

主播历史交易数据驱动的带货商品组合确定模型^{*}

禹春霞¹ 张雅楠¹ 王晓欢²

(1. 中国石油大学(北京) 经济管理学院, 北京 102249;

2. 东北大学 工商管理学院, 辽宁 沈阳 110167)

摘 要 随着直播电商销售模式兴起, 带货商品组合的确定成为提升销售效益和主播影响力的关键问题。然而, 现有选品方法往往缺乏对商品利润和综合表现的全面考量。这种因选品考量不全面导致的决策偏差, 不仅会影响直播销量, 还可能对主播的公众形象造成严重的负面影响。因此, 本文提出了一种主播历史交易数据驱动的带货商品组合确定模型以提升选品决策质量与带货效益。首先, 构建带货商品销量预测指标体系, 采用皮尔逊(Pearson)方法计算商品受众与主播受众相似度, 并采用 Transformer 预测销量; 其次, 采用基于指标相关性的权重确定方法-多属性边界近似区域比较法(criteria importance through intercriteria correlation-multi-attributive border approximation area comparison, CRITIC-MABAC)评价商品综合表现; 再次, 构建以利润最大化和商品综合表现最大化为目标的商品组合确定模型, 并采用 LHIS^①-NSGA-II算法求解。最后, 通过一个主播案例验证了提出的选品模型的有效性和优越性。

关键词 直播, 商品组合确定, 销量预测, NSGA-II, Transformer

中图分类号 C939

1 引言

在互联网快速发展的背景下, 直播电商已经成为推动消费增长的重要引擎^[1]。然而, 日益增长的商业合作也为主播带来了日益严峻的选品问题。选品恰当能够精准匹配消费者需求, 增强用户黏性, 提高消费者信任度; 而选品不当不仅会影响销量和利润, 还可能严重损害主播声誉, 进而引发消费者信任危机。消费者投诉报告数据显示, 2019年至2024年直播电商的投诉增幅高达47.1倍, 明显高于传统电商平台^[2]。主播直播带货频繁“翻车”, 根源在于带货商品选择失误。因此, 如何挑选恰当的带货商品或商品组合, 已成为主播亟待解决的问题。

在直播电商环境下, 主播通过交易商品赚取经济利润, 并依赖交易中形成的声誉资本增强客户的信任^[3]。因此, 主播选品应该着重考虑带货商品的预期利润和商品综合表现两方面。预期利润是主播商业价值的直接体现, 提高利润是直播带货的根本目标^[4]。带货商品综合表现, 包括商品质量、物流效率和售后服务等, 直接影响消费者的购买意愿和主播口碑^[5]。当前学术研究在直播带货领域主要聚焦于消费者感知^[6]、购买意愿^[7]以及主播互动性对直播效果的影响^[8], 侧重直播模式的概述和营销策略的分析。然而, 现有研究尚未充分关注以下方面: 第一, 尚未充分聚焦以主播为决策主体的带货商品组合确定研究, 相关工作多围绕消费者行为、主播表现等外部变量展开, 对主播视角下的多目标选品逻辑关注有限。

^{*} 基金项目: 国家自然科学基金面上项目(72271244)。

通信作者: 张雅楠, 中国石油大学(北京), 博士生, E-mail: zhangyanan2255@163.com。

① latin hypercube sampling-importance sampling, 拉丁超立方抽样-重要性抽样, 简称 LHIS。

第二,尚未充分利用直播带货场景下产生的海量多模态数据,对历史信息挖掘深度不足,不足以有效应对实际选品场景中的复杂挑战。

为解决上述问题,本文构建了融合 Transformer 架构、CRITIC-MABAC 与 LHS-NSGA-II 的综合选品决策模型。该模型包括三个阶段:第一,销量预测阶段,基于主播历史带货数据的多模态信息,采用 Transformer 预测商品销量;第二,商品综合表现评价阶段,基于多维指标体系,采用 CRITIC-MABAC 方法评价商品综合表现;第三,商品组合优化阶段,构建以利润最大化和商品综合表现最大化为目标的商品组合优化模型,采用 LHS-NSGA-II 算法求解,确定最优商品组合。

本文具有以下创新点:首先,针对直播带货商品销量预测困难问题,引入带货商品受众与主播受众相似度作为销量预测的输入变量之一,通过用户画像,比较主播受众和商品受众在性别、年龄和消费水平等维度的相似度,提高销量预测模型的准确性。其次,针对直播带货中备选商品数量庞大,商品组合空间数量呈指数级增长^[9]的问题,在商品组合确定阶段引入拉丁超立方抽样-重要性抽样方法生成初始化解种群。其中,拉丁超立方抽样将随机抽样和分层抽样相结合^[10],可以保证初始解在搜索空间均匀分布;重要性抽样可以增加有效样本的发生概率,从而减小方差,加快算法收敛速度^[11]。最后,引入自适应权重计算每个解的重要性得分,筛选保留得分高的个体,淘汰低质量解,并在筛选前使用局部归一化策略以解决多目标量纲不一致问题^[12]。

本文具有以下贡献:第一,提出了一种主播历史交易数据驱动的带货商品组合确定模型。该模型综合考虑了销量预测、商品综合表现评价与商品组合优化三个阶段,基于“先预测、再评价、后优化”的逻辑链条,保障了决策的科学性与客观性,为主播确定带货商品组合提供了决策支持。第二,构建了一种数据驱动的选品研究视角。基于直播平台的实际数据,充分挖掘历史交易信息的价值,爬取并整理了带货商品历史销量数据及品牌商信息,构建了可用于预测与评价的数据集,实现了对带货商品销量的精准预测,为主播确定带货商品组合提供了数据支持。第三,提出了一种多模态销量预测方法。针对直播带货场景中数据来源复杂、模态多样的特点,设计了基于 Transformer 架构的多模态融合方法,综合利用数值数据、文本数据以及图像数据,实现更高精度的销量预测,为直播电商多源异构数据的深度利用提供了新思路。

2 文献综述

本文旨在构建主播历史交易数据驱动的带货商品组合确定模型,本节将从选品优化、销量预测和商品综合表现三个方面梳理相关研究。

2.1 选品优化

直播带货中的“选品问题”即主播为直播筛选和确定带货商品组合的决策过程。精准选品不仅有助于提高直播转化率,还可以增强消费者满意度,为主播带来更高的经济效益,树立良好形象^[13]。传统电商选品研究大多以经济利益为核心,通过销量与利润确定最优商品组合^[14]。然而,随着电商生态的不断发展,学者逐渐意识到,单纯追求利润最大化难以满足复杂多变的市场环境。Shi 和 Sun^[15]强调产品质量对消费者满意度与复购率起决定性作用,Xu 等^[16]指出优质的产品性能、及时的物流效率以及完善的售后服务能够有效提升消费者的信任与忠诚度。Deng 等^[17]则关注消费者需求与偏好匹配度,提出基于用户画像或市场调研分析的方法能够提高商品与目标受众之间的契合度,进而增强消费者购买意愿。此

外,商品热度也是选品决策中重要的影响因素之一,可以为主播提供有效的决策支持^[18]。综上所述,现有研究多是从单一维度探讨选品问题,缺乏对利润和商品综合表现的系统性整合。然而,选品问题不是单一维度的问题,一方面,利润在选品策略中具有重要价值,直接关系到主播平台的经济效益。另一方面,商品综合表现也是选品策略的核心要素之一。优质的商品能够赢得消费者信赖,形成口碑效应,从而提高消费者的购买意愿和忠诚度,为直播间的长期稳定发展奠定基础。综上所述,本文提出一种综合考虑商品利润和综合表现的商品组合确定模型,以更好地满足直播电商复杂环境下的科学选品需求。

2.2 销量预测

在直播电商销售迅猛发展的背景下,精准的销量预测有助于主播选择预期销量高的商品,提高直播流量及商业变现效率^[19],从而提高直播带货利润。直播平台的动态性和互动性要求使用更先进的预测方法,以应对复杂的新兴市场环境。

传统的销售预测方法主要依托历史销售数据,运用统计和经济模型进行预测^[20-22],在市场条件相对稳定、数据模式特定的情景中具备一定有效性。然而,随着直播电商的兴起,市场环境日益复杂,销售过程涉及大量非结构化信息和高度动态的互动行为^[23,24],传统方法局限性愈发显著,难以适应直播电商的动态性和互动性。

近年来,深度学习和多模态学习技术的发展为销量预测提供了新的思路。Loureiro 等^[25]系统评估了人工神经网络、逻辑回归、支持向量机和随机森林等机器学习模型在新品销量预测中的表现。Joseph 等^[26]使用卷积神经网络(convolutional neural network, CNN)提取特征向量,结合双向长短期记忆(bidirectional long short-term memory, BiLSTM)学习时间模式以预测商品需求。学者在销量预测方面取得了丰硕成果,但直播电商场景的特殊性为销量预测带来了新的挑战,如何充分挖掘和融合非结构化数据中的多模态信息,成为直播商品销量预测的新问题。因此,多模态销量预测方法逐渐被应用到直播电商领域中。Xu 等^[13]基于主播声誉构建基于注意力机制的图文融合相似度推理(attention-based text-image fusion with similarity reasoning, A-tiFSR)框架,将文本、图像及弹幕等细粒度信息进行整合,验证了主播声誉与销量之间的正向关系;Lin 等^[27]设计两阶段全连接神经网络(fully connected neural network, FCNN)模型,融合用户行为与短视频广告信息,从营销模型视角预测直播流量及销量;Xu 等^[24]提出多实体多模态融合(multi-entity multimodal fusion, MEMF)框架,使用改进 Transformer 架构融合主播、商品和直播特征,在淘宝直播数据中表现优异。

上述研究虽在多模态数据销量预测方面取得重要进展,但尚未考虑商品受众与主播受众的相似度问题。而主播粉丝特性与商品目标市场的匹配度,直接决定了直播带货的购买转化率^[28]。因此,本文提出了一种考虑相似度的多模态预测框架。在特征提取阶段,使用基于 Transformer 的双向编码器表示(bidirectional encoder representations from transformers, BERT)技术提取文本特征,利用其强大的双向 Transformer 架构实现深度上下文语义理解^[29];使用残差网络(residual network, ResNet)提取商品图像特征,凭借残差结构有效解决深层网络训练中梯度消失的问题^[30,31]。在特征融合和预测阶段,引入注意力机制实现多模态特征的融合,并通过粉丝画像等信息量化主播受众与商品受众的相似度,精准捕捉主播与商品市场契合度之间的关系,有效弥补传统销量预测方法的不足。

2.3 商品综合表现

商品综合表现是消费者对商品及其相关服务的综合反馈。消费者在购买决策过程中会综合考虑商品

的功能属性与非功能属性,并通过评价表达其满意度^[32]。在直播电商场景下,商品综合表现的理论框架进一步扩展,涵盖了主播信誉和互动体验等独特因素,形成了服务、质量和物流等多维度的评价体系。首先,服务质量是直播带货商品综合表现的核心维度之一。优质完善的售后服务能够及时解决消费者问题,显著提升消费者满意度与忠诚度^[33]。高效的退换货服务有利于减少消费者的购买顾虑,提高其购买意愿^[34]。其次,商品质量是消费者评价的决定性因素,直接影响消费者的感知价值与购买行为^[35]。高质量商品不仅能够满足消费者的功能需求,还能降低退货率与投诉率,从而提高平台的整体口碑。最后,物流效率是直播电商的核心竞争力之一,高效的订单处理与配送服务能够显著提升消费者的满意度与复购意愿^[36]。

带货商品综合表现是一个多属性决策问题。属性权重在大多数多属性决策方法中起着至关重要的作用,权重确定方法可以分为主观赋权和客观赋权两种^[37]。主观赋权法根据专家的经验,比较各指标重要程度并对其进行赋权,结果过多地受到决策者主观偏好影响^[38]。客观赋权法基于各评价指标的数据差异对其进行赋权,结果具有客观优势,更加公正^[39]。其中,CRITIC 是熵权的增强迭代,在评估相互依赖的指标方面具有明显优势^[40]。多属性决策方法可以分为可补偿决策方法、不可补偿决策方法和行为决策方法三类。可补偿决策方法即决策者接受某方案属性的积极评价可以与其他劣势属性消极评价互相补偿。可补偿决策方法包括逼近理想解排序法(technique for order of preference by similarity to ideal solution, TOPSIS)^[41]、多准则妥协解排序法^[42]和 MABAC^[43]等,可以综合考虑方案的各种属性,更好地考虑全局状况,让商品的优劣属性相平衡。不可补偿决策方法即决策者不接受属性之间互相补偿,需要考虑属性的优先级,对备选方案进行排序。不可补偿决策方法包括消去选择转换法^[44]和偏好顺序结构评估^[45]方法等,方案的优缺点无法互补,即只要方案中的某一属性低于标准,即淘汰该备选方案。行为决策方法包括前景理论^[46]和交互式多属性决策^[47]方法等,是结合个人认知和行为规律理论的决策方法,更侧重于个人认知。商品综合表现需要考虑商品的整体信息对其进行评价,因此本文选择可补偿决策方法,其中 MABAC 考虑备选方案与边界近似区域(border approximation area)的距离^[48],结果精确,计算过程简单且易于和其他方法结合。因此,本文基于服务质量、商品质量和物流效率三个维度,使用 CRITIC 确定商品综合表现指标权重,使用 MABAC 计算商品综合表现值。

3 问题描述和假设

3.1 问题描述

如图 1 所示,带货商品组合确定过程涉及四个主体:品牌商、主播、平台和消费者。品牌商向主播提供商品,主播基于自身考量从中筛选商品,并借助平台向消费者销售商品。平台不仅承担交易与流量分发的功能,还提供了丰富的数据支持,基于平台可以获取主播带货商品的历史销量、商品属性以及品牌商相关资料,为科学选品提供数据依据。消费者观看主播的直播内容后,下单购买符合自身需求及购买意愿的商品。在这一过程中,为了提高利润和增强口碑,实现长期利益最大化,主播应充分利用商品数据和历史直播信息,审慎确定合适的商品组合。

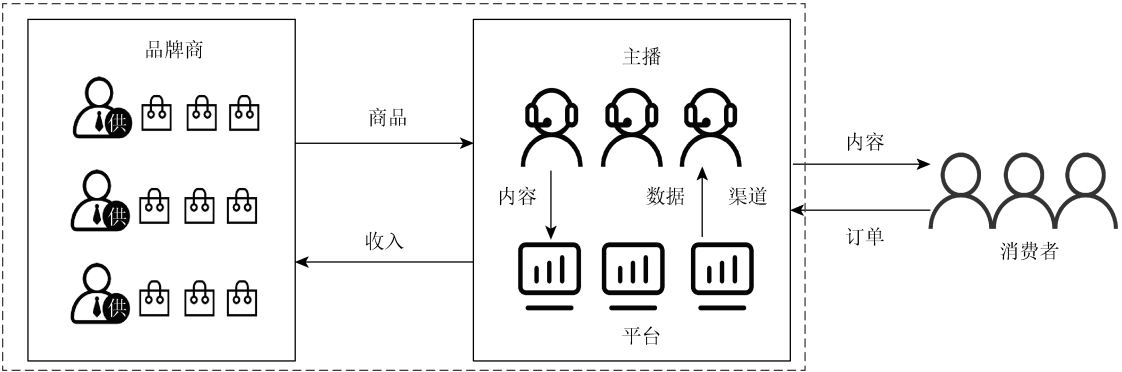


图 1 商品组合确定过程

3.2 符号

本文所使用的主要参数符号及其含义如表 1 所示。

表 1 参数符号

符号	含义
i	商品序号, $i = 1, 2, \dots, N$
j	评价指标序号, $j = 1, 2, \dots, M$
k	相似度影响因素序号, $k = 1, 2, \dots, K$
v	品牌商序号, $v = 1, 2, \dots, V$
c	品类序号, $c = 1, 2, \dots, C$
S_k	主播受众影响因素 k 的值
S_{ik}	商品 i 受众影响因素 k 的值
R_i	商品 i 受众与主播受众的相似度
B_i	商品 i 的佣金率
D_i	商品 i 的预测销量
P_i	商品 i 的单价
W_i	商品 i 的评价值
G_v	品牌商 v 提供的商品集合
F_c	属于品类 c 的品牌商集合
Q	单次直播最大可讲解商品数量
x_i	是否选择商品 i
δ_v	是否选中品牌商 v

3.3 假设

本文基于以下假设解决带货商品组合确定问题：

- (1) 主播选择的商品可以属于同一品牌商，也可以属于不同品牌商。
- (2) 商品不存在缺货问题，每个商品的库存水平都可以满足消费者的需求。
- (3) 由于时间限制，主播每场直播的带货商品数量有限。
- (4) 在同一场直播中，品牌商可以提供不同品类商品。
- (5) 模型求解所用商品单价、客户满意度和延迟交货率等信息均为已知且在下次直播前保持不变。

4 主播历史交易数据驱动的带货商品组合确定模型介绍

本文提出了一个三阶段带货商品组合确定模型，如图 2 所示。

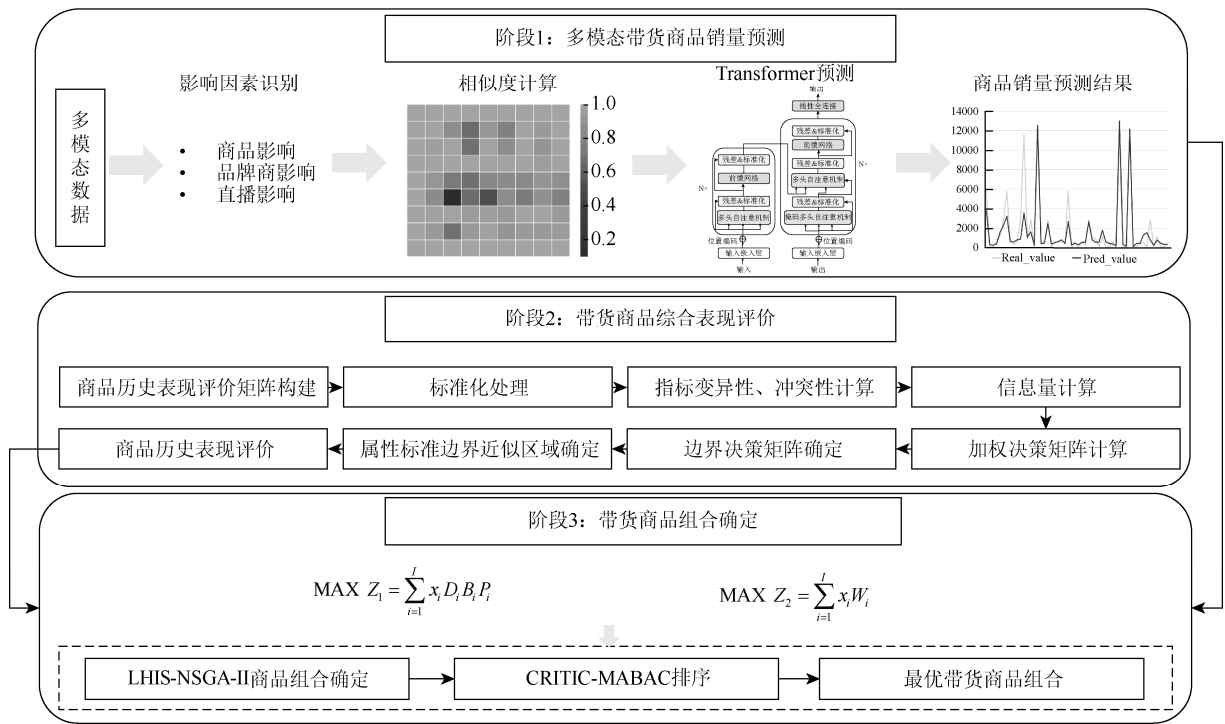


图 2 主播历史交易数据驱动的带货商品组合确定模型

(1) 阶段 1: 多模态带货商品销量预测。基于商品历史销量数据、商品信息，使用 Transformer 方法对未来销量进行预测。

(2) 阶段 2: 带货商品综合表现评价。使用 CRITIC-MABAC 方法对商品综合表现进行评价。

(3) 阶段 3: 带货商品组合确定。构建以商品利润最大化和综合表现最大化为目标的商品组合确定模型，使用 LHS-NSGA-II 算法求解，使用 CRITIC-MABAC 方法对帕累托解集进行排序，确定最优带货组合。

4.1 阶段 1：多模态带货商品销量预测

本节提出一种基于多模态数据融合的带货商品销量预测框架，通过整合数值数据、文本数据和图像数据，采用 Transformer 方法进行预测。

4.1.1 带货商品销量预测因素识别

本文基于商品属性、直播属性和品牌商属性三个维度进行商品销量预测。首先，商品属性涵盖商品标题、图片和价格等。商品标题在商品详情页中具有重要作用，清晰且具有吸引力的标题能够提高消费者的购买意愿；商品图片等视觉展示是最直观的信息，高质量图片和清晰的展示可以显著提升消费者的购买意愿；价格和价格折扣是消费者购买决策的核心影响因素之一^[49]，主播通常会通过强调价格优势，激发消费者的购买意愿；此外，商品与主播形象的高相似度可以提高消费者的信任感和认同感^[50]，进而提高其购买意愿。其次，直播属性主要包括讲解时长、上架时长和直播商品数量。主播对商品的详细讲解能帮助消费者更好地了解商品的功能和优势，从而增加其购买的可能性^[51]。商品在直播中的展示时间越长，消费者接触到的机会越多，进而可能提高销量。单次直播中展示的商品数量会影响消费者的注意力和决策效率。最后，品牌商属性主要包括关联达人和客户满意度。与商品关联的达人数量越多，商品的曝光率和可信度越高，但同时也会造成消费者分流；高客户满意度评分能增强消费者对商品的信任，进而提高消费者的购买意愿^[50]。综上所述，本文选择的带货商品销量预测影响因素如表 2 所示。

表 2 带货商品销量预测影响因素

类型	影响因素	指标解释	度量单位
商品属性	商品标题	主播选择的商品标题	
	商品图片	主播选择的商品图片	
	价格	商品在直播间的销售价格	元
	价格折扣	(商品日常价格-商品价格)/商品日常价格	%
	带货商品受众与主播受众相似度	带货商品受众与主播受众的相似度	%
直播属性	讲解时长	主播讲解商品或内容的总时长	秒
	上架时长	商品在直播中展示的时间长度	秒
	直播商品数量	单次直播中展示的商品总数	个
品牌商属性	关联达人	该商品所关联的达人数量	位
	客户满意度	客户满意度评分	分(1~5分)

4.1.2 带货商品受众与主播受众相似度计算

带货商品受众和主播受众的契合程度计算是一个相似性计算问题，其中，带货商品受众指购买该商品用户的群体特征^[52]，主播受众指在该主播直播间活跃的典型观众的群体特征^[53]。在直播带货中，商品受众与主播受众的契合程度，会在很大程度上左右消费者的购买决策。当商品受众与主播受众存在显著差异时，消费者的购买意愿往往会受到抑制^[54]。反之，当商品受众与主播受众具有较高的相似度时，则商品更有可能契合消费者的心理预期，有利于增加直播销量。本文选取的相似度影响因素如表 3 所示。

表 3 带货商品受众与主播受众相似度计算影响因素

类型	影响因素	含义
消费者因素	性别	消费者的性别比例
	年龄	消费者的年龄比例, 分为 5 个年龄段
	人群分布	消费者的类型比例, 分为 8 个类型
	活跃度分布	消费者的活跃度
地域因素	城市线分布	消费者所在城市的类型比例, 分为 8 个级别
商品因素	成交类目	消费者所购买的商品类型
	商品价格水平分布	商品的价格水平, 分为 6 个区间

本文基于商品受众信息和主播受众信息,使用 Pearson 方法计算商品受众与主播受众相似度。Pearson 相似度能够刻画两组数据之间所显现的线性相关程度,其取值范围在-1 和 1 之间。具体过程如下。

Step 1: 数据均值处理。使用 Pearson 方法进行计算时,需要减去该主播或商品受众相似度影响因素数据的平均值。

Step 2: 计算相似度。通常使用以下公式计算相似度。

$$R_i = \frac{\sum_{k=1}^K (S_k - \bar{S}) \times (S_{ik} - \bar{S}_i)}{\sqrt{\sum_{k=1}^K (S_k - \bar{S})^2 \times \sum_{k=1}^K (S_{ik} - \bar{S}_i)^2}} \quad (1)$$

其中, S_k 表示主播受众在第 k 个影响因素上的数值; S_{ik} 表示商品 i 受众在第 k 个影响因素上的数值; $\bar{S}$ 和 $\bar{S}_i$ 分别表示主播受众和商品受众所有影响因素的平均值。

4.1.3 多模态销量预测

本文使用基于多头自注意力机制的 Transformer 架构进行多模态销量预测,具体流程如图 3 所示。

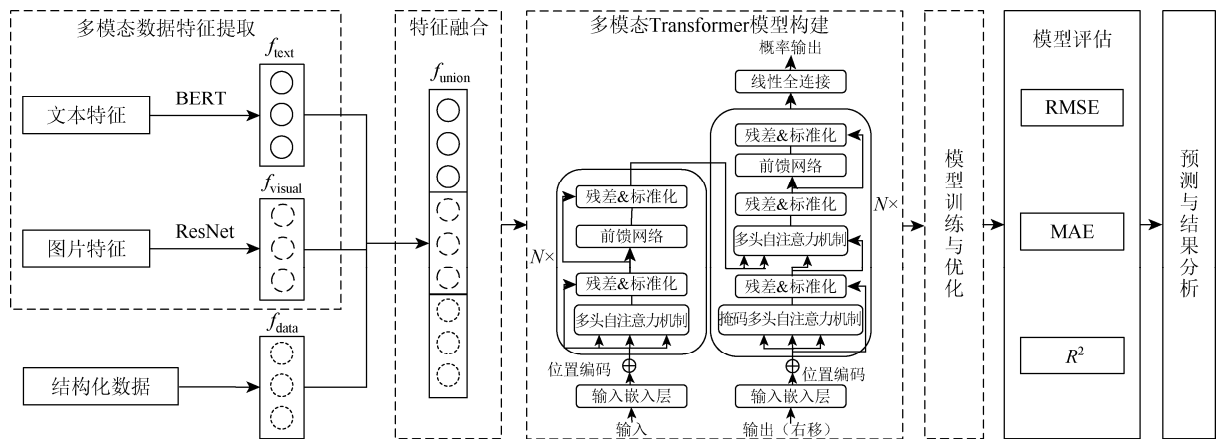


图 3 多模态预测模型架构

Step 1: 多模态数据特征提取。首先,使用 BERT 提取文本数据特征;然后,使用 ResNet 提取图像数据特征。

Step 2: 多模态特征融合。利用自注意力机制,对不同模态间的交互关系进行建模,捕捉各模态间的关联,有效整合多模态信息,为后续模型提供更丰富的特征表示。

Step 3: 多模态 Transformer 模型构建。定义 Transformer 模型的结构,模型由三个模态的特征投影层、多层 Transformer 编码层和输出层构成,以最小化预测销量与真实销量误差为训练目标。

Step 4: 模型训练与优化。使用训练集数据对多模态 Transformer 模型进行训练:使用 Adam 优化器更新模型参数,设置权重衰减防止过拟合。在每个训练周期内,计算训练集上的损失并调整参数。

Step 5: 模型评估。使用测试集数据评估模型的性能,计算均方根误差(RMSE)、平均绝对误差(mean absolute error, MAE)和 R^2 等指标,全面衡量模型预测精度与稳定性。

Step 6: 预测与结果分析。使用训练好的多模态 Transformer 模型对未来的带货商品销量进行预测。分析预测结果,评估模型的预测精度和泛化能力。

4.2 阶段 2: 带货商品综合表现评价

4.2.1 带货商品综合表现指标选取

如表 4 所示,本文基于服务质量、商品质量和物流效率三个维度进行带货商品综合表现评价。首先,服务质量涵盖店铺体验和商家服务等。店铺体验注重消费者购物过程中的主观感受;商家服务涵盖了售前、售中和售后全流程的服务水平^[55]。其次,商品质量涵盖商品体验和好评率。商品体验直观展示了商品自身特性是否符合消费者期望;好评率是消费者对商品质量的综合评价^[56]。最后,物流效率即消费者物流体验,主要衡量消费者对商品运输效率的满意程度^[57]。

表 4 带货商品综合表现指标

类型	影响因素	含义	度量单位
服务质量	店铺体验	消费者浏览选购的页面体验,分为 5 个等级	级(1~5 级,1 级最低,5 级最高)
	商家服务	消费者咨询售后的沟通服务水平	分(1~5 分)
商品质量	商品体验	消费者使用产品的实际感受	分(1~5 分)
	好评率	商品的好评比例	%
物流效率	物流体验	物流时效与配送服务质量	分(1~5 分)

4.2.2 带货商品综合表现评价

本文提出了一种基于 CRITIC-MABAC 的带货商品综合表现评价方法,使用 CRITIC 计算商品综合表现指标的权重,使用 MABAC 计算商品综合表现值,进而对商品进行排序。具体实施步骤如下:

Step 1: 构建商品综合表现矩阵。

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1M} \\ r_{21} & r_{22} & \cdots & r_{2M} \\ \vdots & \vdots & & \vdots \\ r_{N1} & r_{N2} & \cdots & r_{NM} \end{bmatrix} (i=1,2,\cdots,N; j=1,2,\cdots,M) \quad (2)$$

其中, M 表示商品个数; N 表示评价指标选取个数; r_{ij} 表示矩阵中的第 i 行第 j 列数据。

Step 2: 标准化处理。对原始数据进行正向化或逆向化无量纲化处理,消除评价指标属性不同造成的影响。

$$r'_{ij} = \frac{r_j - r_{\min}}{r_{\max} - r_{\min}} \quad (3)$$

$$r'_{ij} = \frac{r_{\max} - r_j}{r_{\max} - r_{\min}} \quad (4)$$

Step 3: 计算指标变异性。

$$\begin{cases} \bar{r}_j = \frac{1}{n} \sum_{i=1}^n r_{ij} \\ S_j = \sqrt{\frac{\sum_{i=1}^n (r_{ij} - \bar{r}_j)^2}{n-1}} \end{cases} \quad (5)$$

其中, S_j 表示第 j 个指标的标准差。

Step 4: 计算指标冲突性。

$$H_j = \sum_{i=1}^P (1 - h_{ij}) \quad (6)$$

其中, h_{ij} 表示评价指标 i 和 j 之间的相关系数。

Step 5: 计算信息量。

$$C_j = S_j \times H_j \quad (7)$$

其中, C_j 越大, 则第 j 个评价指标在整个评价指标体系中的作用越大。

Step 6: 计算各指标权重。

$$W_j = \frac{C_j}{\sum_{j=1}^P C_j}, j=1, 2, \dots, M \quad (8)$$

Step 7: 计算加权决策矩阵。

$$v_{ij} = (x_{ij} + 1) \omega_j \quad (9)$$

$$V = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1M} \\ v_{21} & v_{22} & \cdots & v_{2M} \\ \vdots & \vdots & & \vdots \\ v_{N1} & v_{N2} & \cdots & v_{NM} \end{bmatrix} \quad (10)$$

Step 8: 计算每个指标的边界近似区域 g_j , 确定边界决策矩阵 G 。

$$g_j = \left(\prod_{i=1}^M v_{ij} \right)^{\frac{1}{M}} \quad (11)$$

Step 9: 计算元素 v_{ij} 与 g_j 之间的距离, 得到距离矩阵。

$$q_{ij} = v_{ij} - g_j \quad (12)$$

$$Q = \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1M} \\ q_{21} & q_{22} & \cdots & q_{2M} \\ \vdots & \vdots & & \vdots \\ q_{N1} & q_{N2} & \cdots & q_{NM} \end{bmatrix} \quad (13)$$

Step 10: 确定指标标准边界近似区域。

若 $q_{ij} > 0$ ，则该方案位于上近似区域；若 $q_{ij} = 0$ ，则该方案位于边界上；若 $q_{ij} < 0$ ，则该方案位于下近似区域。商品综合表现的评价结果位于上近似区域的方案越多，则说明该商品的综合表现越好。

$$A_i \in \begin{cases} G^+, & q_{ij} > 0 \\ G, & q_{ij} = 0 \\ G^-, & q_{ij} < 0 \end{cases} \quad (14)$$

Step 11: 评价商品综合表现。

$$W_i = \sum_{j=1}^M q_{ij} \quad (15)$$

4.3 阶段 3：带货商品组合确定

4.3.1 带货商品组合确定模型构建

本文构建的带货商品组合确定模型旨在实现最大化利润和商品综合表现。模型如下：

$$\text{MAX } Z_1 = \sum_{i=1}^N x_i D_i B_i P_i \quad (16)$$

$$\text{MAX } Z_2 = \sum_{i=1}^N x_i W_i \quad (17)$$

Subject to

$$x_i = \begin{cases} 1, & \text{选择商品 } i \\ 0, & \text{其他} \end{cases} \quad (18)$$

$$\sum_{i=1}^N x_i \leq Q \quad (19)$$

$$\sum_{v \in R_c} \delta_v \leq k \quad (20)$$

$$\delta_v = \begin{cases} 1, & \sum_{i \in G_v} x_i \geq 1 \\ 0, & \text{其他} \end{cases} \quad (21)$$

目标函数（16）表示最大化利润，即主播通过佣金获得的利润。

目标函数（17）表示最大化商品综合表现。

变量约束（18）表示是否选择商品。

商品数量约束（19）表示在一场直播中主播详细展示的商品数量不可以超过某个固定值，该值由主播历史直播中详细展示的商品数量决定。

品牌商约束（20~21）表示在一场直播中主播最多只能详细展示 k 家品牌商同一品类的商品。

4.3.2 带货商品组合确定模型求解

带货商品组合确定模型求解具体步骤如下所示。

Step 1: 模型构建，输入商品数据，构建带货商品组合确定模型。

Step 2: 初始化种群，确定种群规模，基于 LHIS 方法生成一组初始解作为初始化种群。

Step 3: 对初始种群进行快速非支配排序以及选择、交叉和变异等操作。

Step 4: 条件判断，如果生成第一代种群则当前进化代数=2，如果没有生成第一代种群则重复 Step 3。

Step 5: 合并父代种群与子代种群。

Step 6: 对新种群进行快速非支配分层排序并计算拥挤度，归一化处理目标函数，筛选出新的种群。

Step 7: 对新种群进行选择、交叉、变异等操作。

Step 8: 判断是否达到终止条件或最大进化代数。若已达到，结束进化，输出帕累托最优解集；否则迭代次数增加一次，重复执行 Step 5~Step 7。

在获得帕累托最优解集后，需要在解集中筛选契合主播需求的最优商品组合确定方案。本文使用 CRITIC-MABAC 方法对帕累托解集中的商品组合进行排序择优。

5 实验分析

本节通过抖音平台案例证明提出的带货商品组合确定模型的实用性。首先，爬取相关数据并进行数据预处理；其次，计算带货商品受众与主播受众相似度，进而预测带货商品销量；再次，评价带货商品综合表现；最后，构建多目标优化模型确定直播商品组合，帮助主播更好地满足消费者需求，提高市场竞争力。

5.1 实验说明

Y 主播是一名拥有 347.7 万名粉丝的美妆品类头部主播，带货口碑评分为 4.9，主要带货品类为美妆个护与食品百货等。目前，与 Y 主播合作的品牌商超 60 家。随着 Y 主播知名度提高，越来越多的品牌商希望与 Y 主播达成合作。因此，Y 主播面临着商品组合确定的难题。本文以该主播为例，假设该主播自主运营账号（拥有选择带货商品的权利）且团队不经营制造工厂（不生产商品），通过爬取该主播历史带货商品数据及合作品牌商数据，帮助其完成商品受众与主播受众相似度计算、带货商品销量预测、商品综合表现评价以及商品组合确定等工作。

本文使用抖音直播电商平台、抖查查和飞瓜数据等平台的 Y 主播 2024 年 3 月至 2024 年 4 月的直播交易数据。数据集包含 28 场直播、306 种商品，共 1203 条数据。关键变量的描述性统计如表 5 所示。

表 5 关键变量描述性统计

变量名称	均值	标准差	最小值	最大值	中位数
价格	156.71	200.36	15	3 300	109
价格折扣	0.86	0.23	0	1	1
带货商品受众与主播受众相似度	0.73	0.31	0	0.96	0.85
讲解时长	900.08	1 294.58	0	15 818.00	741.50
上架时长	15 668.97	9 248.13	3 000.00	45 369.00	13 269.00
直播商品数量	72	21	5	115	68
店铺体验	4.89	0.13	4.50	5.00	4.90
商家服务	4.81	0.29	3.70	5.00	4.90
商品体验	4.89	0.16	4.00	5.00	4.80

续表

变量名称	均值	标准差	最小值	最大值	中位数
好评率	94.21	5.31	50.00	100.00	94.84
物流体验	4.91	0.20	3.30	5.00	4.80

5.2 实验实现

5.2.1 阶段 1：带货商品销量预测

如表 6 所示，基于式（1）计算带货商品受众和主播受众的相似度。

表 6 带货商品受众与主播受众各因素的相似度

商品	性别	年龄	人群分布	活动分布	城市线分布	成交类目	商品价格水平分布	总相似度
G01	1.000	0.646	0.959	0.930	0.620	0.548	0.544	0.799
G02	1.000	0.991	0.971	0.980	0.992	0.922	0.844	0.959
G03	1.000	0.991	0.971	0.980	0.992	0.922	0.844	0.959
G04	1.000	0.379	0.962	0.950	0.713	0.578	0.927	0.804
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
G62	1.000	0.865	0.987	0.990	0.996	0.734	0.095	0.816
G63	1.000	0.865	0.987	0.990	0.996	0.734	0.095	0.816

带货商品销量预测模型的输入变量如表 2 所示。其中随机选取 70%的数据集作为训练集，进行预测模型训练；剩余 30%数据集为测试集，进行模型精度检验。

参数设置如下：学习率为 0.0001，批量大小为 8，训练轮数为 200，学习率调度器步长为 2000，隐藏层维度为 256，注意力头数为 8。基于上述参数，使用 Transformer 方法预测销量， R^2 为 0.871，平均绝对误差为 78.20。带货商品销量预测结果如表 7 所示。

表 7 带货商品销量预测结果

商品	预测销量
G01	7508
G02	2345
G03	2674
G04	298
⋮	⋮
G62	60
G63	75

5.2.2 阶段 2：带货商品综合表现评价

如表 8 所示，对带货商品综合表现进行评价。

表 8 带货商品综合表现信息

商品	店铺体验	商品体验	物流体验	商家服务	好评率
G01	5.00	5.00	5.00	5.00	0.95
G02	4.90	4.95	4.95	4.95	0.94
G03	4.90	4.95	4.95	4.95	0.94
G04	4.80	4.80	4.80	4.80	0.94
⋮	⋮	⋮	⋮	⋮	⋮
G62	5.00	5.00	5.00	5.00	0.95
G63	5.00	5.00	5.00	5.00	0.95

根据式（2）～式（8），计算带货商品综合表现各指标权重，其中店铺体验 0.15，商品体验 0.19，物流体验 0.23，商家服务 0.20，好评率 0.24。根据式（9）～式（15），评价商品综合表现，评价结果如表 9 所示。

表 9 带货商品综合表现评价结果

商品	评价值
G01	99.97
G02	81.15
G03	81.15
G04	79.84
⋮	⋮
G62	99.96
G63	100.00

5.2.3 阶段 3：带货商品组合确定

为主播选择带货商品，使该次直播达到利润和商品综合表现最大化的目标。商品信息如表 10 所示，其中销量和商品综合表现值分别来源于前文的销量预测和商品综合表现评价；基于主播历史直播信息设置 k 值为 2。使用 LHis-NSGA-II 算法求解带货商品组合优化模型。设置运行参数，种群大小为 500，最大迭代数为 2000，变异概率为 0.3，交叉概率为 0.8，终止进化条件为达到最大进化代数，基于 Geatpy 遗传工具包求解。

表 10 带货商品组合确定信息

商品	价格	预测销量	佣金率	商品综合表现
G01	39.80	7508	1%	99.97
G02	100.00	2345	1%	81.15
G03	100.00	2674	1%	81.15
G04	150.00	298	0.5%	79.84
⋮	⋮	⋮	⋮	⋮
G62	27.90	60	1%	99.96
G63	54.80	75	1%	100.00

使用 CRITIC-MABAC 方法对帕累托解集进行排序, 确定最优商品组合方案为 G1, G14, G19, G28, G34, G38。

5.3 对比分析

本节通过对比实验证明所提出的带货商品组合确定模型的优越性。

5.3.1 多模态带货商品销量预测模型对比分析

本节旨在通过对比实验证明本文提出的考虑带货商品受众与主播受众相似度的多模态带货商品销量预测方法的优越性。控制其他变量, 分别进行考虑带货商品受众与主播受众相似度和忽略带货商品受众与主播受众相似度的带货商品销量预测。预测结果评价如表 11 所示。

表 11 预测方法对比

	考虑相似度	忽略相似度
MAE	78.20	192.52
RMSE	215.25	236.89
R^2	0.87	0.84

基于不同模型的多模态销量预测结果, 得出结论如下: 如表 11 所示, 两种预测方法主要有以下差异: 首先, 考虑相似度的预测模型的预测误差较小, 仅为 78.20。其次, 考虑相似度的预测模型的预测结果与真实值的偏差较小, 仅为 215.25。最后, 考虑相似度的预测模型的拟合效果较好, 可以更好地捕捉数据中的复杂关系, 其 R^2 为 0.87。综合对比实验中各方法预测性能, 考虑相似度的预测模型在 MAE、RMSE 和 R^2 三个指标上均显著优于忽略相似度的预测模型, 表明其预测精度和稳定性更高, 且模型拟合效果更好。因此, 本文选择考虑相似度的模型预测带货商品销量。

5.3.2 带货商品组合确定模型对比分析

本节旨在通过对比实验证明本文提出的带货商品组合确定模型的优越性。因此, 本文基于历史数据构建了五组差异化商品组合模型: ①基于主播经验的商品组合确定模型; ②仅以利润为导向的商品组合确定模型; ③仅以商品综合表现为导向的商品组合模型; ④不考虑销量预测的商品组合确定模型; ⑤本文提出的考虑销量预测的商品组合确定模型。在实验中, 使用 2024 年 3 月 1 日至 3 月 31 日的直播销售记录进行模型训练与商品组合确定, 使用 2024 年 4 月的直播销售记录计算利润和商品综合表现。其中, 利润定义为商品销售的总收入乘以佣金率, 反映主播的短期经济收益; 商品组合表现综合考虑服务质量、物流质量和商品质量, 衡量主播的长期市场竞争力。模型 1 为当次主播根据经验选择直播商品组合, 结果即为当次直播的实际数据结果。模型 4 设置历史平均销量为统一数据, 旨在消除预测销量波动对商品组合确定结果的干扰, 为对比分析销量预测的贡献提供参考。此外, 本文引入综合评分指标用于综合衡量短期收益与长期竞争力。不同模型下的商品组合确定结果如表 12 所示。

表 12 商品组合对比

模型	商品组合	利润	商品综合表现	综合评分
模型 1	G1, G2, G4, G6, G8, G25	4559.00	420.78	0.32
模型 2	G1, G6, G8, G14, G22, G34	7564.73	351.73	0.60

续表

模型	商品组合	利润	商品综合表现	综合评分
模型 3	G1, G15, G28, G39, G51, G61	3242.78	558.78	0.40
模型 4	G1, G14, G26, G28, G34, G56	6193.23	516.76	0.73
模型 5	G1, G14, G19, G28, G34, G38	6424.73	516.84	0.76

注：综合评分=0.5×标准化利润+0.5×标准化商品综合表现，标准化值基于实验数据的归一化处理，其中利润范围为[3242.78, 7564.73]，商品综合表现范围为[351.73, 558.78]

基于不同模型下的带货商品组合确定结果，得出结论如下：首先，引入商品综合表现的模型在整体效果上优于单纯依赖经验或单一目标导向的模型。例如，模型 4 的综合评分为 0.73，显著高于其他对比模型。其次，考虑到销量预测能够更准确地反映市场需求，引入销量预测数据后，商品组合决策效果得到了进一步优化。例如，模型 5 在利润和商品综合表现方面均优于模型 4。综上所述，考虑销量预测的带货商品组合确定模型更有利于提高主播效益，验证了本文提出模型的意义。

5.3.3 带货商品组合确定方法对比分析

本节旨在通过对比实验证明本文提出的带货商品组合确定方法的优越性。因此，本文基于已有文献^[58, 59]分别使用三种方法确定商品组合：①基于阈值筛选的商品组合确定方法；②基于整数规划的商品组合确定方法；③本文提出的考虑销量预测的商品组合确定方法。不同方法的商品组合确定结果如表 13 所示。

表 13 带货商品组合确定方法对比

模型	商品组合	利润	商品综合表现	综合评分
模型 1	G1, G9, G14, G15, G26, G56	5997.25	516.90	0.50
模型 2	G1, G8, G14, G22, G26, G34	7253.73	437.09	0.50
模型 3	G1, G14, G19, G28, G34, G38	6424.73	516.84	0.76

注：综合评分=0.5×标准化利润+0.5×标准化商品综合表现，标准化值基于实验数据的归一化处理，其中利润范围为[5997.25, 7253.73]，商品综合表现范围为[437.09, 516.90]

基于不同模型下的带货商品组合确定结果，可得出如下结论：首先，基于阈值筛选的方法在商品综合表现上具有一定优势，但整体效益提升有限。例如，该方法的商品综合表现为 516.90，但其利润水平仅为 5997.25。其次，基于整数规划的方法能够获得更高的利润，但综合评分相对较低。例如，该方法的利润为 7253.73，但其商品综合表现仅为 437.09。相比之下，本文提出的模型在利润与商品综合表现之间实现了更优平衡，其综合评分达到 0.76，明显优于其他方法。综上所述，现有模型存在单一利润或单一表现导向的局限性，难以同时兼顾主播的短期效益与长期市场竞争力；而本文提出的考虑销量预测的商品组合确定方法在收益与表现两个维度上均取得较好结果，更适用于实际直播带货场景，验证了本文方法的有效性与优越性。

6 管理启示

基于上述实验结论，为切实提高主播效益，本文提出以下两点管理启示。

第一，主播应当构建“利润-口碑”协同优化的商业价值提升路径。在经济层面，需要利用模型协

同优化商品利润与综合表现,实现短期利润最大化;在信誉层面,需要通过高质量商品组合维护长期粉丝信任与积累品牌口碑。实现在持续激励主播保持直播热情的同时,吸引更多消费者、增加消费者留存度,推动主播商业价值螺旋式上升。

第二,主播应当注重直播间商品组合的数据吸收与转化能力,充分发挥运营团队在其中的引导作用。当前电商直播竞争日趋激烈,用户需求呈现多元化特征,商品信息规模呈指数级膨胀。在实践中,高佣金、高利润商品未必能够形成销售与口碑的正向循环,部分低质爆品可能导致粉丝信任下降、退货成本增加。因此,主播应加强对选品团队的培训,使其既能精准解读历史交易数据,又能动态捕捉实时流量趋势,确保直播间能够将数据洞察转化为可持续的销售业绩与用户黏性。

7 总结

主播历史交易数据驱动的带货商品组合确定模型对主播选择商品具有重要意义,合理的商品组合确定有利于主播贴合自身形象,满足消费者需求,降低退货率,提高利润。为了解决直播电商头部主播选品难的问题,本文通过分析历史直播数据,根据品牌商信息、商品信息以及历史销量提出了主播历史交易数据驱动的带货商品组合确定模型,第一,采用 Pearson 方法计算商品受众与主播受众相似度;第二,基于相似度和商品信息等多模态数据,采用 Transformer 方法,预测主播未来带货商品的销量;第三,采用 CRITIC-MABAC 方法评价商品综合表现;第四,基于商品信息和销量预测结果以及商品综合表现评价结果,构建利润和商品综合表现最大化的商品组合确定模型,采用 LHS-NSGA-II 算法求解并确定最优商品组合。第五,本文对提出的模型进行算例分析,验证了模型的有效性,并通过对比性分析验证了不同商品组合模型结果的差异,证明本文模型对直播商品组合确定具有一定意义。首先,销量预测可以预测市场趋势和消费者偏好,适应市场波动,从而更精准地满足消费者需求。其次,商品综合表现可以帮助主播构建品控体系,从海量繁杂的商品中选出品质上乘、性价比高的商品组合。综上所述,由主播历史交易数据驱动的直播商品组合确定模型可以帮助主播更精准地把握市场动态,优化直播策略,提升销售业绩,同时增强与粉丝的互动和黏性,为长期发展奠定坚实基础。

未来可以探索几个研究方向。例如,可以考虑主播直播数据时时更新的问题,利用增量学习等方法提高预测模型的精度,更好地进行销量预测;还可以考虑新商品缺乏历史数据的冷启动问题,通过迁移学习等方法,对新商品的销量进行预测,提高模型精度。

参 考 文 献

- [1] Peng X X, Ren J Q, Guo Y T. Enhance consumer experience and product attitude in E-commerce live streaming: based on the environmental perspective[J]. Industrial Management & Data Systems, 2024, 124 (1): 319-343.
- [2] 国务院政策例行吹风会: 消费者权益保护法实施条例有关情况[EB/OL]. [http://www.scio.gov.cn/live/2024/33697/tw/\[2025.06.12\]](http://www.scio.gov.cn/live/2024/33697/tw/[2025.06.12]).
- [3] 马九杰,陈俊良,赵永华. 直播电商价值链中数字守门人与农产品质量安全把守机制研究[J]. 管理世界, 2025, 41 (4): 175-192.
- [4] Chen H, Chen H T, Tian X X. The dual-process model of product information and habit in influencing consumers' purchase intention: the role of live streaming features[J]. Electronic Commerce Research and Applications, 2022, 53: 101150.
- [5] Wongkitrungrueng A, Dehouche N, Assarut N. Live streaming commerce from the sellers' perspective: implications for online relationship marketing[J]. Journal of Marketing Management, 2020, 36 (5/6): 488-518.

- [6] 漆亚林, 杨婧童. S-O-R 理论视域下情绪、感知对电商直播用户购买行为的中介效应研究[J]. 新闻界, 2024 (8): 55-65.
- [7] 刘玉奇, 张晓芸, 王强. 全民直播时代主播特性对消费者购买意愿影响研究[J]. 价格理论与实践, 2024, (4): 153-157, 226.
- [8] Li G M, Jiang Y, Chang L T. The influence mechanism of interaction quality in live streaming shopping on consumers' impulsive purchase intention[J]. *Frontiers in Psychology*, 2022, 13: 918196.
- [9] Chen Y J, Wu Z M, Wang Y L. Omnichannel product selection and shelf space planning optimization[J]. *Omega*, 2024, 127: 103074.
- [10] Tharwat A, Schenck W. Population initialization techniques for evolutionary algorithms for single-objective constrained optimization problems: deterministic vs. stochastic techniques[J]. *Swarm and Evolutionary Computation*, 2021, 67: 100952.
- [11] 赵渊, 刘祎, 任和, 等. 大电网可靠性评估的最优 ϵ 散度重要抽样法[J]. 中国电机工程学报, 2022, 42 (14): 5067-5078.
- [12] 邵一凡, 禹春霞, 禹嘉诚. 基于改进 NSGA-II 算法的云制造服务组合优化研究[J]. 运筹与管理, 2022, 31 (11): 30-36.
- [13] Xu W, Cao Y, Chen R Y. A multimodal analytics framework for product sales prediction with the reputation of anchors in live streaming e-commerce[J]. *Decision Support Systems*, 2024, 177: 114104.
- [14] Peng J, Zhang J H, Nie T F. Social influence and channel competition in the live-streaming market[J]. *Annals of Operations Research*, 2025, 344 (2): 617-645.
- [15] Shi B L, Sun Z M. Product screening in live stream: an equilibrium strategy analysis considering streamer product selection and platform inspection[J]. *Arabian Journal for Science and Engineering*, 2025, 50 (15): 12199-12217.
- [16] Xu Y F, Qi J Y, Kong J H, et al. Strategic decision making in live streaming e-commerce through tripartite evolutionary game analysis[J]. *PLoS One*, 2024, 19 (7): e0305427.
- [17] Deng F M, Chen R, Wu B G. The effect of product selection in live streaming on firm performance: the roles of interaction, scenario, and brand-consumer congruence[J]. *Journal of Business-to-Business Marketing*, 2025, 32 (4): 479-494.
- [18] Li K L, Yao S, Li Y J, et al. Deliberate shortage in live-streaming commerce[J]. *Omega*, 2025, 131: 103201.
- [19] Gustriansyah R, Ermatita E, Rini D P. An approach for sales forecasting[J]. *Expert Systems with Applications*, 2022, 207: 118043.
- [20] Maaouane M, Zouggar S, Krajačić G, et al. Modelling industry energy demand using multiple linear regression analysis based on consumed quantity of goods[J]. *Energy*, 2021, 225: 120270.
- [21] Ferbar Tratar L, Mojskerc B, Toman A. Demand forecasting with four-parameter exponential smoothing[J]. *International Journal of Production Economics*, 2016, 181: 162-173.
- [22] Ye J M, Xu Z S, Gou X J. An adaptive Grey-Markov model based on parameters self-optimization with application to passenger flow volume prediction[J]. *Expert Systems with Applications*, 2022, 202: 117302.
- [23] Xu W, Zhang X H, Chen R Y, et al. How do you say it matters? A multimodal analytics framework for product return prediction in live streaming e-commerce[J]. *Decision Support Systems*, 2023, 172: 113984.
- [24] Xu G, Ren M, Wang Z H, et al. MEMF: multi-entity multimodal fusion framework for sales prediction in live streaming commerce[J]. *Decision Support Systems*, 2024, 184: 114277.
- [25] Loureiro A L D, Miguéis V L, da Silva L F M. Exploring the use of deep neural networks for sales forecasting in fashion retail[J]. *Decision Support Systems*, 2018, 114: 81-93.
- [26] Joseph R V, Mohanty A, Tyagi S, et al. A hybrid deep learning framework with CNN and bi-directional LSTM for store item demand forecasting[J]. *Computers and Electrical Engineering*, 2022, 103: 108358.
- [27] Lin Q P, Jia N, Chen L, et al. A two-stage prediction model based on behavior mining in livestream e-commerce[J]. *Decision Support Systems*, 2023, 174: 114013.
- [28] 艾瑞咨询. 2020 年中国直播电商生态研究报告 [EB/OL]. <https://report.iresearch.cn/report/202006/3606.shtml> [2020-06-30].
- [29] Devlin J, Chang M W, Lee K, et al. BERT: pre-training of deep bidirectional transformers for language understanding[C]//

- Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers) . Stroudsburg: Association for Computational Linguistics, 2019: 4171-4186.
- [30] Cai G P, Wei X C, Li Y. Privacy-preserving CNN feature extraction and retrieval over medical images[J]. International Journal of Intelligent Systems, 2022, 37 (11): 9267-9289.
- [31] 刘兴旺, 王江晴, 徐科. 一种融合 AutoEncoder 与 CNN 的混合算法用于图像特征提取[J]. 计算机应用研究, 2017, 34 (12): 3839-3842, 3847.
- [32] 库向阳, 杨瑞丽, 董立红. 基于 Sword2vect 的中文在线商品评价情感分析[J]. 西安科技大学学报, 2020, 40 (3): 504-511.
- [33] Rita P, Oliveira T, Farisa A. The impact of e-service quality and customer satisfaction on customer behavior in online shopping[J]. Heliyon, 2019, 5 (10): e02690.
- [34] 吴国英, 闫建钢. “趋利” 还是 “避害”? : 直播电商退换货服务对消费者行为的影响研究[J]. 现代财经 (天津财经大学学报), 2021, 41 (12): 65-77.
- [35] Sullivan Y W, Kim D J. Assessing the effects of consumers' product evaluations and trust on repurchase intention in e-commerce environments[J]. International Journal of Information Management, 2018, 39: 199-219.
- [36] 易舒心. 跨境电商物流服务质量对顾客重复购买意向影响实证研究[J]. 商业经济研究, 2020, (17): 85-88.
- [37] Liu Y T, Sun Z W, Liang H M, et al. Ranking range model in multiple attribute decision making: a comparison of selected methods[J]. Computers & Industrial Engineering, 2021, 155: 107180.
- [38] Dong Y C, Liu Y T, Liang H M, et al. Strategic weight manipulation in multiple attribute decision making[J]. Omega, 2018, 75: 154-164.
- [39] 刘松林, 王坦, 戚琳琳. 基于客观组合赋权的就业质量测度与评价[J]. 统计与决策, 2023, 39 (20): 168-173.
- [40] Sharkasi N, Rezakhah S. A modified CRITIC with a reference point based on fuzzy logic and hamming distance[J]. Knowledge-Based Systems, 2022, 255: 109768.
- [41] Hajiaghaei-Keshteli M, Cenik Z, Erdebilli B, et al. Pythagorean fuzzy TOPSIS method for green supplier selection in the food industry[J]. Expert Systems with Applications, 2023, 224: 120036.
- [42] Zhang N, Zhou Y F, Liu J, et al. VIKOR method for Pythagorean hesitant fuzzy multi-attribute decision-making based on regret theory[J]. Engineering Applications of Artificial Intelligence, 2023, 126: 106857.
- [43] Zhao M W, Wei G W, Chen X D, et al. Intuitionistic fuzzy MABAC method based on cumulative prospect theory for multiple attribute group decision making[J]. International Journal of Intelligent Systems, 2021, 36 (11): 6337-6359.
- [44] Akram M, Ilyas F, Deveci M. Interval rough integrated SWARA-ELECTRE model: an application to machine tool remanufacturing[J]. Expert Systems with Applications, 2024, 238: 122067.
- [45] Zhao J H, Rong H Y. Multi-attribute decision-making based on data mining under a dynamic hybrid trust network[J]. Computers & Industrial Engineering, 2023, 185: 109672.
- [46] Dai W F, Zhong Q Y, Qi C Z. Multi-stage multi-attribute decision-making method based on the prospect theory and triangular fuzzy MULTIMOORA[J]. Soft Computing, 2020, 24 (13): 9429-9440.
- [47] Pramanik S, Mallick R. TODIM strategy for multi-attribute group decision making in trapezoidal neutrosophic number environment[J]. Complex & Intelligent Systems, 2019, 5 (4): 379-389.
- [48] Reimers I, Xie C. Do coupons expand or cannibalize revenue? Evidence from an e-market[J]. Management Science, 2019, 65 (1): 286-300.
- [49] 陈敏灵, 路子盈. 淘宝直播电商对消费者购买意愿影响研究[J]. 价格理论与实践, 2024, (6): 172-177.
- [50] 宋思根, 赵庆明. 明星主播对观众购买意愿的影响机制[J]. 中国流通经济, 2023, 37 (3): 61-71.
- [51] 范文芳, 王千. 个性化智能推荐对消费者在线冲动购买意愿的影响研究[J]. 管理评论, 2022, 34 (12): 146-156, 194.
- [52] Yu S S, Jiang Z X, Chen D D, et al. Leveraging tripartite interaction information from live stream E-commerce for improving product recommendation[C]//Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery &

- Data Mining. New York: Association for Computing Machinery, 2021: 3886-3894.
- [53] Wohn D Y, Freeman G. Audience management practices of live streamers on Twitch [C]//Proceedings of the 2020 ACM International Conference on Interactive Media Experiences. New York: Association for Computing Machinery, 2020: 106-116.
- [54] Yue X. Research on the influence of inconsistency between anchor image and product image on consumers' purchase intention[J]. Advances in Economics, Management and Political Sciences, 2023, 2 (1): 16-21.
- [55] Yang Q, Wang Z S, Feng K, et al. Investigating the crucial role of logistics service quality in customer satisfaction for fresh e-commerce: a mutually validating method based on SERVQUAL and service encounter theory[J]. Journal of Retailing and Consumer Services, 2024, 81: 103940.
- [56] Rashid A, Rasheed R. Logistics service quality and product satisfaction in e-commerce[J]. Sage Open, 2024, 14 (1): 1-12.
- [57] Wang L, Tang Y M, Chau K Y, et al. Empirical research of cold-chain logistics service quality in fresh product E-commerce[J]. Journal of Theoretical and Applied Electronic Commerce Research, 2024, 19 (3): 2543-2556.
- [58] Yang L, Zheng C, Hao C X. Optimal platform sales mode in live streaming commerce supply chains[J]. Electronic Commerce Research, 2024, 24 (2): 1017-1070.
- [59] Heger J, Klein R. Assortment optimization: a systematic literature review[J]. OR Spectrum, 2024, 46 (4): 1099-1161.

A Novel Streamer Historical Trading Data-driven Marketing Product Combination Determination Model

YU Chunxia¹, ZHANG Yanan¹, WANG Xiaohuan²

(1. School of Economics and Management, China University of Petroleum, Beijing, 102249, China;
2. School of Business Administration, Northeastern University, Shenyang, 110167, China)

Abstract With the rapid rise of livestreaming e-commerce, determining optimal product portfolios has become a critical issue for improving both sales performance and streamer influence. However, existing product selection methods often fail to comprehensively account for both product profitability and comprehensive performance. Such inappropriate selection may not only reduce sales volume but also negatively affect the streamer's public image. Therefore, this paper proposes a data-driven livestreaming product combination determination model based on a streamer's historical transaction data, aiming to improve decision-making quality and enhance livestreaming effectiveness. Firstly, a sales prediction index system is constructed, in which the Pearson correlation coefficient is used to assess the similarity between product and streamer audiences, and a Transformer model generates product sales predictions. Secondly, the CRITIC-MABAC method is applied to evaluate the comprehensive performance of candidate products. Thirdly, a multi-objective optimization model is formulated to simultaneously maximize profit and product performance, and solved via LHS-NSGA-II algorithm. A case study demonstrates the proposed model's effectiveness and advantages.

Key words Livestreaming, Product combination determination, Sales prediction, NSGA-II, Transformer

作者简介

禹春霞 (1983—), 女, 中国石油大学 (北京) 经济管理学院教授、博士生导师, 研究方向为智能决策与优化。E-mail: cxyu@cup.edu.cn。

张雅楠（2001—），女，中国石油大学（北京）经济管理学院博士生，研究方向为智能决策与优化。
E-mail: zhangyanan2255@163.com。

王晓欢（1983—），女，东北大学工商管理学院副教授、博士生导师，研究方向为物流与供应链管理、行为运作管理。E-mail: xhwang@mail.neu.edu.cn。