

融合商品潜在互补性发现的个性化推荐方法*

邵英玮^{1,2}, 张敏^{1,2}, 马为之^{1,2}, 王晨阳^{1,2}, 刘奕群^{1,2}, 马少平^{1,2}



¹(清华大学 计算机科学与技术系, 北京 100084)

²(清华大学 北京信息科学与技术国家研究中心, 北京 100084)

通讯作者: 张敏, E-mail: z-m@tsinghua.edu.cn

摘要: 结合领域内知识的个性化推荐算法在近年来受到了广泛关注,许多研究工作尝试将商品之间的关系(如互补关系等)融入到推荐算法中.对于商家而言,了解商品互补的关系能够帮助他们更好地制定定价策略;对于推荐算法而言,结合商品关系的推荐也更有可能会生成令人满意的结果,因此,如何挖掘商品间的互补关系是一个很有意义的研究问题.现有方法大多从用户历史中的“共同购买”发掘商品的互补关系,但是由于真实的购买场景非常复杂,得到的很可能仅仅是共现关系而不是互补关系.借鉴经济学的相关研究,提出了商品潜在互补性发现推荐模型(latent complementarity discovery model,简称 LCDM),试图从另一角度更准确地刻画商品间关系.首先,基于经济学理论中的需求交叉弹性(cross-price elasticity of demand),提出互补性发现模型(complementarity discovery model,简称 CDM)联合商品价格与购买历史来挖掘商品间的互补关系.在用户标注任务中,所提算法较已有方法在用户标注的一致性上提升了 10.6%.随后,基于此提出了融合商品互补关系的双重注意力机制推荐模型 LCDM.最后,在真实数据集上的对比实验结果表明,提出的 LCDM 推荐模型能够显著改善推荐的效果,在 Recall@5 和 NDCG@5 上分别有 54.5%和 125.8%的提升,验证了所提方法的有效性.

关键词: 推荐系统;表示学习;商品关系;计算经济学

中图法分类号: TP391

中文引用格式: 邵英玮,张敏,马为之,王晨阳,刘奕群,马少平.融合商品潜在互补性发现的个性化推荐方法.软件学报,2020,31(4):1090-1100. <http://www.jos.org.cn/1000-9825/5925.htm>

英文引用格式: Shao YW, Zhang M, Ma WZ, Wang CY, Liu YQ, Ma SP. Integrating latent item-item complementarity with personalized recommendation systems. Ruan Jian Xue Bao/Journal of Software, 2020,31(4):1090-1100 (in Chinese). <http://www.jos.org.cn/1000-9825/5925.htm>

Integrating Latent Item-item Complementarity with Personalized Recommendation Systems

SHAO Ying-Wei^{1,2}, ZHANG Min^{1,2}, MA Wei-Zhi^{1,2}, WANG Chen-Yang^{1,2}, LIU Yi-Qun^{1,2}, MA Shao-Ping^{1,2}

¹(Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)

²(Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing 100084, China)

Abstract: Domain-specific personalized recommendation algorithm is getting more popular nowadays. In particular, item-item relationship (e.g. complementary good, substitute good) has already been considered in the development of recommendation algorithms. In terms of its potential application for sellers, the ability to notice actual item-item complementarity from data is of paramount importance, as it helps sellers to gain a market competitive advantage via designing better pricing strategies (e.g. bundling or pricing

* 基金项目: 国家自然科学基金(61672311, 61532011); 国家重点研发计划(2018YFC0831900)

Foundation item: National Natural Science Foundation of China (61672311, 61532011); National Key Research and Development Program of China (2018YFC0831900)

本文由“非经典条件下的机器学习方法”专题特约编辑高新波教授、黎铭教授、李天瑞教授推荐.

收稿时间: 2019-05-30; 修改时间: 2019-07-29; 采用时间: 2019-09-20; jos 在线出版时间: 2020-01-10

CNKI 网络优先出版: 2020-01-14 09:53:27, <http://kns.cnki.net/kcms/detail/11.2560.TP.20200114.0953.010.html>

discount). For recommender systems, integrating algorithms with the item-item relationship is also more likely to generate better recommendation results. Therefore, how to mine item-item complementarity is a research problem deserving of study. Even though most existing methodologies leverage on co-occurrence relationship, yet, the recommendation accuracy might easily be adversely affected by noise data due to the complex dynamics in the online shopping environment. In light of the research on economics, the latent complementarity discovery model (LCDM) is proposed in an attempt to more accurately describe the item-item relationship from a different perspective. Specifically, complementarity discovery model (CDM) is firstly proposed based on cross-price elasticity of demand, which jointly utilizes item pricing and purchase history to discover item-item complementarity relationship. Comparing with existing mining methods based on item co-occurrence relationship, the proposed method yields 10.6% increase in user label consistency. Then, LCDM is constructed by integrating dual-item attention with item-item complementarity insight mined from CDM. Lastly, from the comparison experiments conducted on real-world dataset, LCDM has made a significant improvement in recommendation performance, in which there is a 54.4% and 125.8% increase in Recall@5 and NDCG@5 respectively.

Key words: recommender system; representation learning; item-item relationship, computational economics

在新零售时代,随着各种生鲜购物平台的崛起,如国外的亚马逊生鲜,国内的盒马和每日优鲜等等,购物变得更为便利,越来越多原本在零售店的食物和杂货交易开始转为在线上发生.与传统电商相比,用户在零售店购物时更关注购买的必要性^[1],以满足他们日常所需,所以该领域产品个性化推荐的应用变得尤为重要.精准推荐一方面能够帮助用户节省大量商品比较、搜索时间,另一方面能让用户在海量同类型商品中发现潜在需求.已有推荐中比较关注商品间的“共同购买”关系,它是商品互补性的一种体现.而根据经济学的相关定义,互补品是指两个商品存在依赖关系,互为互补品的两个商品不仅有着正向的需求关系,一个商品的价格上升还会导致另外一个商品的需求量下降^[2].对于商家而言,互补品在经营策略中也得到了广泛的应用,其中两种普遍的策略包括:(1) 捆绑式销售:商家把一组不同类型、属于互补品的产品打包售出^[3];(2) 交叉补贴:商家通过有意把互补品中的一个降价而达到促销盈利的目的^[2].

然而,在真实场景中,策略的表现也受制于互补品的关系紧密程度.其一是用户和互补品的相对性^[2],对于不同的用户人群,在不同场景下有不同的互补商品组合.例如,对于健身人群而言,牛奶和蛋白粉是一种互补品,但是对孕妇人群就可能是牛奶和奶瓶的组合,所以缺乏对互补品和互补者的全面理解会为商家的价格策略带来障碍.其二,一个特定商品可能有多个互补商品^[2],针对不同的用户重要性也有所不同.大部分互补品组合相互依赖程度的微小差异单凭人的经验知识难以区分,但会影响到商家商品降价的选择.即使是同一组互补品(如洗衣机和洗衣粉),由于用户在现今网络发达的年代拥有更完备的商品认识,用户也较倾向于把两者的购买独立考量,对于两者都有个别的品牌偏好,商品与商品之间的互补联系也变得更加模糊,导致辨别互补产品变得更具挑战性.

基于以上所述,商品的互补性发现是一个重要且具有挑战性的问题,本文针对此问题提出了商品潜在互补性发现推荐模型(latent complementarity discovery model,简称 LCDM).首先,与单纯考量共同购买不一样,本文提出的商品的潜在互补品发现算法利用了需求交叉弹性(cross-price elasticity of demand)的经济学原理,综合考虑了商品价格与用户购买行为.随后,根据所发现的互补性关系,本文进一步提出了结合双重商品注意力机制和深度学习网络的 LCDM 模型来个性化预测同一订单用户最有可能会购买的商品,并在真实移动支付数据集上的对比实验验证了所提出模型的有效性.

本文主要有以下 3 方面贡献:第一,将经济学理论中的需求交叉弹性用于商品的互补关系挖掘.第二,我们设计了潜在互补性发现推荐模型,能够将商品的互补关系与双重注意力机制融入到个性化推荐算法中.第三,提出的算法在真实的用户购买记录上较传统方法取得了显著的效果提升.

1 相关工作

1.1 个性化商品推荐

购物平台往往有海量的商品,但每位用户往往只对少部分商品感兴趣.个性化推荐系统(personalised

recommender system,简称 PRS)则扮演了为用户进行商品信息匹配的角色.推荐系统也已被应用到多媒体推荐^[4]、新闻推荐^[5]和商品推荐^[6]等诸多场景.其中,协同过滤(collaborative filtering,简称 CF)方法被广泛使用,特别是在 Netflix 比赛中表现突出的隐性因子模型(latent factor model).矩阵分解(MF)在隐性因子模型中起到了核心作用,也被应用在许多真实场景的推荐系统^[7,8]中.随后,Rendle 等人^[9]提出的依靠成对方法(pairwise approach)学习的贝叶斯个性化排序(Bayesian personalized ranking from implicit feedback,简称 BPR)算法在利用矩阵分解方法的同时,使用成对的损失函数对用户在不同物品的相对偏好作排序优化.而 BPR 在实际场景的成功也使得成对训练的损失函数在推荐系统被广泛用于模型训练^[10].除了分析用户与商品的交互历史,近年来推荐系统也尝试引入额外信息优化算法表现,例如商品评论^[11]、社交关系^[12]、图像信息^[13]等等.物品之间的关系也备受关注,尤其是商品间的互补关系.直观来说,互补品对用户而言是两个功能相互补足的商品(例如乒乓球和乒乓球拍),所以用户往往会倾向于共同购买此类商品^[2].基于这种特点,有研究尝试从用户的浏览或同时购买行为数据中挖掘出此类型的关系^[14],还有一部分的工作是分别考虑商品描述和评论^[15]、商品类别^[16]以及图像的信息来更准确地识别互补品.但是这些工作的识别效果容易受到行为噪声影响,且很多时候共现不等于互补(如:用户可能会同时购买啤酒和纸巾,但是它们并没有互补关系)^[15].此外,基于神经网络的隐向量学习方法^[1,17-19]也被应用于预测商品之间的互补关系,但是这些方法主要针对的是用户每次购买一个商品的场景,难以挖掘订单内商品之间的互补关系^[1].

1.2 基于深度学习的推荐算法

近年来,深度神经网络在各个领域,包括计算机视觉、自然语言处理以及语音识别方面都取得了巨大的成功^[20].因此不少研究工作尝试结合各种不同的神经网络结构来改善推荐系统的性能.在此过程中,He 等人^[21]提出了神经协同过滤(neural collaborative filtering,简称 NCF)框架,旨在通过联合矩阵分解和多层前馈神经网络学习用户和物品的隐向量,从而预测用户对商品的评分.随后,神经分解机(neural factorization machines,简称 NFM)^[22]也被提出用于针对高维和非线性特征交互的建模以增强原有的分解机模型(factorization machine,简称 FM).但是已有的大部分推荐工作都没有考虑商品之间的关系.

特别地,神经网络中的注意力机制已在各种机器学习任务,如视频/图像描述和自然语言处理中显示出它的有效性^[23,24].其中,软注意力的关键思想是给一组特征学习分配合适的权重(0~1 之间),特征权重越高,代表对于学习目标的信息量越大,反之亦然.例如在机器翻译领域,注意力机制能够在生成对应文字的任务上找到句子中哪些文字的重要程度是比较高的.在推荐系统领域中,也有工作在 CF 中组件级别和物品级别的模块分别引入了注意力机制,并应用于多媒体的推荐场景^[25].另外一个例子是 NFM 模型的延伸——AFM 模型(attention factorization machine)^[26],它同时使用了注意力机制区分不同特征交互的重要程度.鉴于之前一系列工作反映出的注意力机制广泛的应用性和可解释性,本文也将应用注意力机制进行特征的融合.

1.3 基于经济学理论的推荐算法

在实际场景中,各种推荐算法能够带来巨大的经济效应,但是就我们所知,已有算法与经济学相关理论结合得并不紧密,仅有少数推荐系统研究进行了尝试.其中有工作^[27]在商品推荐中考虑了边际效用递减法则(the law of diminishing marginal utility)以区分消耗品和耐用品.另外,在考虑价格敏感性(price sensitivity)方面,有部分工作可估计出客户在电商平台的预期支付意愿(willingness-to-pay)^[28]或商品的评论信息区分价格敏感程度相对比较高的客户,从而进行商品的个性化定价^[29],以提高商家的收益.也有研究^[30]为了提高整体社会公益(social surplus),在推荐系统中融入了总盈余(total surplus)的概念以更好地匹配卖家和用户.此外,也有工作^[7,31]以电商平台中“共同购买”和“共同浏览”的信息作为互补品(complementary good)和替代品(substitute good)的标签,然后用监督学习算法来识别商品间的关系类型.本文的研究重点在于互补品的关系挖掘,它也在经济学领域中被广泛关注^[32-34].受已有经济学理论的启发,本文将尝试融合需求交叉弹性理论挖掘商品的互补关系,并将其应用在推荐系统中.据我们所知,这是属于少数融合经济理论的推荐算法研究工作.

2 基于需求交叉弹性的互补品发现方法

本文主要利用经济学理论的需求交叉弹性来进行互补品的挖掘^[2].需求交叉弹性主要描述的是两个商品价格与需求量变化的联系.具体来说,假设其他因素保持不变,另外一种商品的价格变化百分之一会导致本产品需求量百分比的改变^[2].这一概念用于计算两种商品之间是否存在互补关系,如果有互补关系,一种商品价格的变化会让另外一种商品的需求量产生交叉影响.需求交叉弹性的定义如下:

$$E_{AB} = \frac{\frac{\Delta Q_{dA}}{Q_{dA}}}{\frac{\Delta P_B}{P_B}} = \frac{\Delta Q_{dA}}{\Delta P_B} \cdot \frac{P_B}{Q_{dA}} \quad (1)$$

其中, E_{AB} 为商品 A 相对于商品 B 的需求交叉弹性, Q_{dA} 为商品 A 的需求量, ΔQ_{dA} 为商品 A 需求的变化量, P_B 为商品 B 的价格, ΔP_B 为商品 B 的价格变化量.实际计算时,我们使用弧弹性计算公式计算 E_{AB} :

$$E_{AB} = \frac{\frac{\Delta Q_{dA}}{Q_{dA1} + Q_{dA2}}}{\frac{\Delta P_B}{P_{B1} + P_{B2}}} = \frac{\Delta Q_{dA}}{\Delta P_B} \cdot \frac{P_{B1} + P_{B2}}{Q_{dA1} + Q_{dA2}} \quad (2)$$

其中, Q_{dA1} 为商品 A 开始时的需求量, Q_{dA2} 为商品 A 结束时的需求量, P_{B1} 为商品 B 开始时的价格, P_{B2} 为商品 B 结束时的价格.需求交叉弹性的计算结果可正可负,为正,代表两个商品价格和需求量的变化为同方向(例:一种商品价格的上升会带来另一种商品需求量的上升),说明两种商品有着替代品的关系;为负,代表商品价格和需求量的变化为反方向(例:一种商品价格的上升会带来另外一种商品需求量的下降),说明两种商品有着互补品的关系.此外,数值的大小也同时说明商品间的依赖程度, E_{AB} 的数值越大,商品 B 对 A 的替代性越强.

我们将这种基于需求交叉弹性理论的互补品计算方法命名为潜在互补品发现算法(complementarity discovery model,简称 CDM),它可以从商品的历史购买数据中计算不同商品之间的潜在关系.相较于已有的互补品发现算法,这一方法并不需要依赖于用户的评论内容,也不是单纯的计算商品共现,而是从经济学角度计算商品间是否存在互补关系.此外,对于每个商品 A 而言,我们可以将其他商品根据 E_{AB} 排序而得到互补性序列,从而为任意商品得到互补商品的结果,弥补了已有互补品发现算法的数据稀疏性问题.

3 融合互补品关系的双重注意力机制推荐模型

本文所使用的变量符号表定义见表 1.

Table 1 Symbol annotation

表 1 符号描述

符号	描述	符号	描述
U	用户的集合	V	商品的集合
T	商品订单的集合	R	用户和商品的交互(购买关系)
X	用户和订单的交互(购买关系)	Y	商品与订单的关系(属于哪个订单)
Z	商品与商品的互补关系	t_k	商品订单 k 的隐因子
u_i	用户 i 的隐因子	v_j	商品 j 的隐因子
$u_i^{(j)}$	针对商品 j 用户 i 的隐因子	$v_j^{(i)}$	针对用户 i 商品 j 的隐因子
$score_j$	正例商品 j 的分数	$score_b$	负例商品 b 的分数
$r(i,j)=1$	用户 i 购买过商品 j	$x(i,k)=1$	用户 i 已下订单 k
$y(j,k)=1$	商品 j 在订单 k 出现过	$z_{(j_1,j_2)}=1$	商品 j_1 和 j_2 有互补关系
$\bar{v}_{(j,l)}$	商品 j 第 l-th 个互补商品的隐因子	$t(j,s)$	商品 j 所在的第 s-th 个订单的隐因子
$\gamma_{(j,l,i)}$	对于用户 i 商品 j 的第 l-th 个互补商品的影响程度	$\beta_{(j,s,i)}$	对于用户 i 商品 j 的第 s-th 个所在订单的影响程度

本文的目标是将商品的互补关系计算结果融入到个性化推荐算法中,算法主要利用了商品的互补关系、所在订单以及用户历史购买来进行用户偏好的建模.假设有 n 个用户 $U=(u_1, u_2, \dots, u_n)$, m 个商品 $V=(v_1, v_2, \dots, v_m)$ 和 w 个订单 $T=(t_1, t_2, \dots, t_w)$.其中,用户、商品、订单存在 4 种关系,分别是用户和商品的交互关系 $R=[r_{ij}]_{n \times m} \in \{0,1\}$,代表 u_i 是否购买过 v_j ,用户和订单的交互关系 $X=[x_{ik}]_{n \times w} \in \{0,1\}$,代表 u_i 是否下过订单,商品与订单的关系 $Y=[y_{js}]_{m \times w} \in \{0,1\}$,代表商品 v_j 是否在订单 t_k 中出现过,商品与商品的互补关系 $Z=[z_{(j_1, j_2)}]_{m \times m} \in \{0,1\}$,代表商品 v_{j_1} 与 v_{j_2} 是否有互补关系.给定一个目标用户 u_i ,任务是要推荐一系列同一订单用户 u_i 感兴趣的物品,可定义为:

输入:用户集合 U ,商品集合 V ,订单集合 T ,用户和商品交互的关系 R ,用户和订单的交互关系 X ,商品与订单的关系 Y 以及商品与商品的互补关系 Z .

输出:针对每个用户 u_i 的个性化排序结果.

3.1 潜在互补性发现模型概述

潜在互补性发现推荐模型(latent complementarity discovery model,简称 LCDM)是一个端到端的深度学习模型(如图 1 所示),特点在于融合了商品之间的互补关系、商品订单信息,同时加入了双重注意力机制.模型主要分为用户和商品两大模块,然后通过成对学习的训练方法学习能够最佳区分正例和负例商品的参数.下面我们将以先商品、后用户的顺序分别加以详细的介绍.

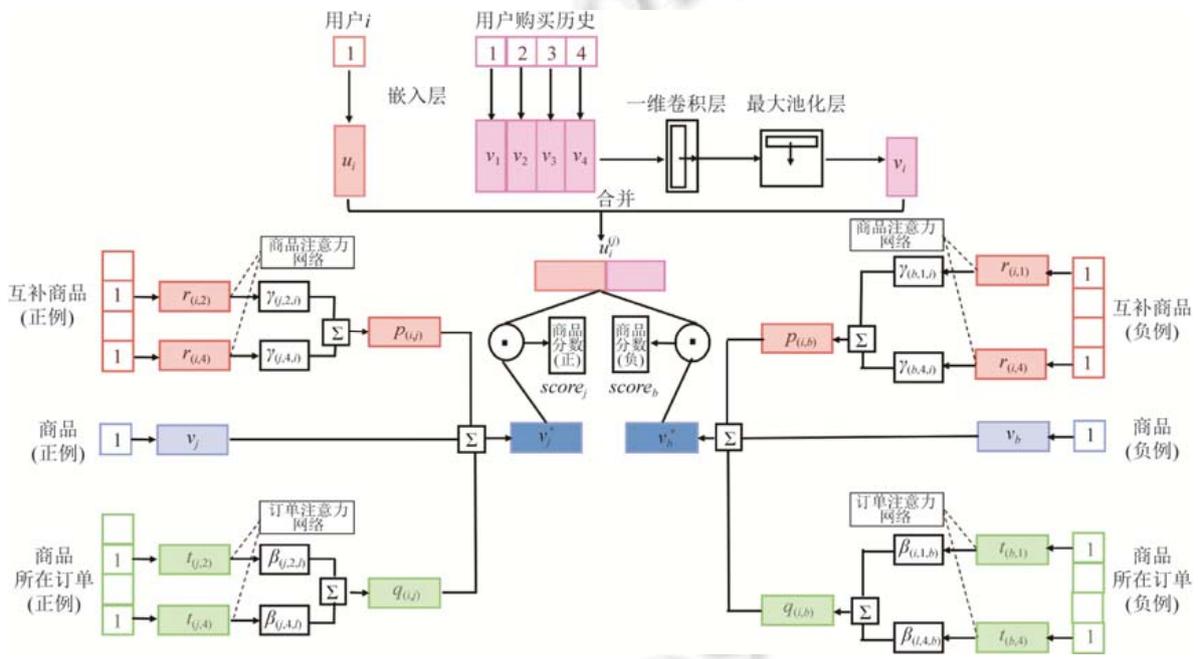


Fig.1 The graphical representation of latent complementarity discovery model (LCDM)

图 1 潜在互补性发现推荐模型(LCDM)的示意图

3.2 候选商品的多粒度特征提取

3.2.1 商品互补关系的融合

为了降低数据稀疏性和噪声的影响,我们先对数据进行了预处理.首先,为了避免冷启动商品影响推荐结果,我们保留了最少被 M 次购买的商品.其次,在计算互补品时我们只保留最少被共同购买过 N 次的商品,因为一方面存在互补关系的商品更倾向于被共同购买或使用,另外一方面也可以减少其他因素带来的随机性.在筛选过后,本文利用第 2 节提到的互补性发现算法 CDM 对每一个候选目标商品计算其与其他商品的需求交叉弹性,并且取值最小的 k 个商品作为与该商品具有互补关系的商品集合 R_j .每一个商品我们用一个 d 维的隐向量

加以表示(该向量将在模型训练时学习更新).为了进一步考虑不同互补商品造成的影响程度,本文加入了注意力机制,学习每一个互补商品特征的权重,计算公式如下:

$$\gamma_{(j,l,i)}^* = h_r^T \text{ReLU}(W_1 \tilde{v}_{(j,l)} + W_2 v_j^{(i)} + b_r) \quad (3)$$

其中, $W_1 \in R^{d_{xw}}$, $W_2 \in R^{d_{xw}}$, $b_r \in R^k$, $h_r \in R^k$ 是注意力模型的参数.为了得到商品互补品的特征,最后得出的商品注意力会被指数函数归一化(softmax),即:

$$\gamma_{(j,l,i)} = \frac{\exp(\gamma_{(j,l,i)}^*)}{\sum_{l \in R_j} \exp(\gamma_{(j,l,i)}^*)} \quad (4)$$

得到注意力网络的权重后,本文通过加权求和计算出当前待推荐商品互补品的特征向量 $p_{(i,j)}$,并在对商品推荐时加以使用,计算公式如下:

$$p_{(i,j)} = \sum_{l \in R_j} \gamma_{(j,l,i)} \tilde{v}_{(j,l)} \quad (5)$$

3.2.2 商品所在订单特征

在实际购物场景中,同一个订单的商品整体更可能反映用户的单一购买意愿,例如在准备火锅的时候,订单大部分以新鲜的蔬菜、鲜肉、丸子、火锅汤料和饮料等等有互补功能的商品组合为主.基于这一考虑,本文也将为每一个订单学习一个 d 维的隐向量,然后同样使用注意力网络学习所在订单里哪些更能反映待推荐商品的功能,计算公式如下:

$$\beta_{(j,s,i)}^* = h_t^T \text{ReLU}(W_1 t_{(j,s)} + W_2 v_j^{(i)} + b_t) \quad (6)$$

其中, $W_1 \in R^{d_{xw}}$, $W_2 \in R^{d_{xw}}$, $b_t \in R^k$, $h_t \in R^k$ 是订单注意力机制的参数,最后的输出会被指数函数归一化:

$$\beta_{(j,s,i)} = \frac{\exp(\beta_{(j,s,i)}^*)}{\sum_{s \in T_j} \exp(\beta_{(j,s,i)}^*)} \quad (7)$$

得到注意力网络的权重后,本文通过加权求和计算出待推荐商品曾出现的订单特征 $q_{(i,j)}$,公式如下:

$$q_{(i,j)} = \sum_{s \in T_j} \beta_{(i,j)} \mathbf{t}_{(i,s)} \quad (8)$$

其中, T_j 代表了包含商品 j 的订单集合.

3.2.3 商品基本特征

我们为每个商品 j 学习一个 d 维的隐向量 v_j 作为商品的基本特征.在推荐中我们会将它与互补商品的特征向量和商品所在订单的特征向量相加,从而得到最终商品的表示向量 V_j^* .

$$V_j^* = v_j + p_{(i,j)} + q_{(i,j)} \quad (9)$$

3.3 用户的长短期偏好建模

3.3.1 用户长期基本特征

我们为每位用户学习一个 d 维的隐向量,作为用户的基本特征 u_i .

3.3.2 用户短期购买历史

为了考虑用户的短期购买需求,本文在模型中考虑了每一个用户最近期购买的 Y 个商品,每个商品用 d 维的隐向量表达拼接成矩阵,之后将该矩阵输入到卷积神经网络(convolutional neural network,简称 CNN)中^[35],将其压缩学习成一个短期兴趣向量,然后再和用户的基本特征(用户长期偏好)进行求和,得到最终的用户表示向量 u_i^j .

3.4 模型训练

我们使用用户向量与商品向量点积的方式预测用户对该商品的评分,即:

$$\text{score}_j = V_j^* \odot u_i^j \quad (10)$$

本文采用了成对的训练方法,故此我们按照上述方法分别得到正例商品(被用户购买过的商品)以及被随机

负采样的负例商品(没有被用户购买过的商品)的评分输出.由于训练目标是希望正例商品的得分应该尽可能地比负例商品的得分要高,因此目标函数如下:

$$L(Y, f(x)) = -\log(\sigma(\text{score}_j - \text{score}_b)) \quad (11)$$

其中, score_j 和 score_b 分别代表了正例商品和负例商品的最终分数.我们使用 Adam 算法进行优化.

4 实验分析

4.1 数据集

本文使用了真实场景下的用户购买记录数据集,包含多个线下零售商店的移动支付记录.该数据集的时间跨度从 2017/5/28~2018/3/6(约 10 个月),用户共购买过 734 066 件商品,数据的统计信息见表 2.在构造数据集时,我们先把购买过多及过少的用户先过滤掉,然后对于每一个订单随机抽取一个商品加入到验证集,一个商品加入到测试集,之后,其他商品用于训练.本文在训练时会为每一个正例商品都随机地在整个训练集中负采样一个负例商品.在验证和测试时,考虑到若对所有商品的得分作预测的话时间较长,于是我们预先对每个验证和测试样本构造了大小为 100 的候选商品集合,集合中包含用户在这个订单中购买过的正例(仅 1 个),其余则是随机采样的未被用户购买过的商品.

Table 2 Dataset statistics

表 2 数据统计

用户数量	商品数量	订单数量	订单平均商品数量
53 247	29 993	63 831	11.5

4.2 互补关系预测效果对比

为了验证提出的算法(CDM)在商品互补性发现任务上的有效性,本文设计了人工标注实验,具体在相关性、多样性和惊喜性 3 个维度上进行评估.两种对比方法分别是基于同时购买频率的互补品发现(norm_prob)和基于同时购买条件概率的互补品发现(cooccur).在实验中,我们首先利用 3 种方法(包括 CDM)分别计算商品间的互补关系,然后分别取出每种方法得到的 Top 150 个互补性最高的商品,把它们合并成一个包含 450 个商品对的列表.经过样本随机后,我们招募了 5 名被试者针对每个商品对回答 4 个问题,以评价互补性表现,其中,1 分代表同意,0 分代表不同意.问题包括:(1) 我有可能同时购买商品 A 和商品 B;(2) 我倾向于同时使用/食用商品 A 和商品 B;(3) 商品 A 和商品 B 很不相似;(4) 商品 A 和商品 B 的搭配出乎我意料,却感觉挺合适的.最终得到的实验标注结果 Kappa 系数>0.3,说明实验结果的一致性较高,可以用于之后的分析.3 种方法在 4 个问题上的得分对比结果如图 2 所示.

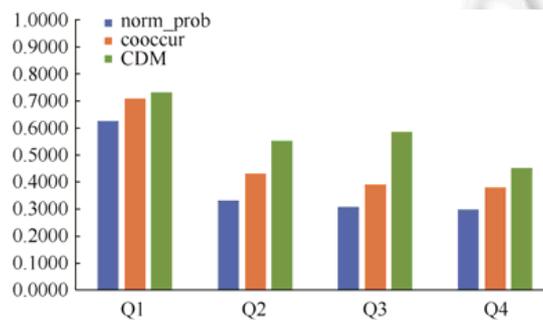


Fig.2 The results of the labeling experiment

图 2 人工标注实验的结果

图 2 的结果说明,本文提出的互补性发现算法 CDM 在互补商品对发现问题上,相对基于同时购买

(norm_prob)和同时购买频率的方法(cooccur),在商品相关性、多样性和惊喜度方面皆有显著的提升.在商品相关性方面,本文的方法在问题 1 和问题 2 上的显著提升说明了所提方法不仅能够发现拥有同时购买特性的互补品,还能发现更符合经济学定义的功能互补的互补品.在商品多样性方面,CDM 在问题 3 上取得了接近 50%的提升(相对于第二高的方法),说明了本文的方法有更大的概率能够发现不同种类的互补品(例如不同牌子/种类),也更贴近互补品的真实情况.此外,在惊喜度方面,CDM 在问题 4 上的表现更是说明了本文的方法除了能够发现更多样的互补品外,商品的组合同时也让用户觉得有新鲜感且是合理的.具体的例子如 3 种方法找出前 10 位的互补品中,CDM 能找到(火锅料,羊肉),基于同现购买频次能找到(可口可乐,雪碧)以及基于购买条件概率能找到(卡布奇诺咖啡,拿铁咖啡).本文方法找出的商品对明显更符合用户于火锅的场景下互补品.相反,另外两种方法找出的商品对都属于同一类别,两种商品更接近于替代品的关系.

4.3 实验对比的基准算法

- 通用矩阵分解(GMF)^[21]:此方法把用户和商品的嵌入向量按元素相乘,然后把其输出结果通过一层神经网络得到商品的最终评分;
- 多层感知机(MLP)^[21]:此方法把用户和商品的隐向量拼接输入到多层神经网络得到商品的最终评分;
- 基于深度神经网络的协同过滤(NCF)^[21]:此方法是当前最好的基于深度模型的推荐方法之一,其核心结合了矩阵分解(GMF)和多层感知器(MLP).在本实验中最终使用的参数如下:两个多层感知器里的嵌入层大小分别为 128 和 64,以及其正则化皆为 0.001,模型学习率为 0.000 1,L2 正则化为 0.05,dropout 为 0.25;
- 贝叶斯个性化排序(BPR)^[9]:此方法通过贝叶斯个性化排序的目标函数把矩阵分解(MF)进行优化.最终使用的参数如下:L2 正则化为 0.05,学习率为 0.000 7;
- 共现发现推荐模型(LCDM_{cooccur-noAtten}):此方法使用商品同现信息代替了 LCDM 的商品互补信息,但没有加上双重注意力机制;
- 基于注意力机制的共现发现推荐模型(LCDM_{cooccur}):此方法使用商品的共现信息代替了 CDM 计算的商品互补信息,并使用了双重注意力机制;
- 互补性发现推荐模型(LCDM_{comple-noAtten}):此方法只使用了基于 CDM 的商品互补关系,但没有加上双重注意力机制;
- 基于注意力机制的潜在互补性发现推荐模型(LCDM_{comple}):此方法是融合 CDM 计算的商品互补关系和双重注意力机制的推荐算法,也是本文提出的完整算法.模型的参数设置如下:用户、商品、订单的嵌入层大小皆为 16,批量大小为 256,学习率为 0.000 8.

4.4 评价指标

我们使用了召回(Recall@K)和归一化折损累计增益(NDCG@K)来进行推荐算法评价.Recall@K 是考虑了有多少正例在 Top K 个位置被召回,而 NDCG@K 则同时考虑正例在 Top K 个的排序的具体位置.其中,IDCG 是指最理想情况下的 DCG.两种指标的计算公式如下:

$$\text{Recall}@K = \frac{\sum_{i=1}^K \text{rel}_i}{\min(K, |y_u^{\text{test}}|)}, \text{DCG}@K = \sum_{i=1}^K \frac{2^{\text{rel}_i} - 1}{\log_2(i+1)}, \text{NDCG}@K = \frac{\text{DCG}@K}{\text{IDCG}@K}.$$

4.5 结果与分析

表 3 显示了 5 个基准模型和本文提出的方法在各个指标上的表现,其中,LCDM_{comple} 的效果相比所有其他模型都有显著的提升.首先,我们发现 NCF 要优于它的分割模型 GMF 和 MLP.比较 NCF 和 BPR 的算法表现,虽然 NCF 有比较强的表达能力,但是 NCF 需要更大规模的用户和交互数据用于参数学习,再加上本数据集相对较为稀疏,所以 NCF 效果要略低于 BPR.随后,我们将 NCF 和 BPR 与 LCDM_{comple} 对比后发现,随着商品互补关系的引入,推荐算法在 Recall 和 NDCG 指标上均取得了显著提升.Recall 的提高说明,本文提出的方法可比其他方法找到更多的正例商品,而 NDCG 由于考虑了排名的位置,该指标上更为显著的提升也说明了模型对于商品的

排序性能有大幅提升,整体返回列表的质量更高.

实验中我们也通过对 $LCDM_{comple}$ 的核心模块进行删减构建对比模型,主要针对模型的两个创新点进行效果验证:(1) 基于需求交叉弹性发现互补品的效果;(2) 双重注意力机制的影响.第一,通过与 $LCDM_{coocur-noAtten}$ 和 $LCDM_{comple-noAtten}$ 的对比,我们发现使用需求交叉弹性得到的商品互补信息效果明显比使用商品同现信息要好,说明使用本文的 CDM 方法能够比商品同现关系更有效地刻画商品间的互补关系,并且能够在推荐系统取得更好的表现.第二,从 $LCDM_{coocur}$ 和 $LCDM_{comple}$ 的效果对比可以看出,注意力机制能够有效帮助模型在商品和订单层面上学习出刻画商品的相关特征.

Table 3 Comparison of different methods on dataset

表 3 不同实验方法在各数据集中的表现

Model	Recall@1	Recall@5	Recall@10	NDCG@1	NDCG@5	NDCG@10
GMF	0.076 8	0.270 5	0.408 2	0.076 8	0.174 5	0.218 8
MLP	0.096 7	0.305 6	0.448 7	0.096 7	0.202 6	0.248 7
NCF	0.115 4	0.338 0	0.482 8	0.115 4	0.228 1	0.272 5
BPR	0.133 4	0.348 5	0.495 2	0.133 4	0.236 4	0.283 7
$LCDM_{coocur-noAtten}$	0.439 5	0.443 5	0.454 4	0.437 7	0.439 7	0.443 4
$LCDM_{coocur}$	0.472 1	0.476 5	0.488 9	0.470 5	0.472 8	0.477 0
$LCDM_{comple-noAtten}$	0.451 3	0.455 3	0.466 8	0.449 5	0.451 6	0.455 4
$LCDM_{comple}$	0.532 9	0.538 3	0.551 9	0.531 0	0.533 8	0.538 5

总结来看,本文提出的方法对于新零售线上平台食物杂货购买的场景具有独特意义.因为在线上平台带来便利的前提下,大部分的客户在进行购买时已经较清楚自己需要什么类别的商品,相对传统电商的探索式购物来说有更为单一的购买倾向.所以用户注重的是在众多相似的同类型商品下能否寻找到与同一订单下功能/体验互补的商品.所以 $LCDM_{comple}$ 更高的 recall 说明本文的方法能够找到更多与用户偏好相匹配的商品.此外,由于用户在较高频的购物情况下需要快捷的搜索体验,加上大部分用户习惯从上而下阅读,相关商品在位置上排得明显靠前能够有效地缩短用户浏览的时间.实验结果显示,本文的方法在 NDCG 上有大幅提升,说明 $LCDM_{comple}$ 能够更准确地为用户在靠前的位置提供相关的商品,缩短浏览时间.因此本文的方法确实能够在杂货购物的实际场景中为用户带来便利,显著优化线上购物的体验.

5 结论与展望

在推荐领域,除了挖掘用户兴趣与购买规律,商品间的关系能够进一步帮助提升其表现.本文提出潜在互补性发现推荐模型(LCDM),能够从经济学理论中的需求交叉弹性出发,重点使用商品价格和交易信息计算得到商品之间的互补关系.我们基于这些信息设计的双重注意力机制能够同时考虑互补商品以及商品订单的信息,更好地刻画了商品特征.最终提出的算法较已有的推荐算法在真实数据集上取得了显著提升.未来,我们会尝试将时间因素加入到模型中,以进一步优化推荐系统的表现.

References:

- [1] Wan M, Wang D, Liu J, Bennett P, McAuley J. Representing and recommending shopping baskets with complementarity, compatibility and loyalty. In: Proc. of the 27th ACM Int'l Conf. on Information and Knowledge Management. ACM, 2018. 1133-1142.
- [2] O'sullivan A, Sheffrin SM. Economics: Principles in Action. Prentice Hall, 2003.
- [3] Bakos Y, Brynjolfsson E. Bundling and competition on the Internet. Marketing Science, 2000,19(1):63-82.
- [4] Covington P, Adams J, Sargin E. Deep neural networks for Youtube recommendations. In: Proc. of the ACM RecSys. 2016. 191-198.
- [5] Yi X, Hong L, Zhong E, Liu NN, Rajan S. Beyond clicks: Dwell time for personalization. In: Proc. of the ACM RecSys. 2014. 113-120.
- [6] Wang P, Guo J, Lan Y, Xu J, Cheng X. Your cart tells you: Inferring demographic attributes from purchase data. In: Proc. of the WSDM. 2016. 173-182.

- [7] LiorRokach FR, Shapira B. Introduction to recommender systems handbook. In: Recommender Systems Handbook. 2011. 1–35.
- [8] Su XY, Khoshgoftaar TM. A survey of collaborative filtering techniques. In: Advances in Artificial Intelligence. 2009. 4.
- [9] Rendle S, Freudenthaler C, Gantner Z, Schmidt-Thieme L. BPR: Bayesian personalized ranking from implicit feedback. In: Proc. of the 25th Conf. on Uncertainty in Artificial Intelligence. AUAI Press, 2009. 452–461.
- [10] Chen JY, Zhang HW, He XN, Liu W, Liu W, Tat SC. Attentive collaborative filtering: Multimedia recommendation with item- and component-level attention. In: Proc. of the SIGIR. 2017. 335–344.
- [11] Chen C, Zhang M, Liu Y, Ma S. Neural attentional rating regression with review-level explanations. In: Proc. of the 2018 World Wide Web Conf. 2018. 1583–1592.
- [12] Chen C, Zhang M, Liu Y, Ma S. Social attentional memory network: Modeling aspect-and friend-level differences in recommendation. In: Proc. of the 12th ACM Int'l Conf. on Web Search and Data Mining. ACM, 2019. 177–185.
- [13] Kwon OB. “I know what you need to buy”: CONTEXT-aware multimedia-based recommendation system. Expert Systems with Applications, 2003,25(3):387–400.
- [14] Zheng J, Wu X, Niu J, Bolivar A. Substitutes or complements: Another step forward in recommendations. In: Proc. of the 10th ACM Conf. on Electronic Commerce. ACM, 2009. 139–146.
- [15] McAuley J, Pandey R, Leskovec J. Inferring networks of substitutable and complementary products. In: Proc. of the 21st ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. ACM, 2015. 785–794.
- [16] Wang Z, Jiang Z, Ren Z, Tang J, Yin D. A path-constrained framework for discriminating substitutable and complementary products in e-commerce. In: Proc. of the 11th ACM Int'l Conf. on Web Search and Data Mining. ACM, 2018. 619–627.
- [17] Barkan O, Koenigstein N. Item2vec: Neural item embedding for collaborative filtering. In: Proc. of the 26th IEEE Int'l Workshop on Machine Learning for Signal Processing (MLSP). IEEE, 2016. 1–6.
- [18] Dong Y, Chawla NV, Swami A. Metapath2vec: Scalable representation learning for heterogeneous networks. In: Proc. of the 23rd ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. ACM, 2017. 135–144.
- [19] Vasile F, Smirnova E, Conneau A. Meta-prod2vec: Product embeddings using side-information for recommendation. In: Proc. of the 10th ACM Conf. on Recommender Systems. ACM, 2016. 225–232.
- [20] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature, 2015,521(7553):436.
- [21] He XN, Liao LZ, Zhang HW, Nie LQ, Hu X, Tat SC. Neural collaborative filtering. In: Proc. of the WWW. 2017. 173–182.
- [22] He XN, Tat SC. Neural factorization machines for sparse predictive analytics. In: Proc. of the SIGIR. 2017. 355–364.
- [23] Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate. arXiv Preprint arXiv: 1409.0473, 2014.
- [24] Rush AM, Chopra S, Weston J. A neural attention model for abstractive sentence summarization. arXiv Preprint arXiv: 1509.00685, 2015.
- [25] Chen JY, Zhang HW, He XN, Liu W, Liu W, Tat SC. Attentive collaborative filtering: Multimedia recommendation with item- and component-level attention. In: Proc. of the SIGIR. 2017. 335–344.
- [26] Xiao J, Ye H, He X, Zhang H, Wu F, Chua TS. Attentional factorization machines: Learning the weight of feature interactions via attention networks. arXiv Preprint arXiv:1708.04617, 2017.
- [27] Wang J, Zhang Y. Utilizing marginal net utility for recommendation in e-commerce. In: Proc. of the SIGIR. ACM, 2011. 1003–1012.
- [28] Zhao Q, Zhang Y, Friedman D, Tan F. E-commerce recommendation with personalized promotion. In: Proc. of the RecSys. 2015. 219–226.
- [29] Shiu YW, Guo C, Zhang M, Liu YQ, Ma SP. Identifying price sensitive customers in e-commerce platforms for recommender systems. In: Proc. of the 24th China Conf., CCIR 2018. 2018. [doi: 10.1007/978-3-030-01012-6_18]
- [30] Zhang Y, Zhao Q, Zhang Y, Friedman D, Zhang M, Liu Y, Ma S. Economic recommendation with surplus maximization. In: Proc. of the WWW. 2015. 73–83.
- [31] Ma W, Zhang M, Cao Y, *et al.* Jointly learning explainable rules for recommendation with knowledge graph. In: Proc. of the World Wide Web Conf. ACM, 2019. 1210–1221.

- [32] Shocker AD, Bayus BL, Kim N. Product complements and substitutes in the real world: The relevance of “other products”. Journal of Marketing, 2004,68(1):28–40.
- [33] Russell GJ, Bolton RN. Implications of market structure for elasticity structure. Journal of Marketing Research, 1988,229–241.
- [34] Russell GJ, Petersen A. Analysis of cross category dependence in market basket selection. Journal of Retailing, 2000,76(3): 367–392.
- [35] LeCun Y, Bengio Y. Convolutional networks for images, speech, and time series. In: The Handbook of Brain Theory and Neural Networks, 1995,3361(10).



邵英玮(1993—),男,香港人,硕士,主要研究领域为个性化推荐,计算机经济学.



王晨阳(1996—),男,博士,CCF 学生会会员,主要研究领域为个性化推荐.



张敏(1977—),女,博士,教授,博士生导师,CCF 高级会员,主要研究领域为信息检索,个性化推荐,用户行为分析,机器学习.



刘奕群(1981—),男,博士,教授,博士生导师,CCF 杰出会员,主要研究领域为网络信息检索,网络用户行为分析,自然语言处理.



马为之(1992—),男,博士,助理研究员,CCF 专业会员,主要研究领域为用户建模,个性化推荐.



马少平(1961—),男,博士,教授,博士生导师,主要研究领域为智能信息处理,信息检索.