴趣的评分 $h_\omega$ 表示为:

$$h_\omega(q_v) = \frac{1}{1 + e^{-\frac{s_q^{merge}}{q_v}}} \quad (14)$$

对于 $e$ 个样本,最终的损失函数 $J(\omega)$ ,定义如下:

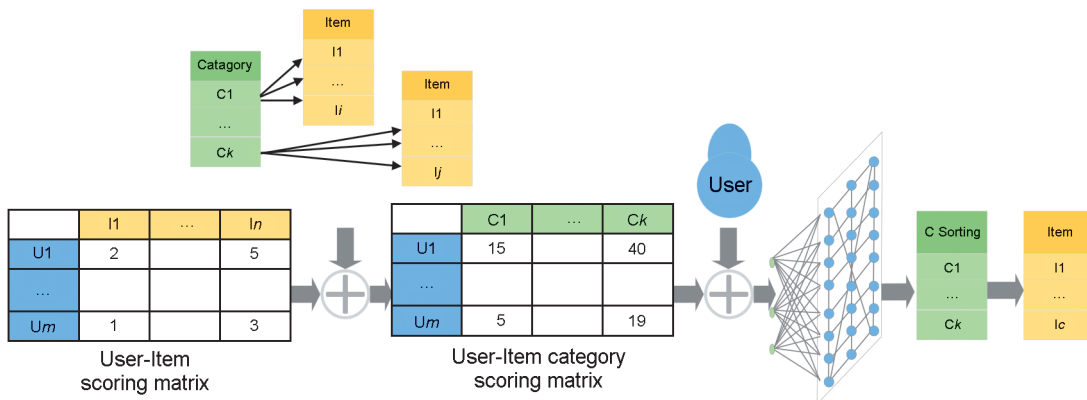


图3 SOM网络推荐产品品类的框架

Fig. 3 The framework of SOM network recommending commodity categories

$$J(\omega) = -\frac{1}{e} \sum_{i=1}^e [y^v \log h_\omega(q_v) + (1 - y^v) \log(1 - h_\omega(q_v))]. \quad (15)$$

实验过程中,我们采用梯度下降法求解得到列表融合的权重。在梯度下降法求解后,  $\omega = \{\omega_1, \omega_2\}$  以及  $\omega_0$ , 可以计算目标用户的推荐集合中不同产品的最终得分, 即  $s_q^{\text{merge}}$ , 根据该值选取 Top-N 的产品作为用户的最终推荐列表。

## 2 实验设置

### 2.1 实验数据集

为验证所提出模型的有效性,使用亚马逊 Music 和 Beauty 数据集进行一系列实验。数据集中,产品包含了丰富的属性描述信息,如产品类别与风格等。用户对产品的评分范围是 1~5,评分越高,说明用户对产品的兴趣越大。表 3 给出了实验数据集的统计数据。

表 3 数据集大小说明

Table 3 Size description of dataset

数据集	用户	产品	产品类别	所有行为数量
Music	5 541	3 568	309	64 786
Beauty	8 159	5 862	248	98 566

### 2.2 对比模型

为了验证本文模型的有效性,本文采用了常见的基于用户的协同过滤算法(User Collaboration Filter, UserCF)<sup>[18]</sup>、贝叶斯个性化排序(Bayesian Personalized Ranking, BPR)<sup>[19]</sup>、非负矩阵分解(Non-negative Matrix Factorization, NMF)<sup>[20]</sup>、深度语义匹配模型(Deep Structured Semantic Model, DSSM)<sup>[21]</sup>、最大边际相关性(Maximal Marginal Relevance, MMR)<sup>[22]</sup>、DPP<sup>[23]</sup>推荐算法在亚马逊数据集上进行了实验。对比模型介绍如下:

1) UserCF: 通过计算余弦距离找出目标用户相似的用户集,计算物品得分推荐给用户。

2) BPR: 对每个用户都重建一个偏序关系,其中有交互的产品优先于未交互的产品。

3) NMF: 结合了广义矩阵分解(Generalized Matrix Factorization, GMF)和多层感知机 MLP, 提取出复杂的用户-产品交互关系。

4) DSSM: 构建用户和产品两个独立网络,将双“塔”中的用户向量和产品向量缓存到内

存数据库中,预测时只需在内存中计算相似度运算。

5) MMR: 通过独立的度量表示相关性和多样性,并通过权衡参数来最大化边际相关性

6) DPP: 应用确定点过程来优化准确性和多样性之间的权衡,并通过平均准确率均值(Mean Average Precision, MAP)推断生成建议列表。

### 2.3 参数设置

实验过程中,本文将数据集的 80% 划分为训练集,剩余的 20% 划分为测试集。在每种对比模型中,我们将不同推荐模型的嵌入维度在 {16, 32, 64, 128, 256} 进行网格搜索,正则化参数在 {0.01, 0.02, ..., 0.05} 进行搜索,学习率为 {0.001, 0.003, ..., 0.009, 0.01}, 通过网格调优方法, BPR 模型嵌入维度 64, 正则项参数为 0.025, 学习率 0.01, DSSM 模型嵌入维度为 50, 学习率为 0.001, batch size 128; NMF 模型隐向量维度为 120, 学习率为 0.001, 本算法中 SOM 竞争层  $x=y=10$ , 领域更新参数 sigma 为 1, 学习率为 0.01。

### 2.4 评价指标

为了评估模型的准确性,本研究采用如下指标所示:

1) HR@K: 命中率,刻画了用户的需求项是否包含在模型的推荐项中。命中率越高,说明推荐准确性越高,定义如下:

$$HR@K = \frac{1}{|U|} \sum_{i=1}^{|U|} hits(i), \quad (16)$$

其中 K 为推荐列表长度, U 为全部用户, hits(i) 表示第 i 个用户交互的产品是否在推荐列表中,是为 1 否则为 0。

2) NDCG@K: 归一化折损累计增益(NDCG), 排名越靠后则折损值就越大。定义如下:

$$DCG@K = \sum_{i=1}^K \frac{2^{r_i-1}}{\log_2^{i+1}}, \quad (17)$$

$$NDCG@K = \frac{DCG@K}{IDCG@K}, \quad (18)$$

其中 K 为推荐列表长度,  $r_i$  为相关因子, 如果产品  $i \in y_u^{\text{test}}$  则  $r_i = 1$ , 否则  $r_i = 0$ 。

3) CC@K: 类别覆盖率,刻画了推荐列表中产品覆盖的类别数量。类别覆盖率越高,说明推荐

列表中产品的类别多样化程度越高。定义如下:

$$CC@K = \frac{C_k}{C_{\text{tol}}}, \quad (19)$$

其中  $C_k$  为 top-k 个产品所覆盖的类别数量,  $C_{\text{tol}}$  为类别总数。

4) ILD@K: 整体多样性, 衡量了所有用户推荐列表的多样性。定义如下:

$$ILD(L) = 1 - \frac{\sum_{b_i \in L} \sum_{b_j \in L, b_j \neq b_i} S(b_i, b_j)}{\sum_{b_i \in L} \sum_{b_j \in L, b_j \neq b_i} L}, \quad (20)$$

其中  $S(i, j)$  为两个产品的相似度,  $L$  为推荐列表。

### 3 实验结果分析与讨论

#### 3.1 推荐结果准确性分析

实验过程中, 我们将推荐列表长度  $K$  分别设置为 5 和 10, 通过在 7 个对比模型上进行实验, 结果如表 4 和表 5 所示。

表 4 在 Music 数据集中的准确性推荐指标 HR 和 NDCG 结果  
Table 4 HR and NDCG results for accuracy recommendation indicators in the Music dataset

Models	HR@5	HR@10	NDCG@5	NDCG@10
UserCF	0.101	0.144	0.149	0.171
BPR	0.113	0.131	0.153	0.179
NMF	0.119	0.179	0.179	0.203
DSSM	0.154	0.213	0.217	0.242
MMR	0.057	0.059	0.129	0.131
DPP	0.077	0.117	0.097	0.124
SOMDR	0.083	0.121	0.117	0.127

表 5 在 Beauty 数据集中的准确性推荐指标 HR 和 NDCG 结果  
Table 5 HR and NDCG results for accuracy recommendation indicators in the Beauty dataset

Models	HR@5	HR@10	NDCG@5	NDCG@10
UserCF	0.049	0.063	0.059	0.068
BPR	0.054	0.062	0.061	0.072
NMF	0.061	0.089	0.084	0.095
DSSM	0.068	0.098	0.087	0.099
MMR	0.043	0.055	0.052	0.064
DPP	0.040	0.057	0.049	0.059
SOMDR	0.048	0.059	0.054	0.061

在表中, 交互式增强算法 DSSM 在所有模型中表现最好, 特别是在 HR@5、HR@10、NDCG@5 和 NDCG@10 上表现最优。这一结果可能是由于 DSSM 分别使用相对独立的两个复

杂网络构建用户相关特征的 user embedding 和 item 相关特征的 item embedding, 线上预测的时候只需要在内存中进行相似度运算, 可以缓解数据的稀疏性问题。

此外, 本文所提自组织神经网络多样化产品推荐 (Self-Organizing Map for Diversified Recommendations, SOMDR) 算法与其他算法相比也有较强的表现, 这说明自组织映射神经网络对用户聚类的方式, 能够有效引入用户群组信息, 对于缓解数据稀疏性有一定的帮助。最后, 推荐多样性算法 MMR、DPP 整体表现不佳, 本文所提算法 SOMDR 相比传统推荐多样性算法 DPP, 在 Music 数据集上准确性指标 HR@10 和 NDCG@10 相对提升 3.4% 和 2.4%, 在 Beauty 数据集上准确性指标 HR@10 和 NDCG@10 相对提升 3.5% 和 3.3%, 相对指标提升说明 SOMDR 模型相比传统推荐多样性模型, 可以更好地权衡推荐精度。

#### 3.2 推荐结果多样性分析

为了评估模型的推荐多样性, 下图给出了 Top-5、Top-10 的推荐性能。图 4 展示了在数据集 Music 下 CC@5、CC@10、ILD@5、ILD@10 的结果, 图 5 展示了在数据集 Beauty 下 CC@5、CC@10、ILD@5、ILD@10 的结果。

从图 4 和图 5 中观察到, 本文所提出的 SOMDR 模型表现最好。这一结果可能是由于 SOMDR 首先引入自组织映射神经网络来针对用户-产品类别信息聚类, 然后通过二次排序, 来将多样化产品推荐给用户, 进一步调整推荐产品种类的多样性。

此外, 在数据集 Music 上推荐多样性算法 MMR 表现次佳, 这是因为该方法通过独立的度量表示相关性和多样性, 并通过权衡参数来最大化边际相关性, 从而提高推荐的个体多样化。SOMDR 算法相比次优 MMR 算法, 在 CC@10、ILD@10 上, 相对提升了 1.6%、1.9%。在数据集 Beauty 上推荐多样性算法 DPP 表现次佳, 这是因为 Beauty 数据集较 Music 数据集更加稀疏, DPP 算法确定点过程来优化准确性和多样性之间的权衡, 并通过 MAP 推断生成推荐列表可以在一定程度上解决数据稀疏的问题。SOMDR 算法相比次优 DPP 算法, 在 CC@10、

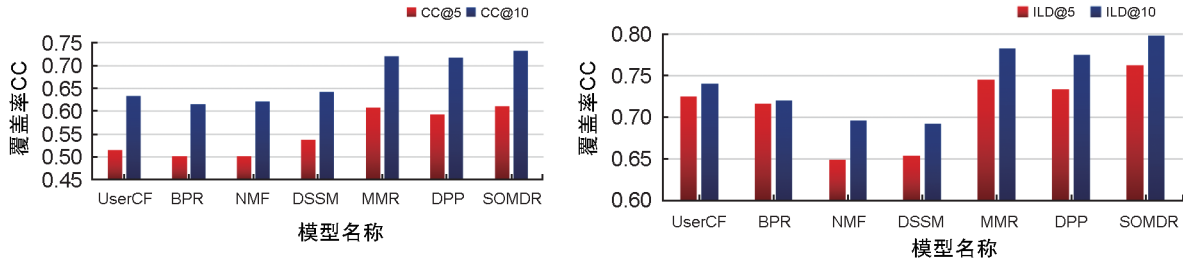


图4 在Music数据集中的多样性推荐指标CC和ILD结果

Fig. 4 CC and ILD results for diversity recommendation indicators in the Music dataset

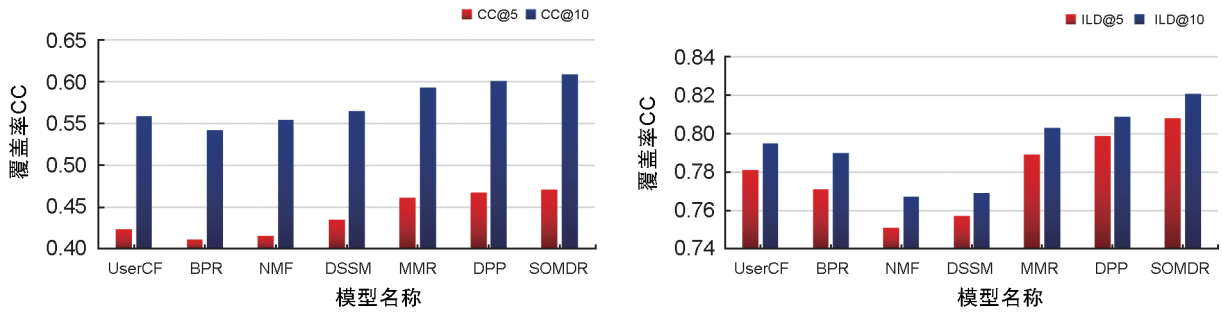


图5 在Beauty数据集中的多样性推荐指标CC和ILD结果

Fig. 5 CC and ILD results for diversity recommendation indicators in the Beauty dataset

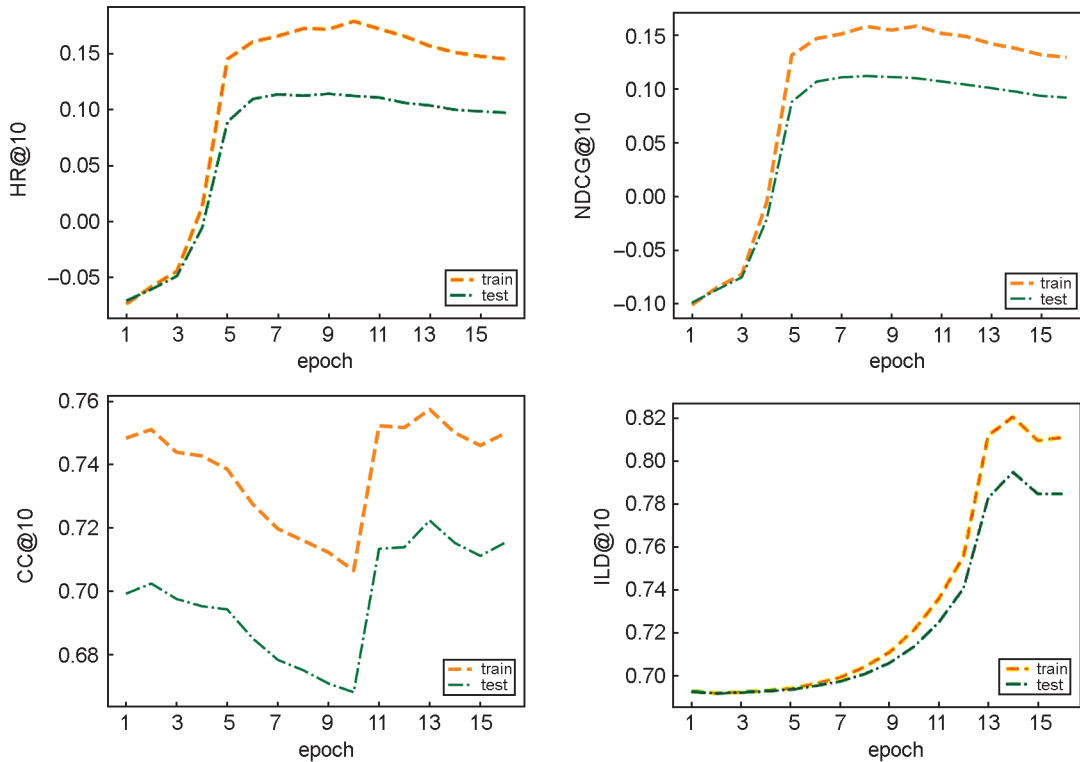


图6 在Music数据集中的epoch敏感性分析

Fig. 6 Epoch sensitivity analysis in the Music dataset

ILD@10上,相对提升了1.3%、1.4%。

综上所述,在所有对比算法中,本文所提模型SOMDR在保证推荐准确性的同时,能较好地实现推荐多样性。

### 3.3 敏感性实验分析

在这一部分中,分析了四个实验评估指标的训练代数epoch的敏感性,将SOMDR的敏感性结果展示于图6。

从图6中可以看出, epoch大小对实验结果有一定影响。当 epoch 为 10 时, SOMDR 模型 HR@10 指标结果最高, NDCG@10 指标也在最大值附近。之后模型各个精度变化不大, 模型收敛, 本研究取 epoch=13 下的结果为最终实验结果。此外, 随着 epoch 的上升, ILD@10 指标在 epoch 为 13 附近上升最快, 当 epoch>14 时, 结果变化不大; CC@10 类似, 在 epoch=12 附近收敛, 随后在 epoch=14 附近缓慢震荡变化。

#### 4 结论

本文考虑了用户对产品的多类别兴趣, 提出了一种融合协同过滤的自组织神经网络的多样化产品推荐模型。该方法根据用户对产品的多类别评分, 利用协同过滤算法实现基于评分相似用户的产品推荐; 并基于用户对产品的类别评分矩阵进行相似用户的聚类, 发现用户可能感兴趣的产品类别, 构建多样化产品列表; 进而将相似性推荐列表和多样化产品推荐列表进行融合, 为用户提供多样化的产品推荐服务。在亚马逊 Music 数据集下多样化推荐指标 CC 和 ILD 上分别达到了 0.73 和 0.79, 在亚马逊 Beauty 数据集下多样化推荐指标 CC 和 ILD 上分别达到了 0.61 和 0.82。实验结果表明本模型与其他传统模型相比, 在多样化推荐性能上表现更好。未来工作中, 进一步考虑到如何平衡推荐模型的推荐结果准确性与多样性, 实现推荐结果性能的全方位提升, 提高推荐结果的可解释性以及用户对推荐结果的满意度。

#### 参考文献:

- [1] CHO Y H, KIM J K. Application of Web Usage Mining and Product Taxonomy to Collaborative Recommendations in E-commerce[J]. *Expert Syst Appl*, 2004, **26**(2): 233-246. DOI: 10.1016/S0957-4174(03)00138-6.
- [2] SUN J N, GUO W, ZHANG D C, et al. A Framework for Recommending Accurate and Diverse Items Using Bayesian Graph Convolutional Neural Networks[C]//Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM, 2020: 2030-2039. DOI: 10.1145/3394486.3403254.
- [3] 王绍卿, 李鑫鑫, 孙福振, 等. 个性化新闻推荐技术研究综述[J]. *计算机科学与探索*, 2020, **14**(1): 18-29. DOI: 10.3778/j.issn.1673-9418.1908024.
- [4] 叶凡, 孙玉, 陈崇成, 等. 基于地理标记照片的个性化景点推荐方法[J]. *地球信息科学学报*, 2021, **23**(8): 1391-1400. DOI: 10.12082/dqxxkx.2021.200608
- YE F, SUN Y, CHEN C C, et al. A Personalized Attraction Recommendation Method Based on Geotagged Photos[J]. *J Geo Inf Sci*, 2021, **23**(8): 1391-1400. DOI: 10.12082/dqxxkx.2021.200608
- [5] 王伟, 郭丽环, 何翎, 等. 众筹项目的个性化推荐: 考虑本地偏好的协同过滤算法[J]. *管理工程学报*, 2022, **36**(2): 204-214. DOI: 10.13587/j.cnki.jieem.2022.02.018.
- WANG W, GUO L H, HE L, et al. A Home Bias Enhanced Collaborative Filtering Algorithm for Crowdfunding Campaigns[J]. *J Ind Eng Eng Manag*, 2022, **36**(2): 204-214. DOI: 10.13587/j.cnki.jieem.2022.02.018.
- [6] LI W J, LUO P H. Diversified Recommendation Algorithm Integrating Kernel Density Estimation and Singular Value Decomposition[J]. *J Chin Comput Syst*, 2020, **41**(1): 56-60. DOI: 10.3969/j.issn.1000-1220.2020.01.011
- [7] GOGNA A, MAJUMDAR A. DiABIO: Optimization Based Design for Improving Diversity in Recommender System[J]. *Inf Sci*, 2017, **378**: 59-74. DOI: 10.1016/j.ins.2016.10.043.
- [8] AN Y H, DONG Q, YUAN Q, et al. Improving Recommendation Diversity by Highlighting the ExTrA Fabricated Experts[J]. *IEEE Access*, 2020, **8**: 64422-64433. DOI: 10.1109/ACCESS.2020.2984365.
- [9] YU T, GUO J P, LI W H, et al. Recommendation with Diversity: an Adaptive Trust-aware Model[J]. *Decis Support Syst*, 2019, **123**: 113073. DOI: 10.1016/j.dss.2019.113073.
- [10] WANG X K, SHENG Y Q, DENG H J. Joint Deep Network with Auxiliary Semantic Learning for Popular Recommendation[J]. *IEEE Access*, 2020, **8**: 41254-41261. DOI: 10.1109/ACCESS.2020.2976498.
- [11] DONG X, NI J C, CHENG W, et al. Asymmetrical Hierarchical Networks with Attentive Interactions for Interpretable Review-based Recommendation[J]. *Proc AAAI Conf Artif Intell*, 2020, **34**(5): 7667-7674. DOI: 10.1609/aaai.v34i05.6268.
- [12] FU K, LIANG S Q, LI B. Commodity Recommendation Model Based on Improved Deep Q Network Structure[J]. *J Comput Appl*, 2020, **40**(9): 2613-2621. DOI: 10.11772/j.issn.1001-9081.2019112002.
- [13] QIN J, ZHANG Q B, WANG B. Recommendation Method with Focus on Long Tail Items[J]. *J Comput*

- Appl*, 2020, **40**(2): 454–458. DOI: 10.11772/j.issn.1001-9081.2019091665.
- [14] GAN L, NURBAKOVA D, LAPORTE L, *et al.* Enhancing Recommendation Diversity Using Determinantal Point Processes on Knowledge Graphs[C]//Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2020: 2001–2004. DOI: 10.1145/3397271.3401213.
- [15] LIU Y, XIAO Y T, WU Q, *et al.* Diversified Interactive Recommendation with Implicit Feedback[J]. *Proc AAAI Conf Artif Intell*, 2020, **34**(4): 4932–4939. DOI: 10.1609/aaai.v34i04.5931.
- [16] WU Q, LIU Y, MIAO C Y, *et al.* PD-GAN: Adversarial Learning for Personalized Diversity-promoting Recommendation[C]//Proceedings of the 28th International Joint Conference on Artificial Intelligence. New York: ACM, 2019: 3870–3876. DOI: 10.5555/3367471.3367579.
- [17] LIANG Y L, QIAN T Y, LI Q, *et al.* Enhancing Domain-level and User-level Adaptivity in Diversified Recommendation[C]//Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2021: 747–756. DOI: 10.1145/3404835.3462957.
- [18] 王喆. 深度学习推荐系统[M]. 北京: 电子工业出版社, 2020. WANG Z. Deep Learning Recommender System[M]. Beijing: Publishing House of Electronics Industry, 2020.
- [19] RENDLE S, FREUDENTHALER C, GANTNER Z, *et al.* BPR: Bayesian Personalized Ranking from Implicit Feedback[EB/OL]. arXiv Preprint: 1205.2618, 2012. <https://arxiv.org/abs/1205.2618>.
- [20] HE X N, LIAO L Z, ZHANG H W, *et al.* Neural Collaborative Filtering[C]//Proceedings of the 26th International Conference on World Wide Web. New York: ACM, 2017: 173–182. DOI: 10.1145/3038912.3052569.
- [21] HUANG P S, HE X D, GAO J F, *et al.* Learning Deep Structured Semantic Models for Web Search Using Clickthrough Data[C]//Proceedings of the 22nd ACM International Conference on Information & Knowledge Management. New York: ACM, 2013: 2333–2338. DOI: 10.1145/2505515.2505665.
- [22] CARBONELL J, GOLDSTEIN J. The Use of MMR, Diversity-based Reranking for Reordering Documents and Producing Summaries[C]//Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 1998: 335–336. DOI: 10.1145/290941.291025.
- [23] CHEN L M, ZHANG G X, ZHOU H N. Fast Greedy MAP Inference for Determinantal Point Process to Improve Recommendation Diversity[C]//Proceedings of the 32nd International Conference on Neural Information Processing Systems. New York: ACM, 2018: 5627–5638. DOI: 10.5555/3327345.3327465.