

文章编号: 2095-2163(2022)09-0139-04

中图分类号: TP312

文献标志码: A

航空运输业电商业务推荐算法应用分析

栗巍, 樊重俊, 许康

(上海理工大学管理学院, 上海 200093)

摘要: 近年来航空运输企业不断谋求非航空收入的新出路, 而电商业务是提高非航收入的重要途径, 诸多业内公司已在尝试第三方平台及自有平台的电商模式, 但数字化水平仍然不足。随着数字化技术的不断萌生, 类似智能推荐算法的技术已被第三方电商平台不断深入应用, 但对于航空运输业来说智能推荐算法在自建电商平台的应用还十分欠缺。本文通过分析智能推荐算法在航空运输业自建电商平台中的应用, 以此来探索智能推荐算法在该行业内的应用前景, 帮助行业获取更多营收思路。

关键词: 推荐算法; 电商; 航空运输业

Analysis of recommendation algorithm in e-commerce business of air transport industry

LI Wei, FAN Chongjun, XU Kang

(Business School, University of Shanghai for Science and Technology, Shanghai 200093, China)

[Abstract] In recent years, air transport companies have been constantly seeking new ways of non-aviation revenue, and e-commerce business is an important way to increase non-aviation revenue. Many companies in the industry have been trying e-commerce models on third-party platforms and their own platforms, but the level of digitalization is still insufficient. With the continuous emergence of digital technology, technologies similar to intelligent recommendation algorithms have been continuously applied by third-party e-commerce platforms, but for the air transport industry, the application of intelligent recommendation algorithms in self-built e-commerce platforms is still very lacking. This paper analyzes the application of intelligent recommendation algorithm in the self-built e-commerce platform of the air transportation industry to explore the application prospect of intelligent recommendation algorithm in this industry, and help the industry to obtain more revenue ideas.

[Key words] recommendation algorithm; e-commerce; air transport industry

0 引言

航空运输业的收入主要由航空收入与非航空收入构成, 行业内各大公司在收入结构调整上, 都在努力提高非航空收入转变。在非航业务, 如零售、餐饮、休闲等方面, 均需进一步推动其朝着更加精细化的方向发展^[1], 加之全国经济保持增长, 航空运输业企业所在地的经济对其非航业务的影响也较为显著, 居民生活水平提高对其零售等业务有着积极的促进作用^[2]。因此, 积极利用机场所带来的高客流量, 实现非航收入的高附加值, 从而优化公司整体收入结构是十分重要的发展方式。

相对于美国等发达国家在航空运输业的产业投入结构对比, 国内航空运输业极大依赖第二产业^[3]。因此, 基于本国国情, 国内航空运输业在餐饮、零售等第三产业相关方面做出突破和改变来促

进非航收入是更为现实的。对此, 抖音的推荐算法思路是一种值得研究的好方法。抖音的推荐算法结合用户偏好、社交关系、公共议题、场景、差异化和平台优先级等6个方面将内容精准推送给用户^[4]。推荐算法中基于用户信息与数据的协同过滤机制又将头部流量池的内容精准推荐给用户, 帮助抖音提高新用户留存率^[5]。其算法核心在于协同过滤, 目的在于挖掘用户潜在兴趣, 提高用户粘性。用户潜在兴趣转化与用户留存率提高的这样一种运行模式与航空运输业中伴随的高净值客流量具有较高契合度。

此外, 疫情的不利影响以及近年来国内航空运输业内部成本偏高、融资结构单一、利润率低等因素, 对整体航空运输业产生了较大的不良影响, 根据自身优势设置合理的运营模式至关重要^[6]。而随着数字技术的快速发展, 智慧机场建设进度加快, 智慧商业和全渠道数字化营销的运营模式逐步成为适合行业内

作者简介: 栗巍(1995-), 男, 硕士研究生, 主要研究方向: 数字经济、运营管理; 樊重俊(1963-), 男, 博士, 教授, 主要研究方向: 电子商务与数字经济; 许康(1996-), 男, 硕士研究生, 主要研究方向: 运营管理。

通讯作者: 樊重俊 Email: fan.chongjun@163.com

收稿日期: 2022-03-02

大部分企业的发展方向^[7]。因此,蕴含数字智慧的电商运营模式逐步被航空运输领域的各大企业所重视。从将推荐算法应用推向新高度的抖音兴趣电商直播模式来看,2020年初仅有海南航空一家相关航空运输公司进行了初步尝试,到目前已有超过十余家相关公司涉足其中,其中包括海南美兰机场、南方航空等大体量的公司。内容促进购买欲是吸引各大航空运输公司入驻的根本原因,也是抖音电商的最大优势^[8],究其本质还是推荐算法的应用。

推荐算法在第三方电商平台上的应用已经十分成熟,但智能推荐算法在航空运输业电商业务中的应用还在不断发展中。如抖音这类第三方电商平台,是当下推荐算法在航空运输业应用的雏形,航空运输业应用第三方电商平台的推荐算法,可有效提高非航收入。但对于本身具备较多优质流量的航空运输业来说,第三方平台存在一些问题,是电商业务发展的巨大桎梏:

(1) 仅仅是平台中的一个商铺,缺乏较多的自主控制权。

(2) 流量掌握在第三方平台手中,一旦停止推广投入,用户量将大幅下降。

(3) 不能得到详细用户数据,不利于用户营销分析,用户忠诚度相对较低。

因此,航空运输业发展具备推荐算法的自建电商平台,是未来打通自有商品零售与平台自有优质流量间通道的重要途径。但在自建电商平台不成熟情况下,平台只能起到广告作用^[9],发展自建电商平台还需结合自身优势进一步借鉴第三方平台的经验。

1 神经协同过滤推荐算法

航空运输业自建电商平台初期,自有的优质线下流量无法即时转换到线上平台。为了分析智能推荐算法在航空运输业电商业务中的具体应用,同时基于航空运输业电商零售业务具有的商品单价高、商品种类少而集中、平台流量多且优质等行业特性,应用神经协同过滤推荐算法框架 NeuMF (Neural Matrix Factorization)^[10]对小样本进行分析,从而大致反映整体应用情况。

神经矩阵分解模型 NeuMF,由广义矩阵分解 (Generalized Matrix Factorization, GMF) 与多层感知机 (Multilayer Perceptron, MLP) 相结合得到,是协同过滤算法结合神经网络的智能化推荐算法。在 NeuMF 中,用户和项目的稀疏向量分别通过 GMF 与 MLP 的 embedding 层,映射得到对应的隐式向量。

GMF 将自身中的隐式向量做哈达玛积运算得到交互特征,MLP 则通过感知机训练得到交互特征。在输出前将 GMF 与 MLP 的交互特征进行级联,再经过一层感知机的加权运算,最后通过 *sigmoid* 函数得到一个 0~1 之间的预测分数,来表示用户与物品间的相似度,基于此相似度进行推荐。其网络结构如图 1 所示。

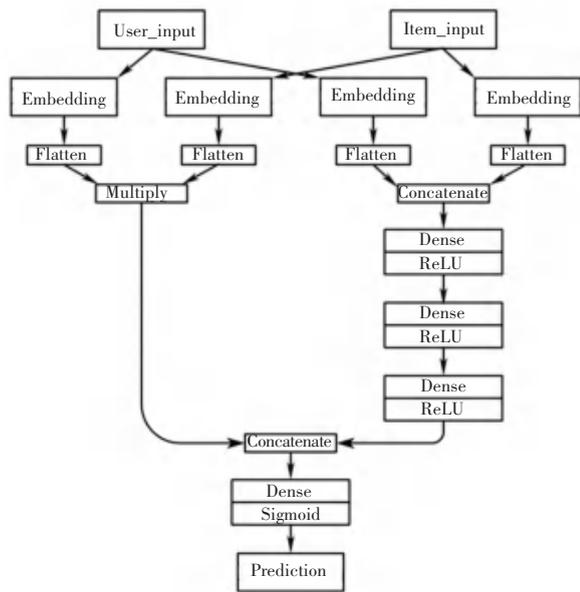


图1 NeuMF 神经协同过滤算法网络结构

Fig. 1 NeuMF neural collaborative filtering algorithm network structure

2 自建电商平台神经协同过滤算法实验

2.1 数据集收集与整理

本文通过问卷调查,获取了 687 份有效问卷。问卷内容包括年龄、性别、乘机次数、乘机目的、候机行为倾向、机场电商平台使用倾向等 6 个基础问题,用于分析问卷分布的合理性。20 个商品购买倾向的打分问题,用于调查购物类别倾向,作为后续实验数据。问卷主要获取了共计 5 181 个打分数据。

样本数据量的合理性根据 10 倍法则来确定,即模型需要超出其自由度 10 倍的数据量进行运算。此处的自由度是影响模型输出的参数,数据集包含 20 类商品,自由度为 20。通过 687 名调查对象得到了 5 181 条用户与商品类目的评分交互数据。数据量已远高于 NeuMF 模型自由度的 10 倍,因此数据样本的数量符合模型基本要求。同时,对于航空运输业中已筛选的优质流量来说,通过少部分的数据就能够对全局进行相对准确的反映。数据整理结果见表 1。

表 1 数据集组成

Tab. 1 Composition of the data set

数据集	交互项	项目	用户	稀疏度/%
问卷数据集	5 181	20	687	62.23

在 NeuMF 模型中进行训练的问卷数据,还要对其做后续的整理和分类,从而获得训练集、测试集、负例集。训练集和测试集中含有用户序号 *userid*、商品序号 *itemid*、打分 *rating* 三列数据,负例集从测试集中将用户序号 *userid* 和商品序号 *itemid* 提取出来,并在后面加上 10 类未与该用户交互过的商品序号 *itemid*。通过 3 个分类好的数据集组成完整的数据集数据结构,如图 2 所示。

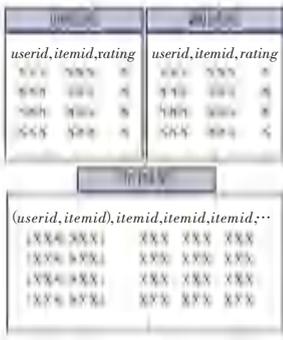


图 2 数据集结构

Fig. 2 The structure of the dataset

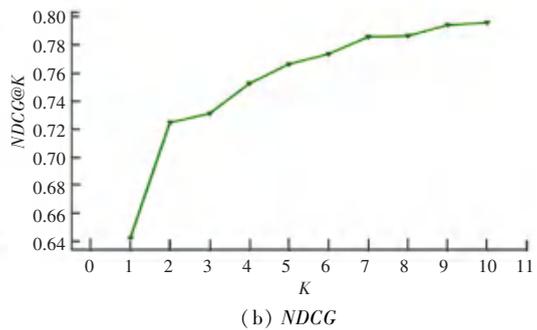
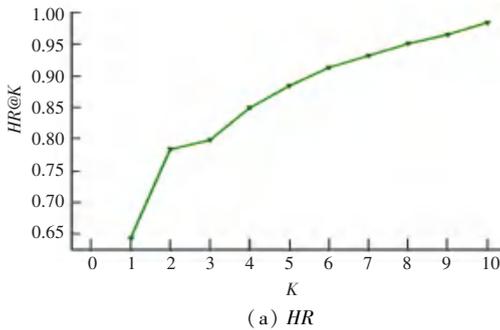


图 3 不同 K 值 (推荐数量) 下评价指标变化曲线

Fig. 3 Changes of evaluation indicators under different K values

由图 3 中可见, *HR* 与 *NDCG* 随 *K* 值的增大而显著提升。但在本例实验中, *K* 值大小对 *HR* 与 *NDCG* 具有不同的意义。 *K* 等价于从负例集的 11 个商品中推荐 *K* 个商品,其中包含测试商品的概率与测试商品在 *K* 个商品中的排名效果用 *HR* 与 *NDCG* 来描述。当 *K* 值较小时, *HR* 相比随机概率(从 11 个负例中随机抽取 *K* 个商品,测试商品在 *K* 个中的概率)优势越明显,越能反映 *HR* 指标。当 *K* 值越大时,推荐列表包含的商品越多,越能反映 *NDCG* 指标。

HR@K 与 *NDCG@K* 在本例实验的小样本问卷数据中表现较好,说明在行业特性的局限下,即使

2.2 评价指标选择

对 NeuMF 模型在数据集上训练的评价,选用常见的命中率 (Hits Ratio, *HR*) 和归一化折损累计增益 (Normalized Discounted Cumulative Gain, *NDCG*)。

Top-*K* 推荐算法 (前 *K* 项推荐算法) 是从用户交互过与未交互过的 *N* 个商品中,得到一个含有 *K* 个商品的推荐列表。在测试中,从该列表中取得交互过的商品即为推荐成功。

HR 表示在模型计算得到的推荐列表内,包含测试集中 *itemid* 个数的占比。就是按预测分数大小,对负例集中的 *itemid* 项进行排序,取前 *K* 个商品进行推荐,形成推荐列表,再计算全部测试集中 *userid* 对应 *itemid* 在推荐列表中的个数与测试集中 *itemid* 总数的比值,属于召回率的评价指标。 *NDCG* 用来评价推荐表的准确度和质量,数值介于 0 ~ 1 之间。其值越大,表明推荐的测试商品在推荐列表中的位置越靠前、评分越高,推荐效果越好。

2.3 训练结果分析

在本例中,由于数据集的限制,每个用户未交互过的负例最大数仅为 10,限制了负例的数量。因此,在较小负例数量中, *K* 值的少量增大会导致 *HR* 与 *NDCG* 的明显提高。在问卷数据集中,取不同 *K* 值的 *HR* 与 *NDCG* 指标如图 3 所示。

是在发展自建平台初期,针对用户数据与商品数据较少的情况下, NeuMF 等智能推荐算法也能够对用户进行高效推荐,发掘用户消费兴趣、促进购买成交,从而增加电商业务的整体营收。

2.4 实例验证分析

在进行了模型训练、指标评估等详尽分析后,为了实证模型的实用性与有效性,随机抽取 5 名用户 *id*, 根据先前的训练模型对这 5 名用户推荐 20 类商品中的 6 类,并与这些用户的实际交互商品进行对比,观察其推荐的准确率和质量。5 名用户的推荐实例情况见表 2。

表2 5名随机用户推荐实例

Tab. 2 Five random users recommendation examples

Userid	推荐列表(上行数据)与实际交互(下行数据)	命中个数	准确率/%
3	(3,0,1,2,8,7) (2,11,3,0,19,18)	3	50.0
56	(3,1,2,0,4,7) (6,4,17,12,3,0,2,7,18)	5	83.3
265	(6,5,8,17,11,7) (7,17,13,6,5,9)	4	66.7
389	(2,3,1,0,8,5) (3,0,2,1,15,12,8,4,5,13)	6	100.0
566	(2,3,1,8,5,0) (8,13,1,3,2,16,18,5,4)	5	83.3

对选取的5名用户做Top-6推荐,分别输出了5名用户的推荐列表。从表2中可以看出,最低的用户6个推荐中命中了3个,最高的用户6个推荐全部命中,说明实际推荐过程中推荐准确性较高。此外,从排序上来看,与实际交互情况有所差异,存在未交互过的商品排名比交互过的商品还高,说明推荐的质量有待提高。造成该问题的主要原因是数据集较小,在数据集不断增大后,该问题能够得到缓解。

在推荐列表中未交互过的商品,就是通过模型训练得出的对用户存在兴趣的推荐,说明了推荐的有效性。综上,NeuMF神经协同过滤推荐系统对航空运输业自建电商平台初期用户的实际推荐过程是有效合理的。

NeuMF这类智能算法在自建电商平台的合理性应用,将为航空运输业在非航收入的提升上带来巨大助力,同时自建电商平台的以下优势将进一步加强航空运输业本身所在平台的优势:

(1)平台独立,具备独立域名;商场个性化、细分化特征明显;平台的推广与普及将增加公司品牌价值;没有中间商赚差价。

(2)利用社交电商、内容创作的思路,能够获取更多更广的私域流量。通过对这些私域流量的经营,获取更高的价值,进一步加强自身平台优质流量优势。

(3)用户数据由自身绝对掌握。通过分析用户行为,实现针对性的推广与服务,提高自有用户粘性,增加自有流量转化率。

3 结束语

航空运输业收入多元化趋势势在必行,智能推

荐算法等数字化技术的融入为其多元化方式带来更多的可能。本文基于航空运输业发展电商业务的趋势,融入当下时兴的智能推荐算法,提出了应用自建电商平台结合智能推荐算法在航空运输业中应用的设想,并通过数据收集、整理形成规范数据集进行实验,初步验证了智能推荐算法结合自建电商平台在航空运输领域中应用的合理性。提出的这一模式将充分发挥航空运输业平台本身具备的有效流量优势。

参考文献

- [1] 陈广杰,张纯良,虞纯一. 机场非航业务精细化管理模式研究—以上海机场为例[J]. 企业科技与发展,2019(02):271-272.
- [2] 章连标,郝飞燕. 我国机场非航业务与腹地经济的互动关系研究—从非航收入视角[J]. 商业会计,2021(13):78-81.
- [3] 姜煦洲,杨文东. 基于投入产出法的中外航空运输业经济效益的对比研究[J]. 物流工程与管理,2022,44(01):109-112.
- [4] 温凤鸣,解学芳. 短视频推荐算法的运行逻辑与伦理隐忧—基于行动者网络理论视角[J]. 西南民族大学学报(人文社会科学版),2022,43(02):160-169.
- [5] 丁合蓉. 抖音短视频智能算法机制及问题研究[J]. 新媒体研究,2021,7(10):10-12.
- [6] 濮晗. 航空运输上市公司盈利能力研究[J]. 物流工程与管理,2020,42(07):63-66.
- [7] 邹立琨. 数字化背景下机场商业智慧化建设的对策[J]. 商业文化,2022(08):95-97.
- [8] 盖孟超,崔强. 抖音平台电商现状及改善措施[J]. 合作经济与科技,2022(08):88-90.
- [9] 王聪,陈燕,杨德礼. 具有自建电商平台的制造商多渠道运营策略研究[J]. 工业工程与管理,2018,23(02):76-84.
- [10] HE X, LIAO L, ZHANG H, et al. Neural collaborative filtering [C]//Proceedings of the 26th International Conference on World Wide Web(WWW 2017). Perth, Australia: ACM, 2017: 173-182.