

doi:10.19920/j.cnki.jmsc.2024.08.007

消费者决策旅程视角基于评论的深度推荐方法^①

张颖异^{1,2}, 李先能^{1,2*}, 郭艳红^{1,2}, 李晓刚^{1,2}, 郑爽^{1,2}

(1. 大连理工大学经济管理学院, 大连 116024; 2. 大连理工大学先进智能研究院, 大连 116024)

摘要: 推荐系统的本质是建模消费者行为中隐藏的兴趣偏好。人的行为离不开心理, 表象的行为背后蕴含着丰富的内在产生动机。然而, 当前的推荐系统主要聚焦于对消费者表象的行为数据建模层面, 很少涉及内在的心理活动和决策中对信息的加工过程, 难以解析行为背后的产生路径。为此, 本文研究了一种新的推荐系统方法创新思路, 聚焦于基于评论的推荐场景, 从消费者决策旅程视角出发, 引入 AIDMA 阶段式决策模型, 将消费者行为的产生过程视为对商品“引起注意→产生兴趣→唤醒欲望→留下记忆→决策行为”的决策旅程, 并将其运用到深度学习模型设计中, 提出一种新的推荐方法。实验表明本研究方法的推荐效果显著优于现有方法。本研究遵循了“模型驱动+数据驱动”的大数据驱动研究范式, 不仅实现了对推荐系统具有消费者行为理论模型支撑的深度学习方法创新, 还为理解消费者在决策过程中对非结构化评论文本的信息加工过程提供了借鉴和参考。

关键词: 推荐系统; 深度学习; 评论文本; 消费者决策旅程

中图分类号: C93 **文献标识码:** A **文章编号:** 1007-9807(2024)08-0105-21

0 引言

推荐系统作为解决电子商务环境下消费者信息过载问题的重要手段, 已成为了各大电商平台的“标配”, 长期以来广受学界和业界的重视。虽有各种不同的切入点, 当前推荐系统的方法研究主要聚焦于对消费者的历史数据进行深入分析与挖掘, 以刻画“千人千面”的个性化偏好, 从而向每一个消费者推荐匹配其兴趣的商品^[1]。在数据规模激增、算力大幅提升和算法不断成熟的背景下, 推荐系统的方法研究已由矩阵分解、决策树等传统机器学习方法全面进入深度学习阶段^[2], 各种深度推荐方法涌现^[3,4]。与此同时, 随着大量的用户评论文本、商品描述图片等非结构化数据的产生, 推荐系统的方法研究迎来了重要的机遇, 即如何从传统意义上难以测量的海量非结构化数据

中发现其蕴藏的稀疏价值, 这同时也是大数据时代下商务智能分析与挖掘的重要挑战^[5-9]。以用户评论文本为例, 社会比较理论^[10]指出消费者在缺乏客观认知的情况下, 往往会利用他人的评价信息以辅助自我评价。在电商环境下, 消费者普遍会通过阅读商品的评论本来获取辅助决策的信息, 将评论文本引入到推荐系统的方法建模中具有显著意义。因此, 基于评论的推荐 (review-aware recommendation)^[11] 近年来获得了越来越多的关注。

推荐系统的方法本质是数据驱动, 通过消费者的行为数据构建其与商品之间的内在联系, 通过对海量数据的特征学习实现消费者兴趣偏好的建模。因此, 推荐系统的方法研究实质上与消费者行为研究息息相关, 彼此的核心均是探究消费者是如何做出购买决策行为的。人的行为离不开心

① 收稿日期: 2021-12-15; 修订日期: 2023-01-27。

基金项目: 国家自然科学基金资助项目(72071029; 71974031; 72231010)。

通讯作者: 李先能(1986—), 男, 广东韶关人, 博士, 教授, 博士生导师。Email: xianneng@dlut.edu.cn

理,表象的行为背后往往蕴含着丰富的内在产生动机。然而,目前主流的推荐系统方法研究主要还是聚焦在对表象的行为数据建模上,通过端到端(end-to-end)的学习建立历史行为与未来决策的联系,往往不会涉及消费者行为背后的心理活动、决策过程中对信息的处理和加工等内容。因此,深度学习类的方法虽然推荐效果很好,但是几乎都不具备可解释性,无法知悉行为背后的产生动机和路径。

而在消费者行为领域,早在1968年恩格尔等^[12]即提出了经典的EKB消费者购买决策模式,从消费者心理活动、信息加工、决策过程以及外部环境四方面阐述外界信息刺激对消费者购买决策的影响。2009年,Court等^[13]进一步提出了消费者决策旅程(consumer decision journey)的概念,将消费者每一次的决策过程比喻为一次从起点到终点的决策旅程,在这个过程中外界信息的刺激潜移默化地匹配了消费者的兴趣,唤醒了消费者的欲望,并深刻的印在了消费者的脑海中,最终影响了决策行为的产生^[14]。消费者决策旅程从理论上诠释了信息内容在消费者体验中的重要性,对理解消费者行为的产生路径提供了理论模型支撑。

然而,虽然两者之间有着高度的关联性,当前模型驱动的消费者行为研究和数据驱动的推荐系统方法研究彼此之间是相对独立的,尚未形成有效的协同,理论模型的发现尚未有力地支撑技术创新。陈国青等^[7]指出将数据驱动和模型驱动相结合,能够形成新型的大数据驱动研究范式,为理论应用和方法创新带来有效的补充扩展。为此,本研究遵循大数据驱动研究范式,探究如何将消费者决策旅程的理论运用到推荐系统的方法创新上。具体来说,聚焦于基于评论的推荐场景,引入消费者决策旅程中的AIDMA阶段式决策模型^[14, 15],在深度学习的框架下提出了一种新的推荐方法,将其命名为DeepReview,其核心思想与主要创新具体如下:

1) 现有的基于评论的推荐方法研究通常仅考虑对消费者亲自撰写的评论文本或社交关系的好友评论进行建模,而忽略了其在信息不确定情况下对评论文本的潜在阅读行为对购买决策的影响。因此,这类方法不仅会因为撰写评论数据的极

度稀疏导致模型训练的困难,而且无法为那些不写评论的绝大多数群体提供推荐。而现有研究^[16]指出,虽然消费者可能不会主动撰写评论,但他们通常会在购买之前阅读他人撰写的评论,并隐式地影响最终的决策。这意味着,消费者的决策过程中往往隐藏着对商品评论的潜在阅读行为。因此,DeepReview着眼于在商品的全量非结构化评论文本中提取消费者的兴趣特征,捕捉消费者对评论文本潜在阅读行为的隐式反馈,并以此对候选商品进行评分预测,以实现基于评论的推荐方法创新。

2) 不同于当前绝大多数的纯算法改进导向的深度推荐方法研究,DeepReview引入了消费者决策旅程中对商品“引起注意→产生兴趣→唤醒欲望→留下记忆→决策行为”(attention→interest→desire→memory→action)的AIDMA阶段式决策模型^[15],将深度学习的模型设计模拟为消费者对评论文本的信息加工过程。整个模型包含五个部分:attention阶段利用GloVe词嵌入方法将非结构化的评论文本映射到统一的潜在空间中,以实现消费者对商品偏好和商品的表征;interest阶段引入卷积神经网络(convolutional neural network, CNN)和注意力(attention)机制,模拟消费者接收到外部信息(阅读商品评论文本)刺激后所产生的即时偏好改变;desire阶段通过自注意力(self-attention)机制,模拟消费者在被唤醒购买欲望后对评论文本的深层语义理解;memory阶段引入协同注意力(co-attention)机制,模拟阅读评论文本后在消费者心中的记忆匹配和记忆更新;最后,action阶段利用上述四个阶段加工后的评论文本信息学习评分预测模型,以模拟消费者的购买决策行为。五个阶段环环相扣,作用明确,功能清晰,以实现“模型驱动+数据驱动”的深度学习方法创新。

在DeepReview中,消费者的每一个决策行为(对商品的每一次评分)被视为一次决策旅程,每个行为背后的产生路径被划分成AIDMA五个阶段,不仅有助于充分理解消费者对评论文本信息的处理加工过程,也使得深度推荐方法的创新不再是“盲人摸象”的试错法,而是具有管理理论支撑的新探索。同时,由于DeepReview的五个组成部分分别对应了消费者决策旅程的AIDMA五个阶段,因此,模型在微观层面上也被赋予了一定的

可解释性. 这是本研究与其他基于评论的推荐方法研究的重要区别.

为了验证研究思路的可行性和研究方法的有效性, 本研究使用了多个亚马逊的公开基准数据集开展实验研究. 实验结果表明, DeepReview 可以从评论文本信息中很好地提取消费者的兴趣偏好, 推荐效果显著地优于其他基于评论的深度推荐方法, 这证明了本研究所提出的研究思路和研究方法是可行的. 同时, 由于 DeepReview 将对评论文本的信息加工分解为五个阶段, 因此消融实验的开展变得十分直观, 在实验过程中, 本研究对 interest、desire 和 memory 三个阶段^②开展了详实的消融实验和可视化工作, 研究表明 DeepReview 能够很好地利用深度学习技术手段模拟消费者对评论文本的信息加工过程: 1) interest 阶段能够成功地模拟外部评论文本对消费者的刺激作用; 2) desire 阶段能够很好地学习到商品的高维有序特征, 以模拟消费者对评论文本的深层语义理解; 3) memory 阶段能够有效地实现消费者与商品特征的深层次交互过程, 并显著地改变了消费者和商品的特征分布; 4) 每一个信息加工阶段均对决策结果产生显著的影响, 但是在不同数据集中, 各个阶段的效用会有所不同. 本研究也对 DeepReview 中部分重要的超参数开展了敏感性分析, 以得到一些有意义的启示: 1) 通过改变消费者历史商品集的输入数量来模拟消费者对历史决策信息的利用程度, 研究发现越多的历史信息并不总会对消费者决策带来提升, 在不同的数据集中, 历史商品集的输入数量对推荐效果分别呈现出倒 U 或倒 N 的影响; 2) 而通过对评论文本的长度进行控制发现, 并非越长的评论 (越大的信息量) 会帮助消费者更好地作出决策, 两者之间呈现了倒 N 的关系. 上述结论, 为理解消费者在决策过程中对评论文本的信息加工过程提供了有益的借鉴和参考.

1 文献综述

1.1 深度推荐方法研究

随着电子商务的蓬勃发展, 大量推荐系统方

法涌现, 该领域的研究主要经历了几个阶段的发展. 早期研究主要聚焦于基于内容的特征匹配和基于矩阵分解的协同过滤^[17, 18]. 矩阵分解类方法计算高效, 但只对“用户-商品”(user-item)评分数据进行拟合, 许多有用的辅助信息均未考虑, 因此更为通用的做法是将推荐转化为回归问题, 并利用逻辑回归、集成学习等进行特征交叉建模, 如因子分解机^[19]和 LR + GBDT^[20]等; 或者将推荐转化为排序问题, 通过隐式反馈信息实现兴趣的偏序建模^[21]. 然而, 经典算法的局限导致其难以对非结构化数据进行高效处理, 而拥有海量的非结构化数据恰恰是推荐系统的重要特性之一.

由于深度学习能够针对海量训练数据很好地学习极其稀疏的原始特征表征, 以刻画数据的丰富内在信息, 为传统的推荐方法带来新的解决思路, 因此, 当前的推荐方法研究已全面进入深度学习阶段^[2]. 深度推荐方法的研究主要包含两个方面: 1) 利用深度学习对非结构化数据进行特征学习, 并将其作为辅助信息应用到传统矩阵分解中^[22-24]; 2) 直接利用深度学习构建推荐模型, 开创性的工作来自于 YouTube 推荐^[25], 验证了基于大规模数据深度学习能够很好地学习到用户的隐含特征表示, 从而对用户和商品之间的非线性交互进行精确建模, He 等^[4]提出 NCF 算法, 通过深度学习的特征学习代替传统矩阵分解的内积计算, Guo 等^[26]将因子分解机引入 Wide&Deep 深度学习框架提出 DeepFM, 学习更高维的特征交叉关系. 许多其他的深度学习模型也已成功应用到了推荐方法中, 如自编码^[27]、卷积神经网络^[2]、注意力机制^[28]、图神经网络^[29]、多任务学习^[30]和多模态学习^[31]等.

总体而言, 深度学习不仅提升了特征学习的效果, 也使得非结构化数据的分析和挖掘变得可行. 研究表明引入更多的辅助信息 (无论是用户侧还是商品侧), 特别是那些传统意义上难以测量的非结构化数据, 能够为推荐效果的提升带来显著的提升. 理由也十分直观, 因为消费者往往会在脑海中对这些看到的非结构化信息进行处理和

^② attention 阶段的主要作用是对评论文本进行向量表征学习, 而 action 阶段是评分预测的学习, 可以理解为 DeepReview 模型的输入 (input) 和输出 (output), 不存在消融实验的意义, 因此该部分实验中未考虑这两个阶段.

加工,以辅助决策行为。但是,现有研究更多的是从算法技术的视角开展,停留在对表象的数据分析建模层面,无法解析决策过程中对辅助信息的加工过程,模型的设计普遍不具有可解释性。

1.2 基于评论的推荐方法研究

由于评论文本中蕴含了丰富的信息,对辅助消费者决策起到了重要的作用,因此在众多深度推荐方法研究中,基于评论的推荐近年来获得了越来越多的关注。HFT模型^[32]提出了一种从语义角度构建评论特征的方法,将评论特征映射到潜在的主题空间以获取消费者兴趣的高维表示。DeepCoNN^[11]通过CNN获取消费者评论中上下文相关的局部兴趣特征。D-Attn^[33]使用双注意力机制学习评论文本中的局部商品特征和全局语义特征。此外,CARL^[34]扩展了DeepCoNN以学习商品表示的上下文感知特性。ANR^[35]考虑了词的多义性,增加了一个潜在语义矩阵,以更好地区分消费者和商品的潜在语义偏好。SIFN模型^[36]包括情感学习和评分学习两部分,强调评论的情感极性,重点建模评论和用户或商品之间的特征交互。DAML^[37]在CARL的基础上进一步使用注意力机制学习消费者与商品的交互特征。此外,从评论信息有用性的视角,也涌现了诸多有价值的研究。Wang等^[38]认为评论属性与用户偏好相关,并提出了基于评论属性的推荐模型。Chen等^[39]发现,将评论信息对消费者的有用性进行建模能够辅助推荐方法分析消费者的购买意图。Guan等^[1]提出了多模态推荐模型,利用消费者认知风格,得到消费者对评论等多模态数据的差异化偏好,并指出评论信息对消费者的影响更强。Shuai等^[29]将评论语义用于组成图的边缘特征,构建了具有评论特征增强边的评论感知用户-商品图。总体而言,现有研究主要从特征工程的角度出发,利用深度学习的从消费者的评论文本中获取商品特征以提升推荐效果。

而大量客观事实观察和研究表明,相比于写评论的行为,大部分的消费者更倾向于阅读他人的评论以获取信息辅助购买决策^[16]。利用他人的评论信息可以提升消费者的信息感知能力,提升购买决策行为的发生^[40]。消费者在缺乏对商品的了解时会去评论信息中寻找商品的特征以匹配兴趣^[41],也会为了验证购买而阅读评论信息以辅助

自己进行效用最大化的决策^[42]。消费者的社交信息也被用于获取额外辅助信息以降低数据稀疏性^[43]。SBPR模型^[44]通过建立朋友反馈集和添加偏好集以提升推荐效果,G_SBPR模型^[45]则进一步扩展至朋友与群体信息。然而,当前基于评论的推荐方法主要考虑对消费者亲自撰写的评论文本进行建模,或者通过社交关系额外引入好友评论,忽略了在信息不确定情况下对他人评论文本的潜在阅读行为对购买决策的影响,导致无法为那些只看评论、不写评论的“沉默的大多数”提供推荐服务。

1.3 消费者决策旅程研究

消费者决策旅程^[13,14]是消费者行为研究中近年涌现出的一个重要方向,它将消费者行为概念化为动态的决策旅程,以助于在结构上更深入地理解消费者的决策过程和行为产生路径。在相关研究中,外部信息的刺激如何影响消费者的决策旅程,受到了众多关注。Hovland和Lumsdaine^[46]指出个人受到外部环境传递的信息影响时需要经过注意、理解和接受的过程。Colicev等^[47]将消费者的决策旅程划分为外部感知、购买意愿和满意度量三个阶段。而Kim等^[48]在最近的一项工作中将消费者在电子商务平台中的购买行为分解为学习、感知和行为三个阶段,并探明各决策阶段之间存在顺序的关联性。最新的研究中^[49],将消费者决策旅程引入公司数字营销战略之中,突显在数字化社会中研究消费者决策旅程的必要性。

在众多消费者决策旅程研究中,效果层次模型^[50]是探究消费者如何在决策过程中处理和利用信息的一类重要方法。此类模型起源于认知和社会心理学,认为消费者会以顺序的形式处理和响应信息,并将其分解为层次阶段。其中最为重要的模型就是AIDA阶段式决策模型^[51],将消费者的决策旅程划分为“引起注意→产生兴趣→唤醒欲望→决策行为”(attention→interest→desire→action)四个顺序阶段。AIDA模型在帮助企业理解消费者购买行为和制定营销策略起到了重要的作用^[52]。此后,大量的拓展工作围绕该模型展开,其中一项重要的工作即Devoe^[15]提出的AIDMA模型,在AIDA的基础上加入一个新的阶段“留下记忆”(memory),以表征消费者利用自身记忆与潜

在商品进行“货比三家”的比较行为,进一步完善了消费者在购买决策前的心理过程.该模型尤其可以解释商品的广告文稿等外部信息刺激,需要能够首先做到引起注意,然后激发消费者的兴趣,进一步刺激消费者的购买欲望,在消费者心中留下记忆,最终促成消费者的购买行为.该模型已被证明可以有效解释消费者在购买决策中对信息的处理加工过程^[53],并应用于在线广告^[54]、品牌营销^[55]、旅游决策^[56]和在线搜索^[57]等诸多场景中^[58-62].

从本质上来说,无论是推荐系统还是消费者决策旅程,两者研究的核心都是消费者行为,旨在通过不同的方式(数据驱动或模型驱动)探究消费者是如何做出购买决策行为的.消费者决策旅程研究为理解消费者决策过程中对信息的处理加工过程提供了重要的理论模型支撑,但是却尚未有效地运用到推荐系统的方法探索中,理论模型的发现未能有力地支撑技术创新.这也是促使本研究开展的根本出发点.

2 问题描述

本研究选取评分预测(rating prediction)任务构建基于评论的推荐方法研究场景,即利用商品的评论文本(输入)对消费者^③的兴趣偏好进行建模,以实现商品评分(输出)的预测.具体问题描述如下.

设 $u \in \mathbb{U}$ 和 $v \in \mathbb{V}$ 分别代表推荐系统中的—个消费者和商品,其中 \mathbb{U} 为所有消费者的集合, \mathbb{V} 为所有商品的集合. $r_{u,v}$ 表示消费者 u 对商品 v 的评分,而 $R_{u,v}$ 表示对应的评论文本.在消费者与商品的交互数据集 \mathbb{D} 中,每个样本 $\mathbf{D} \in \mathbb{D}$ 可以用一个四元组表示,即 $\mathbf{D} = (u, v, R_{u,v}, r_{u,v})$,记录一个消费者对一个商品的一次评论和评分数据.本研究旨在构建一个方法模型 $\mathbf{h}(u, v | \mathbf{W})$ 以预测 $r_{u,v}$,其中 \mathbf{W} 为模型参数,其目标最优解 \mathbf{W}^* 为

$$\mathbf{W}^* = \operatorname{argmin}_{\mathbf{W}} \sum_{u \in \mathbb{U}, v \in \mathbb{V}} (\mathbf{h}(u, v | \mathbf{W}) - r_{u,v})^2 \quad (1)$$

在基于评论的推荐方法研究中,消费者 u 和

商品 v 的表征是通过非结构化评论文本 $R_{u,v}$ 构建而成,因此构建 $\mathbf{h}(u, v | \mathbf{W})$ 的核心是对 $R_{u,v}$ 的处理.但是,不同于现有方法^[11, 32-35, 37]假设所有消费者都会写评论(即 $R_{u,v} \neq \emptyset$),本研究用一个更为宽松且更符合实际的假设取而代之,即消费者可能不会主动撰写评论,但消费者通常会在购买之前阅读他人撰写的评论,并隐式地影响最终的决策行为.因此,本研究允许 $R_{u,v} = \emptyset$ 的存在,故而全量的数据集 \mathbb{D} 均会被使用,那些没有写评论的消费者也能够被服务.

3 DeepReview 推荐方法

3.1 模型思想与方法框架

针对上述问题,在深度学习框架下,本研究提出一种新的推荐方法 DeepReview,其模型思想为:将深度学习的方法设计模拟为消费者决策旅程中对评论文本的信息加工过程.引入 AIDMA 阶段式决策模型,将该过程划分为五个顺序阶段

1) attention 阶段:消费者感知到商品的存在并引起注意;

2) interest 阶段:消费者对商品产生兴趣,阅读评论文本,并在心中留下印象;

3) desire 阶段:消费者购买欲望被唤醒,对评论文本进行深层语义理解,更深入了解商品;

4) memory 阶段:消费者将对商品的深层认知与自身记忆进行匹配;

5) action 阶段:经过前面四个阶段的信息加工,作出此次旅程的最终决策行为.

DeepReview 的方法框架是在深度学习中将上述五个顺序阶段通过算法设计的形式进行模拟,以实现“输入评论文本→输出评分预测”的信息加工过程.图1具体展示了 DeepReview 方法框架及其与消费者决策旅程的逻辑关系.方法层面具体包含五个模块:首尾 attention 和 action 两个模块分别对应整个深度学习方法的输入和输出,中间 interest、desire 和 memory 三个模块则模拟不同阶段的信息加工方式.

^③ 在推荐系统研究中更常用的称呼是“用户”,虽然营销学中它与“消费者”在概念上稍有差异,但是为了保证理论和方法的表述统一性,在本研究场景中两者等同,研究结论不受上述假设影响.

针对不同阶段的信息加工方式模拟,本研究在 interest、desire 和 memory 三个模块中分别采用了注意力机制、自注意力机制以及协同注意力机制三种方法进行实现. 注意力机制最早被广泛应用于视觉图像领域^[63],通过模拟人类视觉所特有的大脑信息处理机制,快速扫描全局获得需要重点关注的目标信息,并抑制其他不相关的信息. 后续研究表明,注意力机制能够有效应用于许多应用领域,如自然语言处理和推荐系统^[11]. 此外,通过对注意力权重的可视化,还可以解释在生成高维特征时更关注低维特征的哪些部分,使得模型一定程度上具有可解释性. 自注意力机制是注意力机制的一种变体,它通过计算文本中单词间的内部相关性,实现句子内部结构的自我学习,最终生成更聚焦和抽象的高维有序特征表示^[64]. 协同注意力机制进一步将单一对象的注意力建模推广到多对象中,实现注意力权重的同步学习^[65].

AIDMA 阶段式模型与 DeepReview 方法框架的逻辑对应关系如下所示

1) attention 与消费者/商品表征: 商品能够被

消费者发现的途径之一是商品的评论信息,因此通过对消费者的历史决策信息和商品的评论信息建模来反映消费者感知商品并引起注意的过程;

2) interest 与注意力机制: 消费者往往存在多样化的兴趣,注意力机制可以计算权重以表示商品对消费者不同兴趣的刺激程度,以此获得消费者的即时偏好改变以模拟其对商品产生兴趣的过程(改变消费者的表征);

3) desire 与自注意力机制: 当消费者产生购买欲望后,会进一步了解更深层次的商品信息,通过自注意力机制能够学习商品评论文本自身的高维有序特征,模拟消费者对评论文本信息的深层语义理解(改变商品的表征);

4) memory 与协同注意力机制: 消费者在决策旅程中会不断将引起购买欲望的商品与自身兴趣进行匹配,并一次更新其兴趣记忆. 协同注意力机制能够实现多对象的注意力建模与同步学习,很好地模拟消费者和商品在交互过程中的记忆匹配和记忆更新过程(同时改变消费者和商品的表征);

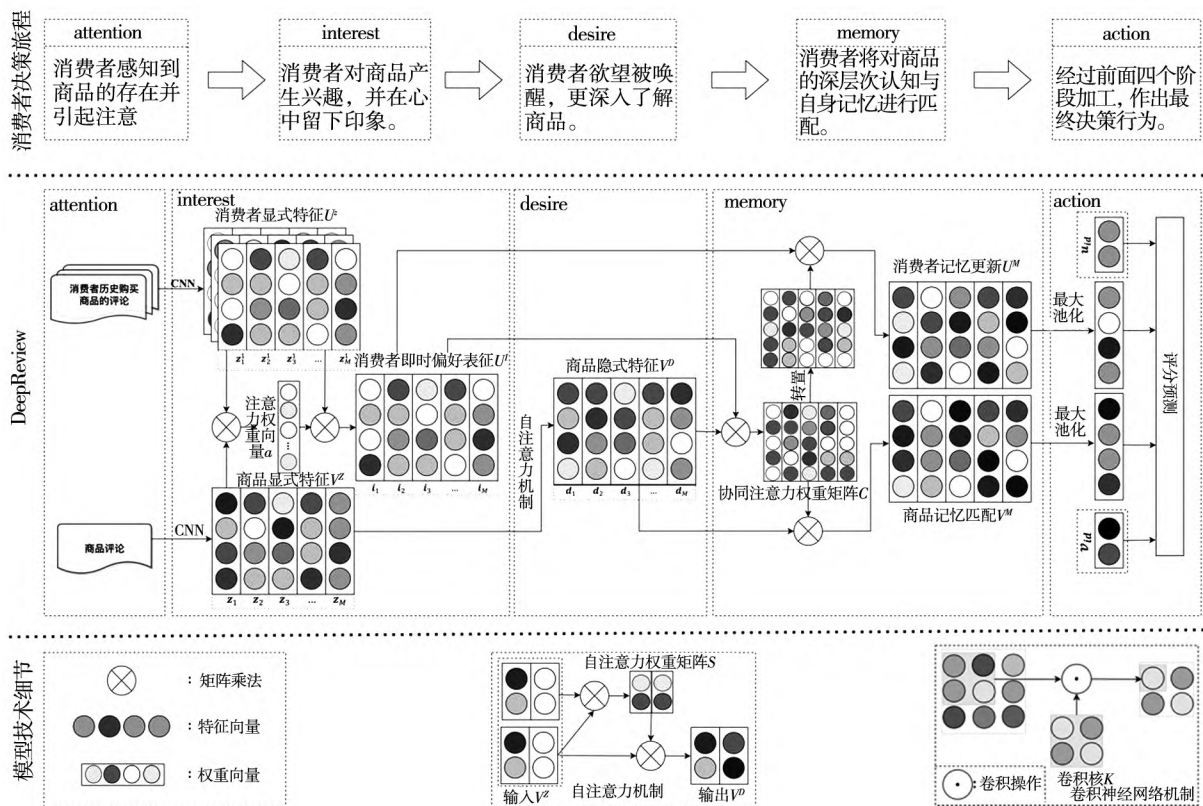


图1 DeepReview 方法框架及其与消费者决策旅程的逻辑关系图

Fig. 1 DeepReview framework and its logical relationship with consumer decision journey

5) action 与评分预测: 决策旅程的最后是决策行为的产生. 消费者对商品的评分能够反映商品购买过程的体验感受, 可以作为一次消费者决策旅程的终点, 因此本研究采用评分预测任务来模拟最终的决策行为的发生.

3.2 attention 模块

attention 模块模拟消费者的商品感知和引起兴趣, 对应 DeepReview 的数据输入阶段. 其核心是通过评论文本构建消费者和商品的统一表征, 即如何利用 $R_{u,v}$ 表征消费者 u 和商品 v .

针对每个商品 v , 可以在交互数据集 \mathbb{D} 中提取其全量的非结构化评论文本, 表示为 \mathbb{D}_v , 具体包含每个消费者 u 对该商品的评论文本 $R_{u,v}$. 本研究使用了 GloVe 方法^[66]对 \mathbb{D}_v 进行词向量的预训练学习, 将非结构化文本转化为向量表征. 具体来说, 通过单词嵌入矩阵 $\mathbf{E} \in \mathbb{R}^{d \times L}$ (d 表示词向量的长度, L 为语料库单词数量, 一般为预设值), 将商品 v 的评论文本表示为下述矩阵

$$\mathbf{V} = [e_1, e_2, \dots, e_T] \quad (2)$$

其中 e_i 表示第 i 个单词的向量; T 为超参数, 表示所使用的商品评论文本长度 (单词数量). 词向量 e_i 可表示为

$$e_i = [e_i^1, e_i^2, \dots, e_i^d] \quad (3)$$

每个商品的表征可以用式 (2) 的矩阵进行表示, 而消费者兴趣偏好是隐藏在其购买过 (评分过) 的商品评论文本中的, 因此可以用消费者历史购买过的商品的矩阵进行表示. 具体方法中, 引入一个新的超参数 N_u , 表示消费者在当前决策过程中会利用到的历史信息 (对历史 N_u 个商品的潜在阅读行为). 消费者 u 的矩阵可表示为下述公式

$$\mathbf{U} = [\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_{N_u}] \quad (4)$$

若 $N_u = 1$, 则表示消费者只利用了他过去评价过的一个商品的评论文本表征其兴趣偏好. N_u 越大表示消费者决策时会利用到越多的历史经验 (模拟消费者的记忆容量). 此外, N_u 个历史商品是按照时间序列的排序进行选择的, 即选择消费者最近的 N_u 个历史商品.

3.3 interest 模块

interest 模块模拟消费者对商品产生兴趣后, 以阅读评论文本的方式接收外部信息刺激, 进而

在心中产生即时偏好改变的过程. 其核心是如何通过 \mathbf{V} 改变 \mathbf{U} .

首先通过卷积神经网络对消费者和商品的矩阵进行显式特征提取. 针对商品 v 的矩阵表征 \mathbf{V} , 其显式特征可表征为矩阵 \mathbf{V}^z

$$\mathbf{V}^z = [z_1, z_2, \dots, z_M] \quad (5)$$

其中每个向量 z_i 通过下述公式计算

$$z_i = \text{ReLU}(\mathbf{V} \odot \mathbf{K}_i + \mathbf{b}_i) \quad (6)$$

M 为卷积核个数, \mathbf{K}_i 表示第 i 个卷积核的参数, \mathbf{b}_i 为对应的偏置项; \odot 表示卷积运算, 激活函数选用了 ReLU 函数. 而针对消费者 u 的矩阵表征 \mathbf{U} , 其显式特征可表征为矩阵 \mathbf{U}^z

$$\mathbf{U}^z = [\mathbf{V}_1^z, \mathbf{V}_2^z, \dots, \mathbf{V}_{N_u}^z] \quad (7)$$

通过上述卷积神经网络操作, 可捕获消费者商品的显式特征, 分别表示为 \mathbf{U}^z 和 \mathbf{V}^z . 在此基础上, 基于注意力机制设计方法以模拟消费者的即时偏好改变过程, 即利用 \mathbf{V}^z 改变 \mathbf{U}^z . 其核心是学习一个注意力权重向量 \mathbf{a}

$$\mathbf{a} = [\alpha_1, \alpha_2, \dots, \alpha_{N_u}] \quad (8)$$

每个 α_i 表示消费者历史购买的第 i 个商品表征 \mathbf{V}_i^z 和当前商品 v 表征 \mathbf{V}^z 的相似度, 其计算如下所示

$$\alpha_i = \text{softmax}(\tanh(\mathbf{W}_u^a \mathbf{V}_i^z) \otimes \tanh(\mathbf{W}_v^a \mathbf{V}^z)) \quad (9)$$

其中 \mathbf{W}_u^a 和 \mathbf{W}_v^a 分别为模型中消费者侧和商品侧需要学习的转换矩阵; \otimes 表示矩阵乘法运算; 激活函数选用了 *softmax* 和 *tanh* 函数. 注意力权重向量 \mathbf{a} 满足下述条件: $\sum_{i=1}^{N_u} \alpha_i = 1$, 在该约束下, 每个注意力权重的取值具有了可以解释的实际意义. 例如, 若通过计算发现 α_i 的取值比向量中的其他值均高, 这意味着商品 v 与消费者 u 历史购买过的第 i 个商品更匹配.

将注意力权重向量 \mathbf{a} 与消费者 u 的显式特征矩阵 \mathbf{U}^z 进行加权求和后, 即可模拟其受到商品 v 刺激后所产生的即使偏好改变, 从而生成新的消费者矩阵, 将其表示为 \mathbf{U}^l , 具体如下所示

$$\mathbf{U}^l = [i_1, i_2, \dots, i_M], \quad (10)$$

其中每个 i_j 通过加权求和方式计算

$$i_j = \sum_{i=1}^{N_u} \alpha_i z_j^i \quad (11)$$

z_j^i 表示消费者历史购买的第 i 个商品的第 j 个显式特征向量,可通过式(6)求得.

3.4 desire 模块

desire 模块模拟消费者若被商品唤醒了购买欲望,会在心中对商品评论文本进行进一步的信息加工,以实现深层语义理解的过程.其核心是如何改变商品显式特征 V^c 以获得更高维有序的隐式特征,在此用 V^D 表示.

此阶段引入自注意力机制以实现深层语义理解过程,旨在学得一个自注意力权重矩阵 S

$$S = [s_1, s_2, \dots, s_M] \quad (12)$$

其中每个向量 s_i 表示商品的第 i 个显式特征(即 z_i)与其他显式特征的相似度,包含 M 个自注意力权重

$$s_i = [\beta_{i,1}, \beta_{i,2}, \dots, \beta_{i,M}] \quad (13)$$

$\beta_{i,j}$ 表示显式特征 z_i 与 z_j 的相似度,计算如下:

$$\beta_{i,j} = \text{softmax}(\tanh(W_q^s z_i) \otimes \tanh(W_k^s z_j)) \quad (14)$$

其中 W_q^s 和 W_k^s 为学习自注意力权重中需要学习的转换矩阵.

将自注意力权重矩阵 S 与商品 v 的显式特征矩阵 V^c 进行矩阵相乘后,即可获得高维有序的隐式特征 V^D ,计算如下

$$V^D = S W_v^s V^c \quad (15)$$

W_v^s 是需要学习的转换矩阵. V^D 可表示为

$$V^D = [d_1, d_2, \dots, d_M] \quad (16)$$

3.5 memory 模块

在唤醒欲望后,消费者会利用对商品的深层语义理解更新自己的记忆.与此同时,商品中那些和消费者记忆匹配的特质会留存下来.因此, memory 模块的核心是模拟消费者和商品的交互过程,通过 U^I 和 V^D 得到记忆更新和记忆匹配后两者的新表征,在此用 U^M 和 V^M 表示.

此交互过程基于协同注意力机制实现.首先,学得一个协同注意力权重矩阵 C

$$C = [c_1, c_2, \dots, c_M] \quad (17)$$

其中每个向量 c_i 表示消费者的第 i 个特征(即 i_i)与商品隐式特征(即 V^D)的相似度,包含 M 个协同注意力权重

$$c_i = [\gamma_{i,1}, \gamma_{i,2}, \dots, \gamma_{i,M}] \quad (18)$$

$\gamma_{i,j}$ 表示消费者的第 i 个特征与商品第 j 个隐式特征的相似度,计算如下

$$\gamma_{i,j} = \tanh(W_u^c i_i) \otimes \tanh(W_v^c d_j) \quad (19)$$

W_u^c 和 W_v^c 为协同注意力模型中需要学习的转换矩阵.

将协同注意力权重矩阵 C 与消费者和商品的特征矩阵(即 U^I 和 V^D)进行矩阵相乘后,即可获得记忆更新和记忆匹配后的表征 U^M 和 V^M ,具体计算如下

$$U^M = \text{softmax}(C^T) \otimes U^I \quad (20)$$

$$V^M = \text{softmax}(C) \otimes V^D \quad (21)$$

3.6 action 模块

通过前面四个阶段对评论文本的信息加工后,可得到消费者和商品的最终矩阵表征 U^M 和 V^M . action 模块旨在利用上述两个矩阵进行评分预测.

在这个过程中,首先引入最大池化(max-pooling)操作,将矩阵转换为向量,以进一步强化消费者和商品的表征,得到消费者和商品的通过评论文本信息加工后的向量表征,用 u_r 和 v_r 表示.其次,借鉴相关研究^[11, 33, 35, 37],将消费者和商品的 ID 信息通过 one-hot 编码引入,用 u_{id} 和 v_{id} 表示.最后,分别将消费者和商品的向量进行拼接,即 u_r 和 u_{id} 拼接得到消费者的最终向量表征 u , v_r 和 v_{id} 拼接得到商品的最终向量表征 v ,构建评分预测函数如下

$$h(u, v | W) = u \cdot v + b + b_u + b_v \quad (22)$$

其中 b, b_u, b_v 为偏置项, W 为 DeepReview 需要学习的所有参数.参数学习使用均方误差来作为目标函数,具体可参考式(1),本研究使用了 Adam 算法^[67]作为学习方法.

4 实验结果与分析

4.1 数据集描述及预处理过程

为了验证 DeepReview 的有效性,根据评论推荐的相关文献调研^[11, 37],本研究选取了 Amazon Review Data^④中 5 个被广泛使用的公开数据集开

④ 公开数据集官方下载链接: <https://nijianmo.github.io/amazon/index.html>.

展实验研究, 分别是 Office Products (OP), Pet Supplies (PS), Video Games (VG), Kindle Store (KS) 和 Amazon Fashion (FS). 在数据新旧程度、数据体量和计算资源等因素的综合考虑下, 本研究选取了公开数据集中最近的三年数据开展实验研究, 即 2016 年—2018 年的数据^⑤.

各数据集的具体统计信息如表 1 所示. 在数据体量视角下, 本研究所选取的 5 个数据集覆盖了 50 万 ~ 300 万量级的评论数量, 以实现中大规模的数据集验证. 由于 Amazon Review Data 是专门应用于评论推荐的基准数据集, 因此绝大部分消费者均有撰写评论的行为, 但是该行为相对稀疏(也符合实际情况), 消费者的平均评论数量在 1 条 ~ 2 条之间. 而在商品视角下, 商品的评论信息则更加丰富, 在 OP、PS 和 VG 数据集中, 每个商品的平均评论数量均超过 10 条. 同时为了验证本研究在不同决策场景中的效果, 本研究也选取了 KS 及 FS 两个评论数量相对稀疏的数据集, 在这两个数据集中, 商品平均评论数量均小于 10 条, 尤其是 FS 数据集中每个商品仅有 3.84 条评论. 不同于传统的评论推荐相关研究^[11, 33, 35, 37]

在数据预处理中普遍选取的 5-core 数据集(即消费者至少要写 5 条评论), 本研究使用了海量且更加稀疏的评论数据, 以更真实模拟模型在实际应用场景中的效果.

在实验中, 本研究对于每一个数据集均按照撰写时间的顺序对评论进行排序预处理, 以模拟实际场景中消费者仅能看到过去的评论信息辅助未来决策, 这也不同于传统相关研究^[11, 33, 35, 37]在模型训练、验证和测试过程中随机选取数据, 不考虑评论的时序性. 传统的数据随机处理方式可能会带来与现实相违背的决策场景, 如预测消费者在 2016 年行为的时候, 使用了 2017 年的未来评论数据, 本研究在实验中规避了这种情况的发生. 因此, 对于每个数据集, 本研究按照时间顺序选取前 80% 的样本作为训练集、之后的 10% 样本作为验证集并将时间最后的 10% 样本作为测试集. 同时, 为了避免头部效应(即少量消费者提供大量评论, 或消费者的一条评论包含过长的文本长度), 数据预处理过程中使用了超参数 N_u 和 T 予以控制, 超参数的具体介绍见第 3.2 节 attention 模块.

表 1 实验数据集的具体统计信息描述(单位: 数量)

Table 1 Statistical information description of the experimental data set (unit: quantity)

数据集(2016 年—2018 年)	消费者	商品	评论	消费者平均评论	商品平均评论
Office Products (OP)	1 645 837	191 538	2 445 502	1.49	12.77
Pet Supplies (PS)	1 614 747	128 239	2 983 238	1.85	23.26
Video Games (VG)	513 472	43 613	734 840	1.43	16.85
Kindle Store (KS)	870 460	248 606	1 747 303	2.01	7.03
Amazon Fashion (FS)	424 732	126 904	487 701	1.15	3.84

4.2 评价指标和对比方法选择

本研究选用评分预测问题中最为常用的均方误差(mean square error, MSE) 作为效果评价指标^[11]. 为了有效验证 DeepReview 的推荐效果, 本研究选取了三类方法进行对比, 最大限度地涵盖了经典的矩阵分解类推荐方法、深度学习类推荐方法和最新的基于评论的推荐方法, 具体如下:

1) SVD: 奇异值分解(singular value decomposition model), 使用矩阵分解方法建模消费者和商品之间的潜在信息^[17];

2) PMF: 概率矩阵分解(probabilistic matrix factorization), 基于高斯分布对消费者和商品的潜在因素进行建模^[68];

3) NCF: 神经协同过滤(neural collaborative filtering), 基于深度学习的协同过滤方法^[4];

4) DeepCoNN: 深度协同神经网络(deep cooperative neural network), 利用消费者写评论的数据构建深度推荐方法^[11];

5) D-Attn: 双重注意力方法(dual attention model), 用全局和局部注意机制改进 DeepCoNN,

⑤ 本论文所使用数据集的下载链接: <https://www.kaggle.com/datasets/deepreview/deepreview>.

以挖掘具有不同信息的单词对评论整体含义的影响^[33];

6) ANR: 基于 aspect 的神经推荐 (aspect based neural recommender), 引入 aspect 概念, 从不同视角提取评论文本特征, 以扩展 DeepCoNN 获取更细粒度的商品表征^[35];

7) DAML: 双重注意交互学习 (dual attention mutual learning), 考虑消费者与商品之间的交互, 并使用欧几里德距离计算消费者特征与商品特征之间的相似性^[37].

方法4) ~ 方法7) 为基于评论的深度推荐方法, 本研究所提出的 DeepReview 与之有两个方面的显著差异: 1) 方法设计原理不同, 本研究模拟了消费者决策旅程中对评论的信息加工过程; 2) 不同于上述四种方法仅利用了消费者亲自撰写的评论进行建模, 本研究同时考虑了消费者对他人评论的潜在阅读行为。

4.3 参数设置

所有的对比方法均使用了最优的参数设置进行训练, 具体设置通过文献参考与实验调优两种方式组合确定, 以保证每个对比方法均在本研究实验中取得各自最优的结果。

在深度学习类方法中, NCF 使用了 3 层感知器计算消费者和商品的 100 维 ID 特征, DeepCoNN 和 DAML 使用了 50 个窗口大小为 3 的卷积核训练, D-Attn 模型使用 200 个局部滤波器和 100 个全局滤波器进行训练, ANR 使用 5 个 aspects 进行训练。同时, 为了公平起见, 在所有的基于评论的推荐方法中, 均截取了评论文本中 T 个单词作为输入, 每个词向量的长度也保持一致 (设为 300)。

表 2 列出了 DeepReview 的具体参数设置。在 attention 模块中, 每个词向量的长度与对比方法保持一致, 设为 $d = 300$; GloVe 语料库选取了 $L = 50\ 000$ 个单词; T 的取值设为 200, 这借鉴了参考文献的设置^[11, 33, 37], 实际上第 4.7.2 节参数敏感性实验表明, 这个取值并非 DeepReview 的最优值, 但是为了保证实验对比的公平性, 本研究针对每个商品选取了 200 个单词作为输入, 与其他对比方法保持一致; 在前四个评论较丰富的数据集中, 消费者平均购买决策数为 8.30, 所以本研究

设置消费者会利用到最多 9 个历史商品的信息, 即 $N_u = 9$, 需要说明的是, N_u 也是一个重要的超参数, 在第 4.7.1 节进行了更详细的敏感型实验。interest 模块与 DeepCoNN 和 DAML 保持一致, 选用了 50 个窗口大小为 3 卷积核, 即 $M = 50$ 。desire 和 memory 模块没有参数需要设置。action 模块的参数设置与对比方法一样, 选择了 3 层感知器及 100 维的消费者和商品 ID 特征; 在参数学习的 Adam 算法中, 学习率设置为 0.05, 训练 10 个 epoch。

表 2 DeepReview 具体参数设置

Table 2 Parameter settings of DeepReview

模块	参数设置
attention	$d = 300, L = 50\ 000, T = 200, N_u = 9$
interest	$M = 50$
desire	——
memory	——
action	$ u_{id} = v_{id} = 100$, 学习率 = 0.05, epoch = 10

同时, 本研究所有实验均执行 10 次后取平均值再进行效果对比, 在每一次实验中, 所有模型参数 (除上述超参数外) 的初始化均采用随机方式产生, 以最大程度降低实验结果的随机性。此外, 为了保证对比实验结果的统计显著性, 本研究也引入 t 检验以验证结果的差异显著性 (包括 95% 和 99% 置信度, 即 $p < 0.05$ 和 $p < 0.01$)。

4.4 对比实验结果分析

在整体推荐效果的实验中, 本研究围绕以下两个问题开展对比分析:

问题 1 和传统基于评分的推荐方法相比, 引入评论文本数据是否会提升推荐效果?

问题 2 将他人撰写的评论用于消费者的偏好建模中, 是否会比仅利用消费者亲自撰写的评论更有效?

如前所述, 本研究选取了三类代表性的推荐方法与 DeepReview 进行对比验证, 分别是经典的矩阵分解类 (SVD 和 PMF)、深度学习类 (NCF) 和基于评论的深度学习类 (DeepCoNN、D-Attn、ANR 和 DAML)。本研究对上述方法在测试集中的效果进行了对比, 表 3 列出了所有方法的对比结果, 以“均方误差 \pm 方差”的形式展示。

表 3 测试集结果对比 (均方误差 ± 方差)

Table 3 Comparison of test set results (mean square error ± variance)

数据集 方法	OP	PS	VG	KS	FS
SVD	2.978 0 ± 0.120 5 **	2.484 4 ± 0.023 6 **	2.516 6 ± 0.068 5 **	2.288 0 ± 0.028 9 **	1.981 5 ± 0.004 9 **
PMF	2.361 2 ± 0.029 6 **	2.126 7 ± 0.018 6 **	1.896 5 ± 0.044 3 **	2.030 5 ± 0.009 1 **	1.990 7 ± 0.004 4 **
NCF	2.150 8 ± 0.007 3 **	1.740 6 ± 0.011 8 **	1.186 9 ± 0.013 2 **	1.901 2 ± 0.010 7 **	1.839 0 ± 0.005 4 **
DeepCoNN	2.123 5 ± 0.007 2 **	1.712 0 ± 0.004 3 **	1.171 8 ± 0.001 9 **	1.833 5 ± 0.004 3 **	1.791 5 ± 0.001 9 **
D-Attn	2.105 7 ± 0.011 9 **	1.705 1 ± 0.009 8 **	1.167 0 ± 0.002 7 **	1.826 7 ± 0.009 5 **	1.783 0 ± 0.013 4 **
ANR	2.160 8 ± 0.093 3 **	1.705 4 ± 0.009 6 **	1.243 8 ± 0.091 5 **	1.884 8 ± 0.015 8 **	1.784 7 ± 0.044 7 **
DAML	2.088 0 ± 0.005 2 **	1.652 6 ± 0.001 2 **	1.071 2 ± 0.001 7 **	1.780 4 ± 0.001 5 **	1.721 8 ± 0.034 5 **
DeepReview	2.065 1 ± 0.007 6	1.588 4 ± 0.015 6	0.985 1 ± 0.021 9	1.757 9 ± 0.008 2	1.656 0 ± 0.022 0

注:1. 表中下划线为对比方法中的最优值;加粗值为本研究方法测算;斜体值表示该评论推荐方法比传统非评论推荐方法效果好。
2. 表中“*”和“**”表示该方法与 DeepReview 的结果对比具有统计显著性,分别代表 95%和 99%的置信度。

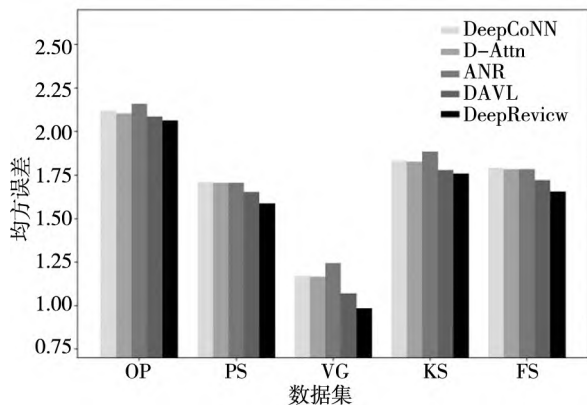
针对第一个问题,实验结果表明,两个经典的矩阵分解推荐方法 SVD 和 PMF 在所有 5 个数据集中的效果均显著地劣于深度学习类方法。而在深度学习类方法中,只有 NCF 未使用评论信息,在其与基于评论的深度推荐方法对比中,可以发现 DeepCoNN、D-Attn、DAML 在所有的数据集中均取得了比 NCF 更好的效果,NCF 仅在 OP 和 VG 两个数据集中比 ANR 效果更好。为了更进一步展示对比方法的综合效果,对各方法在每一个数据集的结果进行排序,具体排序规则为:首先按照均方误差进行排序,再通过 t 检验对相邻的两个方法进行显著性检验,未通过显著性检验 ($p > 0.05$) 的结果将并列排序。具体排序结果如表 4 所示。平均排序结果显示,基于评论的深度推荐方法均比传统的基于评分的推荐方法取得了更好的效

果。该实验结果印证了前面所提出的第一个问题 1,即在绝大多数场景下,引入评论文本数据,是能够有效地提升推荐效果的,此结论也与现有研究结论相吻合^[11, 33, 35, 37]。

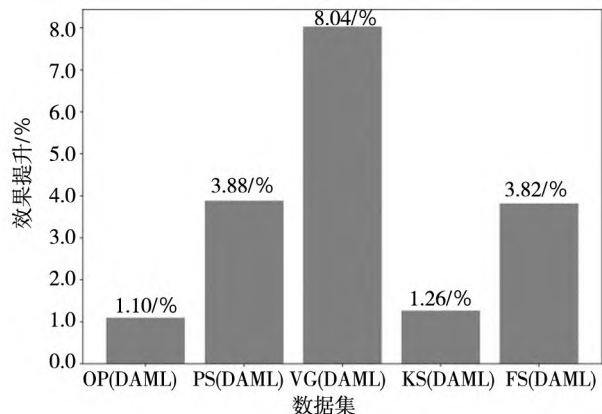
表 4 测试集结果排序对比

Table 4 Ranking of testing results

数据集 方法	OP	PS	VG	KS	FS	平均 排序
SVD	8	8	8	8	7	7.8
PMF	7	7	7	7	8	7.2
NCF	5	6	5	6	6	5.6
DeepCoNN	4	5	4	4	5	4.4
D-Attn	3	3	3	3	3	3
ANR	6	3	6	5	3	4.6
DAML	2	2	2	2	2	2
DeepReview	1	1	1	1	1	1



(a) 基于评论的深度推荐方法结果对比



(b) 与表现最好的方法 (DAML) 对比

图 2 DeepReview 和基于评论的深度推荐方法对比

Fig. 2 Comparison between DeepReview and review-aware recommendation methods

对于第二个问题,本研究对比了 DeepReview 与其他四种基于评论的深度推荐方法,结果如表 3、表 4 和图 2(a) 所示. 实验结果表明,在所有五个数据集中,DeepReview 均取得了最好的推荐效果(排序第 1,且具有 99% 置信度的统计显著性),对比方法中 DAML 和 D-Attn 分列第 2 位和第 3 位. 这表明了消费者的偏好不仅隐藏在自己亲自撰写的评论中,也会受到他人撰写的评论的影响.

图 2(b) 进一步展示了 DeepReview 和最好的对比方法 DAML 相比,其推荐效果的提升比例. 结果表明,相比于每个数据集中表现最好的对比方法 DAML,DeepReview 能够分别达到 1.10%、3.88%、8.04%、1.26% 和 3.82% 的效果提升. 其中,DeepReview 在 PS、VG 和 FS 三个数据集中的效果提升最为显著,在 OP 和 KS 两个数据集中的效果提升相对较低,但是 t 检验结果显示,所有的提升均具有统计意义上的显著性. 效果提升的差异性主要来自于数据集所代表的决策环境差异,在 OP 和 KS 两个数据集中,消费者关注于购买办公用品及电子书籍两类商品. 从商品特性来说,办公用品具有刚需的特性,消费者的决策时效相对较短;电子书籍属于内容型商品,消费者可以通过介绍及预览章节了解书籍内容. 在这两类商品的购买决策中,阅读评论所带来的辅助作用相对是较低的. 因此,在实验结果的对比中可以看到,DeepReview 在 OP 和 KS 两个数据集中的效果提升相对较低. 而在宠物用品 (PS)、电子游戏 (VG)

和时尚品 (FS) 三个购物场景中,消费者往往会更加愿意通过阅读评论来获取额外的辅助决策信息,因此 DeepReview 在这三个数据集中的效果提升更加显著. 此外,从商品评论数据的稀疏性角度分析,实验结果表明,即使在评论数据极度稀疏的 FS 数据集(商品平均仅有 3.84 条评论),DeepReview 也能够很好地使用他人撰写的评论,以显著提升模型的效果.

4.5 消融实验

本节实验围绕第三个问题开展分析:

问题 3 在基于评论的推荐场景中,消费者对评论文本的信息加工,是否遵循了消费者决策旅程中对商品“引起注意→产生兴趣→唤醒欲望→留下记忆→决策行为”的 AIDMA 阶段式过程?

对该问题开展深入的研究,不仅能够检验 DeepReview 方法中各个模块对最终推荐效果所产生的作用,同时也为理解消费者对评论文本的信息加工过程提供了一种可验证的尝试. 为此,本研究设计了消融实验以探索 DeepReview 中各个模块对整体效果的影响. 由于 attention 和 action 两个模块为模型的输入和输出部分,因此仅对 interest, desire 和 memory 三个模块进行消融实验. 具体来说,在 DeepReview 的模型框架下,分别删除了上述三个模块的其中一个,由此设计了三个消融方法:即 ADMA、AIMA 和 AIDA,并选取了整体效果最好的对比方法 DAML 进行比较. 其中,为了便于对比,在本实验中 DeepReview 被命名为 AIDMA. 表 5 和图 3 为消融实验结果的对比.

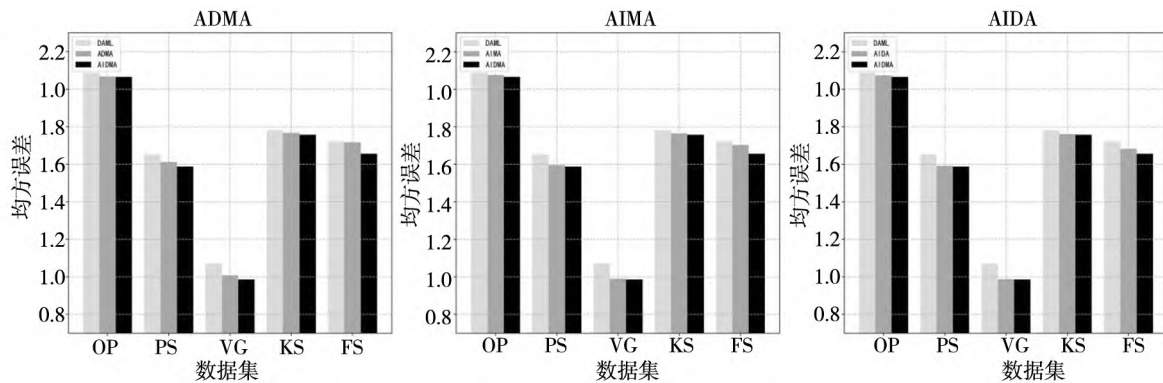


图 3 消融实验结果 (ADMA/AIMA/AIDA 与 DAML 和 AIDMA 的对比)

Fig. 3 Results of ablation experiments (Comparison of ADMA/AIMA/AIDA with DAML and A/DMA)

表 5 消融实验结果对比(均方误差 ± 方差)

Table 5 Results comparison of ablation experiments (mean square error ± variance)

数据集 方法	OP	PS	VG	KS	FS
DAML	2.088 0 ± 0.005 2 **	1.652 6 ± 0.001 2 **	1.071 2 ± 0.001 7 **	1.780 4 ± 0.001 5 **	1.721 8 ± 0.034 5 **
ADMA	2.068 8 ± 0.012 2 **	1.612 2 ± 0.006 6 **	1.009 1 ± 0.004 5 **	1.767 4 ± 0.002 6 **	1.715 8 ± 0.047 2 **
AIMA	2.075 0 ± 0.008 9 **	1.595 0 ± 0.017 3 *	0.989 5 ± 0.007 9 **	1.765 0 ± 0.009 3 **	1.701 6 ± 0.005 6 **
AIDA	2.073 7 ± 0.024 5 *	1.591 0 ± 0.019 1 *	0.988 7 ± 0.030 3 *	1.760 2 ± 0.008 7 *	1.682 4 ± 0.007 8 **
AIDMA	2.065 1 ± 0.007 6	1.588 4 ± 0.015 6	0.985 1 ± 0.021 9	1.757 9 ± 0.008 2	1.656 0 ± 0.022 0

注：1. 表中“*”和“**”表示该方法与 AIDMA 的结果对比具有统计显著性，分别代表 95% 和 99% 的置信度。

2. AIDMA 即 DeepReview 方法。

4.5.1 ADMA 分析: interest 模块的消融

ADMA 假设消费者不经过产生兴趣阶段直接进入唤醒欲望阶段，即当前商品对消费者所有的历史偏好具有同等的刺激效果。在模型设计中，该方法未使用注意力机制模拟 interest 模块，即注意力权重向量 $\alpha = [1/N_u, 1/N_u, \dots, 1/N_u]$ 。

五个数据集的实验结果表明，本研究所设计的 interest 模块发挥了重要的作用，一旦去除效果会显著下降。这说明了当消费者对商品产生兴趣后，阅读评论的行为会潜在刺激消费者的内心世界，并实质上会改变她的即时偏好，进而影响后续的决策行为。与此同时，在与 DAML 的效果对比中发现，即使没有 interest 模块，其余的四个模块依然能够保证本研究所提出的框架能够取得比 DAML 更好的效果。

4.5.2 AIMA 分析: desire 模块的消融

AIMA 去除了自注意力机制提取商品隐式语义特征的过程，即不对评论进行深层语义理解，而直接使用商品浅层的显式特征。

实验结果表明，desire 模块同样起到了重要的作用，一旦去除效果会显著下降。这表明提取商品评论的隐含语义特征是有利于推荐系统效果的提升，消费者一旦被唤醒购买欲望后，会通过对评论的深层语义理解，以获取对商品更深入的了解，而此过程可以通过本研究所设计的自注意力机制模拟实现。

4.5.3 AIDA 分析: memory 模块的消融

AIDA 假设消费者不经过记忆更新和记忆匹

配的过程，从 desire 阶段直接进入最终决策，即在 DeepReview 中去协同注意力机制。

在实验中发现，memory 模块对最终推荐效果同样发挥了正向作用，若去除整体效果会下降。这表明消费者在对评论进行信息加工过程中，会将商品特征与其记忆进行匹配并更新，从而影响最终的购买决策行为。

4.5.4 综合分析

进一步使用第 4.4 节中的排序方法，本研究对各方法的推荐效果进行排序，结果如表 6 所示。对消融实验的综合分析发现，每一个信息加工阶段均对决策结果的产生起到了重要的作用。但是在不同的决策环境下（不同的数据集），各个阶段的效用会有所不同。

表 6 消融实验结果排序对比

Table 6 Ranking of ablation results

数据集 方法	OP	PS	VG	KS	FS	平均 排序
DAML	5	5	5	5	5	5
ADMA	2	4	4	4	4	3.6
AIMA	3	3	2	3	3	2.8
AIDA	3	2	2	2	2	2.2
AIDMA	1	1	1	1	1	1

整体来说，所有的 DeepReview 消融方法，均能够取得比 DAML 更好的推荐效果。而在三个消融方法的比较中发现，AIDA 在整体上会显著地优于 ADMA 和 AIMA，这说明了 interest 和 desire 模块会比 memory 模块更加重要。解释该结果的原因可能

有两个方面:1)在 interest 和 desire 模块的信息加工后,消费者和商品的高维隐式特征已能够被捕获,直接使用它们进行评分预测也能够保障较好的效果;2)memory 模块模拟的是消费者唤醒购买欲望后的进一步信息交互过程,但是在许多场景中,消费者可能在被唤醒欲望后即进入购买决策阶段.这一实验发现也侧面印证了为何消费者决策旅程

研究中会有经典的 AIDA 模型,但却很少提及 AD-MA 和 AIMA 模型.从整体结果来说,AIDMA(也就是本研究提出的 DeepReview)取得了最好的效果,这表明了每一个模块都对最终效果的达成起到了重要的作用,这也印证了消费者在本研究所实验场景中的信息加工过程体现了 AIDMA 消费者决策旅程.

表 7 interest 模块中商品评论对消费者的偏好刺激

Table 7 Stimulation of consumer preference by reviews in the interest module

	消费者历史购买的商品及其评论文本	注意力权重	一件目标商品(短裤)的评论文本
短裤	...非常光滑。跑步用品。他的口袋可以防止你的东西从边缘掉下来...这太完美了!你身材适中的话,它可以使你的大腿感到舒适,这短裤是对 80 年代风格的完美回归...	0.675	...这条裤子有几种设计,有多种口袋可供选择,有多种尺寸和颜色可供选择,可以穿在任何衣服上...这些慢跑裤正是我想要的。我为我得到的东西感到非常自豪。我打算在 7 月 4 日穿上它...
帽子	...这顶帽子远远超出了我的预期!我想要一顶朴素的帽子。不到 6 美元的价格真是无可挑剔。不过,这顶帽子是可调节的,适合两端都戴的人。它的材质柔软轻便...洗起来不太好洗,主要是因为很容易褪色。对比我尝试过的不同品牌,这个帽子总是很可靠,而且有很多颜色...	0.214	
T 恤	...小心购买乐队衬衫。我曾尝试从两家不同的公司购买艾略特-史密斯设计的衬衫(我退了一件,又从另一家公司买了一件),结果发现不同的卖家似乎都在抄袭自己的设计...	0.092	

4.6 可视化实验

4.6.1 interest 模块可视化

为了验证 interest 模块的注意力机制能够模拟商品对消费者的刺激过程,本研究随机从 FS 数据集中提取了一对消费者和商品的相关信息,如表 7 所示.消费者侧展示了三件其历史购买过的商品,分别为短裤、帽子和 T 恤,以及相对应的三个评论文本集合;目标商品侧则展示了一件短裤及对应的评论文本集合.interest 模块旨在将目标商品与消费者的历史经验进行匹配,以刺激消费者的即使偏好改变,具体的刺激程度可以通过注意力权重表示.如表所示,目标短裤对消费历史购买过的短裤、帽子和 T 恤的刺激程度分别为 67.5%、21.4% 和 9.2%.从直观上可以表示,消费者在购买新的短裤时,往往会更多地利用到其过去购买过同类商品的经验;同时,也会利用到跨品类但是具有一定互补特性的商品经验如帽子和

T 恤.上述三个注意力权重的求和为 98.1%,因为

注意力权重需要满足 $\sum_{i=1}^{N_u} \alpha_i = 1$ 条件,这意味着该目标商品对消费者其他的历史购买商品刺激程度仅为 1.8%.这表明了 interest 模块能够很好地模拟商品对消费者历史偏好的刺激过程.

4.6.2 desire 模块可视化

desire 模块的可视化实验中,从 FS 数据集中随机抽取了一个商品(u20829,i11515),使用 t-SNE 算法将学习到的商品显式特征 V^r 、自注意力的权重矩阵 S 和商品的高维隐式特征 V^p 投影到二维空间.如图 4 所示,商品的高维隐式特征相比于其低维显式特征分布更为有序和集中,而自注意力权重矩阵呈现集群式分布,这表明了 desire 模块能够提取到商品信息中最有用的信息特征.其他的消费者和商品特征可视化结果,同样呈现了上述结果.

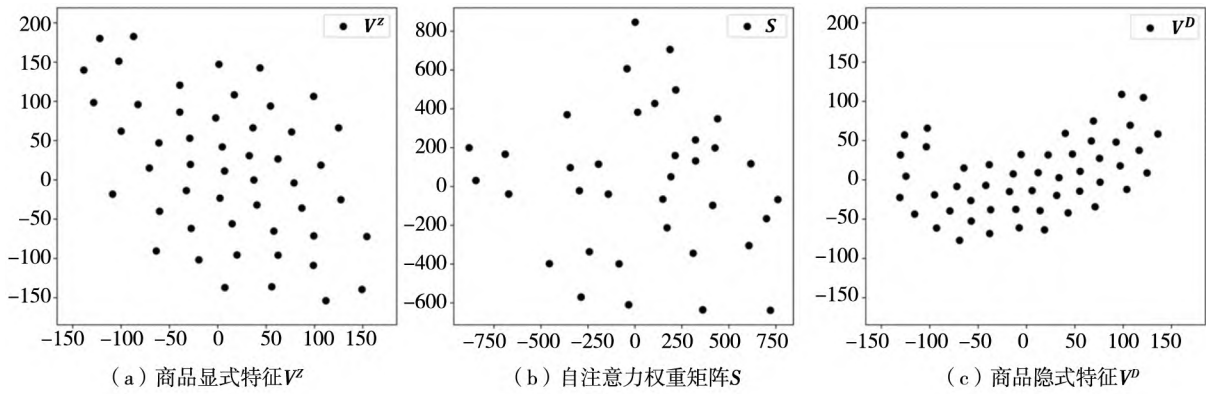


图 4 desire 模块可视化 (u20829, i11515)

Fig. 4 Visualization of desire module (u20829, i11515)

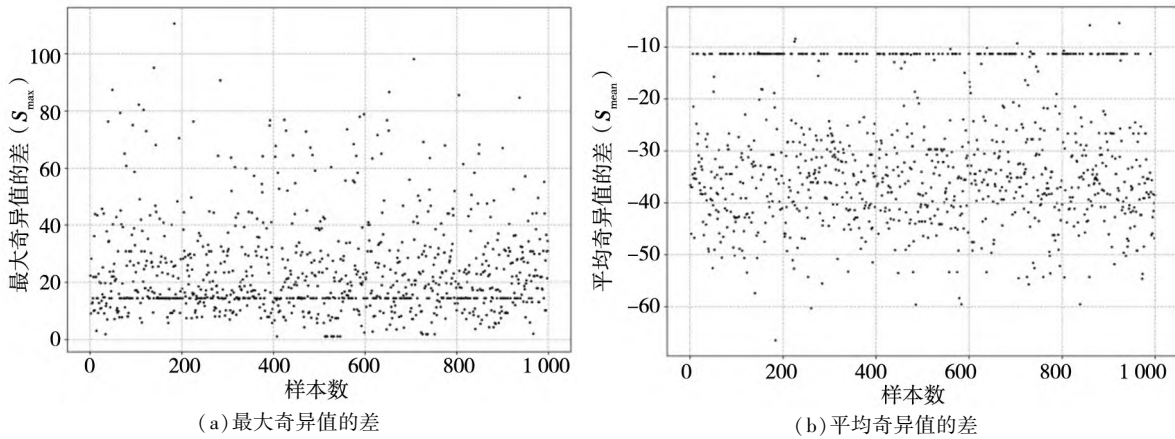


图 5 desire 模块的奇异值差异可视化

Fig. 5 Visualization of singular value difference in desire module

为了进一步验证 desire 模块中商品的高维隐式特征 V^D 相比低维显式特征 V^z 能够表征更为密集有序的信息,本研究引入奇异值分解 (singular value decomposition) 以比较矩阵间的信息差异,如式(23)和式(24)所示,其中 S^D 和 S^z 分别表示两个矩阵的奇异值.

$$V^D = P^D S^D Q^D \tag{23}$$

$$V^z = P^z S^z Q^z \tag{24}$$

奇异值对应着矩阵中隐含的重要信息,且重要性和奇异值大小正相关^[69].因此,若 V^D 的最大奇异值大于 V^z 的最大奇异值,则表明矩阵学得更高维的重要信息.若去除最大值后的 V^D 的平均奇异值小于 V^z 所对应的值,代表矩阵学得更密集的有序信息.最大奇异值的差值 S_{Max} 及平均奇异值的差值 S_{Mean} 可以通过下述两个公式计算

$$S_{Max} = \text{Max}(S^D) - \text{Max}(S^z) \tag{25}$$

$$S_{Mean} = \frac{\text{Sum}(S^D) - \text{Max}(S^D)}{\text{Count}(S^D) - 1} - \frac{\text{Sum}(S^z) - \text{Max}(S^z)}{\text{Count}(S^z) - 1} \tag{26}$$

为此,从 FS 数据集的测试集中随机抽取 1 000 条样本,计算样本中的 S_{Max} 和 S_{Mean} ,结果如图 5 所示.可以看出相比于 V^z , V^D 能够得到更大的最大奇异值 ($S_{Max} > 0$) 和更小的平均奇异值 ($S_{Mean} < 0$),这表明了 desire 模块能够帮助用户获取更加高维的有序特征.

4.6.3 memory 模块可视化

memory 模块可视化中,随机抽取了 FS 数据集的四对消费者和商品,分别为:消费者 u71476 及其购买过的两个商品 i11115 和 i17902、商品 i11515 及购买过它的两位消费者 u34048 和 u20829.使用 t-SNE 算法将学习到的消费者即时

偏好矩阵 U^I 、商品高维隐式特征 V^D 、协同注意力权重矩阵 C 、消费者的更新记忆 U^M 和商品匹配记忆 V^M 投影到二维空间,如图 6 和图 7 所示. 可以发现,协同注意力权重矩阵的分布较为集中,并且消费者和商品的特征分布经过协同交互后,分布方向均发生了显著变化. 如图 6 所示,对于同一个消费者 $u71476$,不同商品与之交互的权重矩阵

C 分布位置不同,使得消费者特征分布转变方向也不同,这表明了 memory 模块通过将商品文本信息引入到用户兴趣特征后,更新了消费者的记忆存储. 而如图 7 所示,对于同一个商品 $i11515$,不同消费者与之交互的权重矩阵分布位置亦不同,商品的特征分布改变也不同,表明 memory 模块同样成功地模拟了不同消费者记忆中的偏好差异.

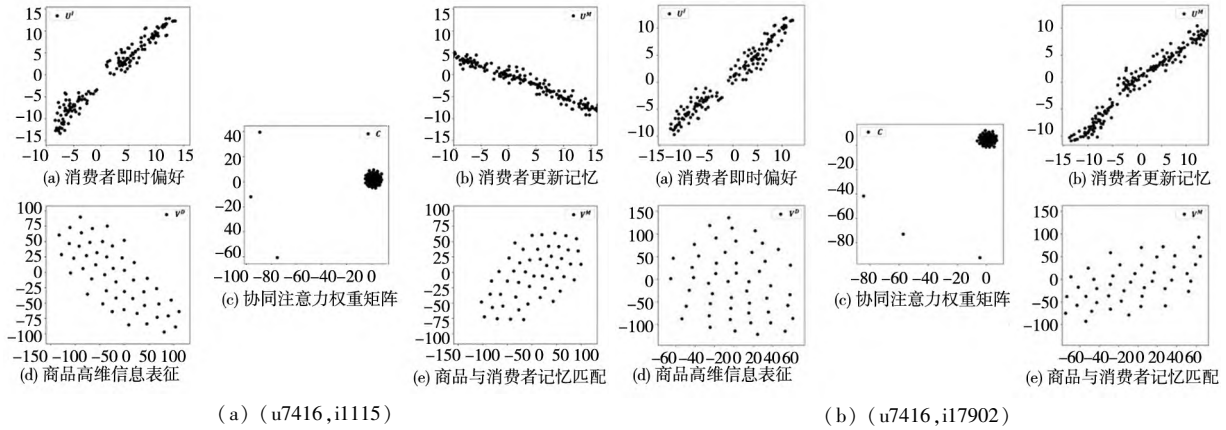


图 6 memory 模块可视化(消费者 $u71476$ 示例)

Fig. 6 Visualization of memory module (consumer $u71476$ example)

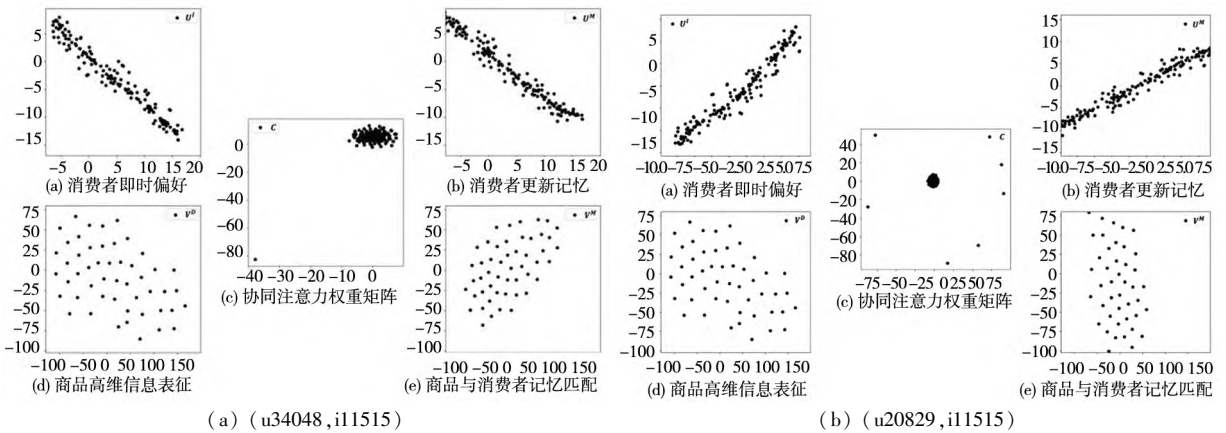


图 7 memory 模块可视化(商品 $i11515$ 示例)

Fig. 7 Visualization of memory module (item $i11515$ example)

为了进一步验证 memory 模块中用户侧学得的 U^M 相比 U^I 、以及商品侧的 V^M 相比 V^D 分布方向发生了转变,本研究引入 KL 散度 (Kullback-Leibler divergence)^[70] 指标以计算分布之间的差异,如式(27)和式(28)所示,其中 KL_u 和 KL_v 分别表示消费者侧和商品侧的 KL 散度. KL 散度值越大,表示分布间的差异越大.

$$KL_u = \frac{\sum U^M \log U^M}{U^I} \quad (27)$$

$$KL_v = \frac{\sum V^M \log V^M}{V^I} \quad (28)$$

为此,从 FS 数据集的测试集中随机抽取 1 000 条样本,计算消费者与商品侧的分布差异. 如图 8 所示,无论是消费者侧还是商品侧的特征分布,经过协同交互后,均发生了显著的变化 (KL 散度均 > 0). 独立分析可以发现,消费者侧的分布发生了较大的变化,而商品侧的分布变化相对较小,其原因主要在于:消费者的偏好往往更加多元,记

忆的更新往往更加多样,因此表现出来的 KL 散度变化更大;而商品有特定的固有特征,记忆匹配的特征往往更加集中,由此而展现的 KL 散度变

化就会更小.综上所述,实验表明了 $memory$ 模块能够成功地挖掘消费者内心深处对特定商品的记忆表征,以及商品中与特定消费者相关的属性表征.

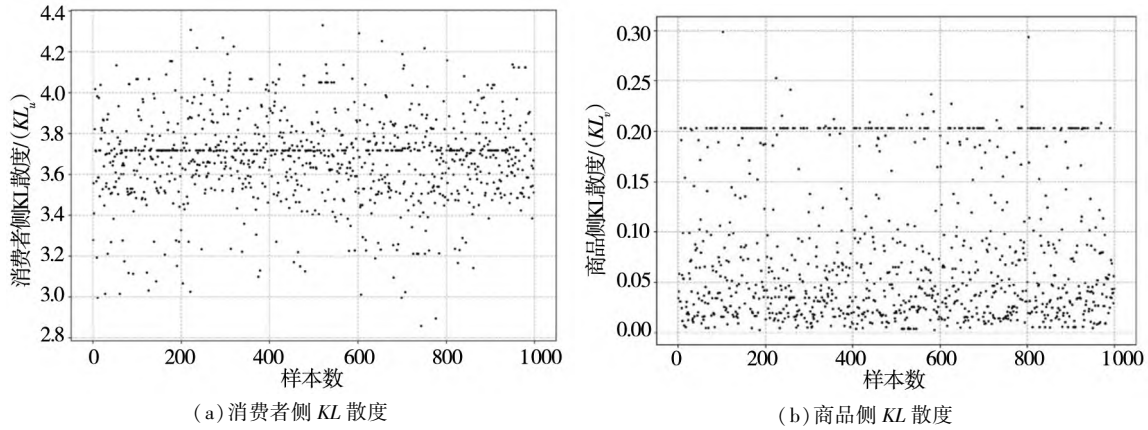


图 8 memory 阶段 KL 散度可视化

Fig. 8 Visualization of KL divergence in memory module

4.7 参数敏感性分析

4.7.1 超参数 N_u

如第 3.2 节所述,超参数 N_u 控制着消费者在每次决策过程中会利用到的历史商品信息数量,本质上即模拟了消费者的记忆容量.在前述实验中, N_u 统一设置为一个较大的数值 9(相对本研究数据集而言).在本节实验中,通过设置不同的 N_u 来改变消费者历史商品集的输入数量, N_u 的取值选取了 [1, 2, 3, 4, 5, 6, 7, 8, 9] 共 9 组实验.

实验结果如表 8 和图 9 所示.实验结果发现在 OP、VG 和 KS 数据集中该参数对消费者购买决策的影响呈现倒 U 关系,在 PS 和 FS 数据集上呈现倒 N 关系.这表明了越多的历史信息并不总会消费者对决策带来提升.但是普遍现象是,当历

史信息足够丰富的时候,对于建模消费者的兴趣偏好往往会更准确.

4.7.2 超参数 T

超参数 T 表示在构建消费者和商品矩阵时,所使用到的商品评论文本长度,即单词数量,模拟了评论文本的信息量.在前述实验中, T 均取值为 200.本节实验将 T 分别取值为 [200, 250, 300, 350, 400],本研究按照评论时间的顺序关系选取距离当前时间点最近的 T 个词.在 FS 数据集中开展敏感性实验,以验证消费者在处理不同信息量的场景下 DeepReview 的推荐效果.实验结果如图 10 所示,可以看出在 T 为 300 时,DeepReview 表现出最好的效果.从整体来看,并非越长的评论(越大的信息量)会帮助消费者更好地作出决策,两者之间呈现了倒 N 的关系.

表 8 超参数 N_u 不同取值下的结果对比

Table 8 Comparison under different values of N_u

N_u 数据集	1	2	3	4	5	6	7	8	9
OP	2.075 1	2.080 1	2.072 0	2.071 1	2.066 0	2.065 2	2.066 1	2.061 2	2.065 1
PS	1.619 2	1.606 1	1.611 0	1.608 2	1.611 1	1.588 7	1.584 2	1.584 9	1.588 4
VG	1.001 1	1.010 1	1.010 3	0.997 0	0.995 2	0.993 0	0.985 9	0.987 7	0.985 1
KS	1.761 2	1.767 1	1.772 3	1.770 9	1.772 9	1.763 3	1.762 1	1.755 2	1.757 9
FS	1.697 3	1.661 0	1.647 1	1.628 1	1.655 1	1.700 2	1.726 0	1.681 1	1.656 0

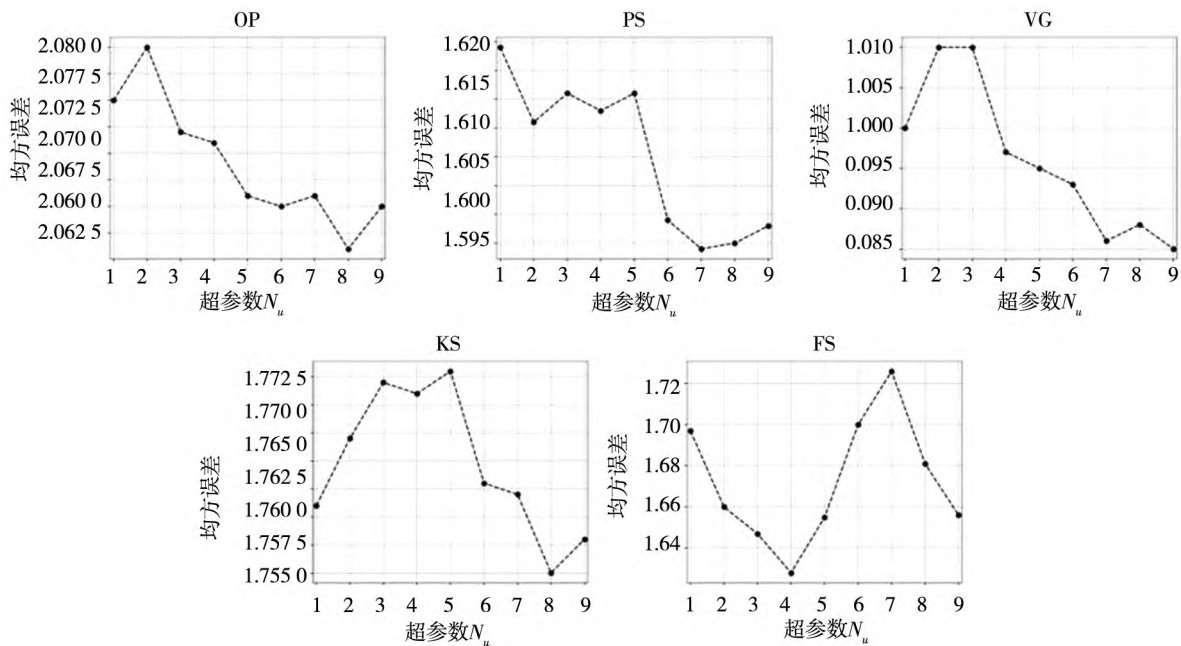


图9 超参数 N_u 的敏感性分析

Fig. 9 Sensitivity analysis of N_u

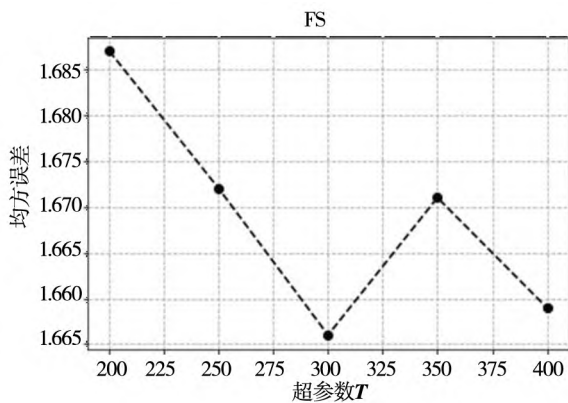


图10 超参数 T 的参数敏感性分析

Fig. 10 Sensitivity analysis of T

4 结束语

本文研究了一种新的推荐系统方法创新思路,提出了一种新的基于评论的推荐方法 Deep-Review. 该方法从消费者决策旅程视角出发,引入 AIDMA 阶段式决策模型,将消费者行为的产生过程视为对商品“引起注意→产生兴趣→唤醒欲望→留下记忆→决策行为”的决策旅程,并将其运用到深度学习的模型设计中,通过注意力机制等方法进行模拟. 不同于当前绝大多数的纯算法改进驱动的推荐方法研究,本研究遵循了“模型驱

动+数据驱动”的大数据驱动研究范式,引入消费者行为理论模型,实现具有理论支撑的方法创新. 在五个基准数据集上的实验结果表明,Deep-Review 可以从评论文本信息中心很好地提取消费者的高维兴趣偏好,推荐效果显著优于对比方法. 这证明了本研究所提出的研究思路和研究方法是可行的. 还验证了消费者对评论文本的信息加工存在 AIDMA 的阶段式过程,并发现了越多的历史信息并不总会对消费者决策带来提升,评论文本蕴含的信息量越大也不总是能够帮助消费者进行更好的决策,这为理解消费者在决策过程中对非结构化评论文本的信息加工过程提供了借鉴和参考.

本研究仅将消费者历史决策经验中的评论文本数据以及 AIDMA 阶段式决策模型进行融合,未来可利用更多模态的数据(如图片、视频等)挖掘消费者的行为偏好,亦可考虑消费者行为和消费者心理领域中的其他理论模型. 此外,本研究在实验中选取了被广泛使用的 Amazon Review Data 公开数据集,该数据集未包含不写评论的消费者数据,针对此类不写评论的消费者群体如何利用他人的评论以影响自身的决策旅程,也将是本文的另一项未来研究工作.

参考文献:

- [1] Guan Y, Wei Q, Chen G. Deep learning based personalized recommendation with multi-view information integration[J]. Decision Support Systems, 2019, 118: 58 – 69.
- [2] Zhang S, Yao L, Sun A, et al. Deep learning based recommender system: A survey and new perspectives[J]. ACM Computing Surveys, 2019, 52(1): 1 – 38.
- [3] Sargin E. Deep Neural Networks for Youtube Recommendations[C]. ACM Recommender Systems Conference, New York: Association for Computing Machinery, 2016.
- [4] He X, Liao L, Zhang H, et al. Neural Collaborative Filtering[C]. ACM International World Wide Web Conference, New York: Association for Computing Machinery, 2017.
- [5] 徐宗本, 冯芷艳, 郭迅华, 等. 大数据驱动的管理与决策前沿课题[J]. 管理世界, 2014, (11): 158 – 163.
Xu Zongben, Feng Zhiyan, Guo Xunhua, et al. Big data-driven management and decision-making frontier topics[J]. Journal of Management World, 2014, (11): 158 – 163. (in Chinese)
- [6] 杨善林, 周开乐. 大数据中的管理问题: 基于大数据的资源观[J]. 管理科学学报, 2015, 18(5): 1 – 8.
Yang Shanlin, Zhou Kaile. Management issues in Big Data: The resource-based view of Big Data[J]. Journal of Management Sciences in China, 2015, 18(5): 1 – 8. (in Chinese)
- [7] 陈国青, 吴刚, 顾远东, 等. 管理决策情境下大数据驱动的研究和应用挑战 —— 范式转变与研究方向[J]. 管理科学学报, 2018, 21(7): 1 – 10.
Chen Guoqing, Wu Gang, Gu Yuandong, et al. The challenges for Big data driven research and applications in the context of managerial decision-making-paradigm shift and research directions[J]. Journal of Management Sciences in China, 2018, 21(7): 1 – 10. (in Chinese)
- [8] 陈国青, 曾大军, 卫强, 等. 大数据环境下的决策范式转变与使能创新[J]. 管理世界, 2020, 36(2): 95 – 105.
Chen Guoqing, Zeng Dajun, Wei Qiang, et al. Transitions of decision-making paradigms and enabled innovations in the context of Big Data[J]. Journal of Management World, 2020, 36(2): 95 – 105. (in Chinese)
- [9] 刘业政, 孙见山, 姜元春, 等. 大数据的价值发现: 4C 模型[J]. 管理世界, 2020, 36(2): 129 – 138.
Liu Yezheng, Sun Jianshan, Jiang Yuanchun, et al. 4C Model: Value discovery in Big Data[J]. Journal of Management World, 2020, 36(2): 129 – 138. (in Chinese)
- [10] Festinger L. A theory of social comparison processes[J]. Human Relations, 1954, 7(2): 117 – 140.
- [11] Zheng L, Noroozi V, Yu P S. Joint Deep Modeling of Users and Items Using Reviews for Recommendation[C]. ACM International Conference on Web Search and Data Mining, New York: Association for Computing Machinery, 2017.
- [12] Hirschman E C. Consumer behavior theories as heroic quest[J]. ACR North American Advances in Consumer Research, 1989, 16: 13639 – 16646.
- [13] Court D, Elzinga D, Mulder S, et al. The consumer decision journey[J]. McKinsey Quarterly, 2009, 3(3): 1 – 11.
- [14] Santos S, Goncalves H M. The consumer decision journey: A literature review of the foundational models and theories and a future perspective[J]. Technological Forecasting and Social Change, 2021, 173: 121117.
- [15] Devoe M. Effective Self-Management in Selling[M]. New Jersey: Prentice-Hall, 1956.
- [16] Mudambi S M, Schuff D. Research note: What makes a helpful online review? A study of customer reviews on Amazon.com[J]. MIS Quarterly, 2010: 185 – 200.
- [17] Koren Y, Bell R, Volinsky C. Matrix factorization techniques for recommender systems[J]. Computer, 2009, 42(8): 30 – 37.
- [18] 朱国玮, 周利. 基于遗忘函数和领域最近邻的混合推荐研究[J]. 管理科学学报, 2012, 15(5): 55 – 64.
Zhu Guowei, Zhou Li. Hybrid recommendation based on forgetting curve and domain nearest neighbor[J]. Journal of Management Sciences in China, 2012, 15(5): 55 – 64. (in Chinese)
- [19] Rendle S. Factorization Machines[C]. IEEE ICDM, Piscataway: IEEE, 2010.
- [20] He X, Pan J, Jin O, et al. Practical Lessons From Predicting Clicks on Ads at Facebook[C]. The Workshop on Data Mining for Online Advertising, New York: Association for Computing Machinery, 2014.
- [21] Rendle S, Freudenthaler C, Gantner Z, et al. BPR: Bayesian Personalized Ranking from Implicit Feedback[C]. The Conference on Uncertainty in Artificial Intelligence, Arlington: AUAI Press, 2009.
- [22] Wang H, Wang N, Yeung D Y. Collaborative Deep Learning for Recommender Systems[C]. ACM SIGKDD Conference on Knowledge Discovery and Data Mining, New York: Association for Computing Machinery, 2015.
- [23] Zhang F, Yuan N J, Lian D, et al. Collaborative Knowledge base Embedding for Recommender Systems[C]. ACM SIGKDD Conference on Knowledge Discovery and Data Mining, New York: Association for Computing Machinery, 2016.
- [24] Li X, Chen H. Recommendation as link prediction in bipartite graphs: A graph kernel-based machine learning approach[J]. Decision Support Systems, 2013, 54(2): 880 – 890.
- [25] Covington P, Adams J, Sargin E. Deep Neural Networks for YouTube Recommendations[C]. ACM Recommender Systems Conference, New York: Association for Computing Machinery, 2016.
- [26] Guo H, Tang R, Ye Y, et al. DeepFM: A Factorization-Machine based Neural Network for CTR Prediction[C]. International Joint Conference on Artificial Intelligence, Melbourne: AAAI Press, 2017: 1725 – 1731.
- [27] Sedhain S, Menon A K, Sanner S, et al. Autorec: Autoencoders Meet Collaborative Filtering[C]. ACM International

- World Wide Web Conference, New York; Association for Computing Machinery, 2015.
- [28] Dong X, Ni J, Cheng W, et al. Asymmetrical Hierarchical Networks with Attentive Interactions for Interpretable Review-based Recommendation [C]. The Association for the Advancement of Artificial Intelligence, Melbourne; AAAI Press, 2020.
- [29] Shuai J, Zhang K, Wu L, et al. A Review-Aware Graph Contrastive Learning Framework for Recommendation[C]. ACM SIGIR Conference on Research and Development in Information Retrieval, New York; Association for Computing Machinery, 2022.
- [30] Sheng X R, Zhao L, Zhou G, et al. One Model to Serve All: Star Topology Adaptive Recommender for Multi-Domain CTR Prediction[C]. ACM International Conference on Information and Knowledge Management, New York; Association for Computing Machinery, 2021.
- [31] Qian Y, Xu W, Liu X, et al. Popularity prediction for marketer-generated content: A text-guided attention neural network for multi-modal feature fusion[J]. Information Processing & Management, 2022, 59(4): 102984.
- [32] McAuley J, Leskovec J. Hidden Factors and Hidden Topics: Understanding Rating Dimensions with Review Text[C]. ACM Recommender Systems Conference, New York; Association for Computing Machinery, 2013.
- [33] Seo S, Huang J, Yang H, et al. Interpretable Convolutional Neural Networks with Dual Local and Global Attention for Review Rating Prediction [C]. ACM Recommender Systems Conference, New York; Association for Computing Machinery, 2017.
- [34] Wu L, Quan C, Li C, et al. A context-aware user-item representation learning for item recommendation[J]. ACM Transactions on Information Systems, 2019, 37(2): 1–29.
- [35] Chin J Y, Zhao K, Joty S, et al. ANR: Aspect-based Neural Recommender[C]. ACM International Conference on Information and Knowledge Management, New York; Association for Computing Machinery, 2018.
- [36] Zhang K, Qian H, Liu Q, et al. SIFN: A Sentiment-aware Interactive Fusion Network for Review-based Item Recommendation[C]. ACM International Conference on Information and Knowledge Management, New York; Association for Computing Machinery, 2021.
- [37] Liu D, Li J, Du B, et al. Daml: Dual Attention Mutual Learning between Ratings and Reviews for Item Recommendation [C]. ACM SIGKDD Conference on Knowledge Discovery and Data Mining, New York; Association for Computing Machinery, 2019.
- [38] Wang X, Ounis I, Macdonald C. Leveraging Review Properties for Effective Recommendation[C]. ACM International World Wide Web Conference, New York; Association for Computing Machinery, 2021.
- [39] Chen C, Zhang M, Liu Y, et al. Neural Attentional Rating Regression with Review-level Explanations[C]. ACM International World Wide Web Conference, New York; Association for Computing Machinery, 2018.
- [40] Herm-Stapelberg N, Rothlauf F. The crowd against the few: Measuring the impact of expert recommendations[J]. Decision Support Systems, 2020, 138: 113345.
- [41] Song H J, Ruan W J, Jeon Y J. An integrated approach to the purchase decision making process of food-delivery apps: Focusing on the TAM and AIDA models[J]. International Journal of Hospitality Management, 2021, 95: 102943.
- [42] He J, Wang X, Vandenbosch M B, et al. Revealed preference in online reviews: Purchase verification in the tablet market [J]. Decision Support Systems, 2020, 132: 113281.
- [43] Symeonidis P. ClustHOSVD: Item recommendation by combining semantically enhanced tag clustering with tensor HOSVD [J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2015, 46(9): 1240–1251.
- [44] Zhao T, McAuley J, King I. Leveraging Social Connections to Improve Personalized Ranking for Collaborative Filtering[C]. ACM International Conference on Information and Knowledge Management, New York; Association for Computing Machinery, 2014.
- [45] Sun J, Ying R, Jiang Y, et al. Leveraging friend and group information to improve social recommender system[J]. Electronic Commerce Research. 2020, 20(1): 147–172.
- [46] Hovland C I, Lumsdaine A A. Experiments on Mass Communication[M]. Princeton: Princeton University Press, 2017.
- [47] Colicev A, Malshe A, Pauwels K, et al. Improving consumer mindset metrics and shareholder value through social media: The different roles of owned and earned media[J]. Journal of Marketing, 2018, 82(1): 37–56.
- [48] Kim H, Jiang J, Bruce N I. Discovering heterogeneous consumer journeys in online platforms: Implications for networking investment[J]. Journal of the Academy of Marketing Science, 2021, 49(2): 374–396.
- [49] Vollrath M D, Villegas S G. Avoiding digital marketing analytics myopia: revisiting the customer decision journey as a strategic marketing framework[J]. Journal of Marketing Analytics, 2022, 10(2): 106–113.
- [50] Barry T E, Howard D J. A review and critique of the hierarchy of effects in advertising[J]. Journal of Advertising, 1990, 9(2): 121–135
- [51] Strong Jr E K. Theories of selling[J]. Journal of Applied Psychology, 1925, 9(1): 75–86.
- [52] Hassan S, Nadzim S Z A, Shiratuddin N. Strategic use of social media for small business based on the AIDA model[J]. Procedia-Social and Behavioral Sciences, 2015, 172: 262–269.
- [53] Swinyard W R, Kenneth A C. Promotional effects on a high-versus low-involvement electorate[J]. Journal of Consumer Research, 1978, 5(1): 41–48.
- [54] Todri V, Ghose A, Singh P V. Trade-offs in online advertising: Advertising effectiveness and annoyance dynamics across

- the purchase funnel[J]. *Information Systems Research*, 2020, 31(1): 102–125.
- [55] Mu J, Zhang J Z. Seller marketing capability, brand reputation, and consumer journeys on e-commerce platforms[J]. *Journal of the Academy of Marketing Science*, 2021, 49: 994–1020.
- [56] Pop R A, Saplacan Z, Dabija D C, et al. The impact of social media influencers on travel decisions: The role of trust in consumer decision journey[J]. *Current Issues in Tourism*, 2022, 25(5): 823–843.
- [57] Humphreys A, Isaac M S, Wang R J. Construal matching in online search: Applying text analysis to illuminate the consumer decision journey[J]. *Journal of Marketing Research*, 2021, 58(6): 1101–1119.
- [58] 邝云娟, 傅科. 考虑消费者后悔的库存及退货策略研究[J]. *管理科学学报*, 2021, 24(4): 69–85.
Kuang Yunjuan, Fu Ke. Inventory and consumer returns policies under consumers' anticipated regret[J]. *Journal of Management Sciences in China*, 2021, 24(4): 69–85 (in Chinese)
- [59] 万谧宇, 柳键, 程永生, 等. 汽车分时租赁下消费者出行模式选择行为分析[J]. *管理科学学报*, 2022, 25(9): 32–51
Wan Miyu, Liu Jian, Cheng Yongsheng, et al. Analysis of consumer travel mode choices in car-sharing[J]. *Journal of Management Sciences in China*, 2022, 25(9): 32–51 (in Chinese)
- [60] 李双琦, 陈其安, 朱沙. 考虑消费与投资者情绪的股票市场资产定价[J]. *管理科学学报*, 2021, 24(4): 86–108
Li Shuangqi, Chen Qi'an, Zhu Sha. Asset pricing in the stock market with consumption and investor sentiment[J]. *Journal of Management Sciences in China*, 2021, 24(4): 86–108 (in Chinese)
- [61] 滕乐法, 吴媛媛, 李峰. 越沉浸越好吗? ——品牌体验中消费者沉浸程度的双重影响研究[J]. *管理世界*, 2020, 36(6): 153–167+251.
Teng Lefa, Wu Yuanyuan, Li Feng. Is immersion better? Research on the dual impact of consumer immersion in brand experience[J]. *Journal of Management World*, 2020, 36(6): 153–167+251. (in Chinese)
- [62] Pop R A, Săplăcan Z, Dabija D C, et al. The impact of social media influencers on travel decisions: The role of trust in consumer decision journey[J]. *Current Issues in Tourism*, 2022, 25(5): 823–843.
- [63] Mnih V, Heess N, Graves A. Recurrent Models of Visual Attention[C]. *Annual Conference on Neural Information Processing Systems*, New York: Curran Associates, 2014.
- [64] Vaswani A, Shazeer N, Parmar N, et al. Attention is All You Need[C]. *Annual Conference on Neural Information Processing Systems*, New York: Curran Associates, 2017.
- [65] Lu J, Yang J, Batra D, et al. Hierarchical Question-image Co-attention for Visual Question Answering[C]. *Annual Conference on Neural Information Processing Systems*, New York: Curran Associates, 2016.
- [66] Pennington J, Socher R, Manning C. GloVe: Global Vectors for Word Representation[C]. *The Conference on Empirical Methods in Natural Language Processing*, Stroudsburg: ACL, 2014.
- [67] Kingma D P, Ba J. Adam: A Method for Stochastic Optimization[C]. *arXiv preprint arXiv: 1412. 6980*, International Conference on Learning Representations, Ithaca: ArXiv, 2015.
- [68] Salakhutdinov R, Mnih A. Probabilistic Matrix Factorization[C]. *Annual Conference on Neural Information Processing Systems*, New York: Curran Associates, 2007.
- [69] Stewart G W. On the early history of the singular value decomposition[J]. *SIAM Review*, 1993, 35(4): 551–566.
- [70] Kullback S, Leibler R A. On information and sufficiency[J]. *Annals of Mathematical Statistics*. 1951, 22(1): 79–86.

Deep review-based recommendation from the perspective of consumer decision journey

ZHANG Ying-yi^{1, 2}, LI Xian-neng^{1, 2*}, GUO Yan-hong^{1, 2}, LI Xiao-gang^{1, 2},
ZHENG Shuang^{1, 2}

1. School of Economics and Management, Dalian University of Technology, Dalian 116024, China;
2. Institute for Advanced Intelligence, Dalian University of Technology, Dalian 116024, China

Abstract: The essence of recommender systems is to model the implicit preferences in consumer behavior. The human behavior is inseparable from psychology, and there are rich internal motives behind the superficial behavior. However, the current studies mainly focus on the behavioral data modeling, rarely involving the internal psychological activities and the information processing process of decision-making. Therefore, this paper studied a new idea of recommender systems by introducing AIDMA decision model from the perspective of consumer decision journey. A new deep review-based recommender system is proposed, which applies the AIDMA decision journey into the deep learning framework. Experiments showed that the recommendation performance of the proposal is significantly better than the state-of-the-art methods. This paper follows the big data-driven research paradigm of “model driven + data-driven”, realizing in-depth methodological innovation with theoretical support.

Key words: recommender system; deep learning; textual review; consumer decision journey