

分类号 _____
U D C _____

密级 _____
编号 10741 _____



硕士学位论文

(专业学位)

论文题目 基于交易数据的JDC集团跨境电商营销策略研究

研究生姓名: _____ 刘阳 _____

指导教师姓名、职称: _____ 牛胜强 教授 _____

学科、专业名称: _____ 统计学 应用统计硕士 _____

研究方向: _____ 市场研究 _____

提交日期: _____ 2021年6月6日 _____

独创性声明

本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果。与我一同工作的同志对本研究所做的任何贡献均已在文中作了明确的说明并表示了谢意。

学位论文作者签名： 刘阳 签字日期： 2021.6.6

导师签名： 牛志强 签字日期： 2021.6.6

导师(校外)签名： 张小弓 签字日期： 2021.6.6

关于论文使用授权的说明

本人完全了解学校关于保留、使用学位论文的各项规定， 同意（选择“同意”/“不同意”）以下事项：

1. 学校有权保留本论文的复印件和磁盘，允许论文被查阅和借阅，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文；

2. 学校有权将本人的学位论文提交至清华大学“中国学术期刊（光盘版）电子杂志社”用于出版和编入 CNKI《中国知识资源总库》或其他同类数据库，传播本学位论文的全部或部分内容。

学位论文作者签名： 刘阳 签字日期： 2021.6.6

导师签名： 牛志强 签字日期： 2021.6.6

导师(校外)签名： 张小弓 签字日期： 2021.6.6

**Research on cross-border E-commerce
Marketing Strategy of JDC Group based on
transaction data**

Candidate : Liu Yang

Supervisor: Niu Shengqiang

摘 要

随着互联网与大数据技术愈益成熟以及我国对外贸易广度和深度不断加强,再加上电子商务与国际贸易活动深度融合,跨境电子商务应运而生并快速发展。统计资料显示,2009年我国跨境电子商务交易总额仅为0.9万亿元,2014年上升为4.2万亿元,而到2019年其交易规模已突破10万亿元,同时,跨境电子商务的迅速发展使其占我国进出口总额的比重也从2009年的6%提高至2014年的15.9%,到2019年其占比已高达33%。可以看出,在我国传统外贸增速放缓的背景下,跨境电子商务作为电子商务的最新发展领域,不仅保持了快速增长态势,而且已成为我国对外贸易发展的重要增长点。然而,随着逆全球化浪潮不断升级,跨境电子商务行业竞争也愈发激烈,在面临行业标准有待完善、政策法规体系不健全、配套服务不成熟、专业人才匮乏、跨境物流运输效率低等诸多发展困境的同时,贸易成本不断增加以及消费市场细分要求进一步提高也使其持续快速发展受到极大制约。因而,在此背景下,为进一步提高我国跨境电子商务营销水平和市场占有率,本文以JDC集团跨境电子商务企业为例,从用户视角出发,充分利用JDC集团跨境电子商务平台交易数据,按照理论研究——现状分析——目标用户定位——预测模型构建——营销建议提出的研究思路,对JDC集团跨境电子商务平台营销策略进行探究。文章首先对数据库营销和跨境电子商务理论研究成果进行阐述,并基于平台基础指标数据对JDC集团跨境电子商务平台发展现状进行分析总结,接着依据现状分析结果进一步归纳平台在产品、渠道、客诉及商家管理、数据四个层面存在的主要问题,然后从用户基本特征、用户价值、用户偏好三个角度出发,全面探究平台目标用户的精准化定位,在总结平台目标用户特征的基础上,得出目前平台用户忠诚度偏低、消费能力处于中下等水平、低价值用户占比大以及不同性别、年龄、地区的用户偏好差异较大等结论,进而以GBDT机器学习算法概述及模型效果评价指标说明为基础,依照数据预处理——样本选取——特征变量构建与选择——模型训练的流程构建用户购买行为预测模型,并得到F1值为92.1%、AUC的值为0.965,模型预测效果较好的结论。最后根据前文分析结果首先从巩固平台已有用户视角提出品类优化、品牌推广、用户个性化推荐、跨境物流服务等方面的营销建议;接着又从拓展新用户视角以产品、市场、运营三方面为出发点提出平台拉新引流的相应策略。

关键词: 对外贸易 跨境电子商务 数据库营销 用户 GBDT 算法

Abstract

As the Internet and big data technology become more and more mature, the breadth and depth of China's foreign trade continue to strengthen, coupled with the in-depth integration of e-commerce and international trade activities, cross-border e-commerce arises at the right moment and develops rapidly. Statistics show that the total amount of cross-border e-commerce transactions in China was only 0.9 trillion yuan in 2009, rising to 4.2 trillion yuan in 2014, and by 2019, the transaction size has exceeded 10 trillion yuan. Meanwhile, The rapid development of cross-border e-commerce has increased its share of China's total import and export volume from 6% in 2009 to 15.9% in 2014 and up to 33% in 2019. It can be seen that in the context of China's traditional foreign trade growth slowing down, cross-border e-commerce, as the latest development field of e-commerce, has not only maintained a rapid growth trend, but also become an important growth point of China's foreign trade development. However, as the tide of anti-globalization continues to escalate, the competition in the cross-border e-commerce industry is becoming increasingly fierce. In the face of many development difficulties, such as unperfected industrial standards, imperfect policies and regulations, immature supporting services, lack of professionals, and low efficiency of cross-border logistics and transportation, The continuous increase of trade costs and the further improvement of consumer market segmentation requirements also make its sustained and rapid development is greatly restricted. Therefore, in this context, in order to further

improve the level of cross-border e-commerce marketing and market share in China, this paper takes the cross-border e-commerce enterprise of JDC Group as an example, and makes full use of the transaction data of the cross-border e-commerce platform of JDC Group from the perspective of users. According to the research ideas of theoretical research, current situation analysis, target user positioning, prediction model construction and marketing suggestions, the marketing strategy of JDC cross-border e-commerce platform is explored. This paper on database marketing and cross-border e-commerce elaborates the theoretical research results, and data index is based on the platform of JDC group cross-border e-commerce platform development present situation carries on the analysis summary, then according to the status quo analysis results further induces platform in the product, channel, customer complaint and business management, main problems of data of four levels, Basic features from a user, the user value, then the user preference three perspective, comprehensive inquiry platform target users precise positioning, on the basis of summarizing the platform target user characteristics, it is concluded that the current platform user loyalty is low, consumption ability is in the lowest level, low value accounted for big users and user preferences of different gender, age, area differences such as conclusion, Then, based on the GBDT machine learning algorithm outline and the model effect evaluation index description, according to the process of data preprocessing -- sample selection -- characteristic variable construction and selection -- model training, the user purchase behavior prediction model is built,

and the F1 value is 92.1%, the AUC value is 0.965, the model prediction effect is good. Finally, according to the above analysis results, the marketing suggestions on category optimization, brand promotion, personalized recommendation of users, cross-border logistics services and other aspects are proposed from the perspective of consolidating the existing users of the platform. Then from the perspective of expanding new users from the product, market, operation as the starting point of three aspects of the platform to draw new drainage of the corresponding strategy.

Keywords : The ministry of foreign trade; Cross-border electricity; Precision marketing; Users; GBDT algorithm

目 录

1 绪论	1
1.1 研究背景及意义.....	1
1.1.1 研究背景.....	1
1.1.2 研究目的.....	2
1.1.3 研究意义.....	2
1.2 文献综述.....	3
1.2.1 国外研究现状.....	3
1.2.2 国内研究现状.....	4
1.2.3 文献述评.....	6
1.3 主要研究内容、方法及创新点.....	6
1.3.1 主要研究内容.....	6
1.3.2 主要研究方法.....	8
1.3.3 主要创新点.....	8
2 数据库营销概述及我国跨境电子商务发展概况	9
2.1 关于市场营销.....	9
2.1.1 市场营销的概念.....	9
2.1.2 市场营销理念演变过程.....	9
2.1.3 市场营销领域发展新趋势.....	10
2.2 数据库营销概述.....	11
2.2.1 数据库营销概念.....	11
2.2.2 数据库营销优点.....	11
2.2.3 数据库营销的运作程序.....	12
2.3 我国跨境电子商务发展概况.....	13
2.3.1 跨境电子商务概念.....	13
2.3.2 我国跨境电子商务发展现状.....	13
2.3.3 跨境电子商务对我国对外贸易的影响.....	16
3 JDC 集团跨境电商营销现状概述	17

3.1 JDC 集团跨境电商平台经营现状分析.....	17
3.2 JDC 集团跨境电商平台存在的主要问题.....	19
3.2.1 产品层面.....	19
3.2.2 渠道层面.....	20
3.2.3 客诉及商家管理层面.....	21
3.2.4 数据层面.....	21
4 立足于交易数据的目标用户精准化定位.....	22
4.1 用户基本特征分析.....	22
4.1.1 用户人口属性.....	22
4.1.2 用户消费习惯.....	25
4.2 用户价值分析.....	28
4.2.1 用户忠诚度及消费能力.....	28
4.2.2 用户复购、留存及新老用户占比.....	32
4.3 用户消费偏好分析.....	35
4.3.1 用户商品品类偏好.....	36
4.3.2 用户品牌偏好.....	40
4.3.3 用户页面点击渠道偏好.....	43
4.4 本章小结.....	46
5 基于机器学习算法的预测模型构建.....	47
5.1 GBDT 模型.....	47
5.1.1 GBDT 算法概述.....	47
5.1.2 GBDT 模型的参数说明.....	49
5.2 模型效果分析指标.....	49
5.2.1 混淆矩阵.....	49
5.2.2 ROC 曲线、AUC 指标.....	51
5.3 模型构建流程.....	52
5.3.1 数据来源及预处理.....	52
5.3.2 样本选取.....	53
5.3.3 特征变量构建及选择.....	53
5.3.4 预测模型构建.....	57
5.4 本章小结.....	59

6 JDC 集团跨境电子商务营销策略提出	60
6.1 巩固平台已有用户视角.....	60
6.2 拓展新用户视角.....	62
7 总结与展望	65
7.1 总结.....	65
7.2 展望.....	65
参考文献	67
后记	70

1 绪论

1.1 研究背景及意义

1.1.1 研究背景

2008 年金融危机爆发以后，全球经济出现了衰退征兆，欧美等发达国家市场有效需求增速下滑，加之原材料价格持续上涨、劳动力和土地成本不断增加以及对外贸易摩擦频繁发生，使得我国传统外贸导向型企业受到很大程度冲击，经济增速也明显放缓^[10]。为突破传统贸易局限和顺应经济全球化发展趋势，跨境电子商务应运而生并迅速发展。跨境电子商务不仅有效实现了网络信息技术与国际贸易活动的深度融合，使对外贸易模式得到进一步转型升级与发展，而且能将厂商与用户通过电商平台直接联系起来，既有效缩短了产品从厂家到全球消费者的空间距离，还降低了中小企业开展跨境贸易的门槛及成本，使“买全球、卖全球”成为可能，对于推动我国对外贸易发展及经济发展方式转型都具有重大意义。此外，在国家一系列利好政策的相继出台、国内电子商务体系的日渐完善、外贸企业家的积极推动以及“中国制造”的优势地位等有利条件的共同作用下，我国跨境电子商务进出口平台迎来了发展新机遇，并在近几年内以迅猛的势头快速兴起。

目前，我国已有一定数量规模的跨境电子商务企业致力于构建跨境进口平台为国内用户直供海外商品，满足国人不出国门购买全球商品的消费需求，如天猫国际、网易考拉海购、洋码头等，该类平台专注于为中国用户打造“一站式”海外购物体验，尤其在 2018 年我国举办中国国际进口博览会后，为响应国家坚定支持贸易自由化和经济全球化的立场，各平台极力满足用户不断增长的多样化需求，持续加大进口平台运营的各方面投入，扩大经营规模。因此，我国当前跨境电子商务进口平台发展已基本能够满足国内消费者“足不出户，淘遍全球”的购买需求。另外，随着近年来全球消费者对“中国制造”的市场需求与日俱增、贸易市场环境明显改善以及受国家利好政策大力扶持，我国出口型跨境电子商务迈入快速发展阶段，其市场规模也远超于进口。统计资料显示，2015—2019 年我国跨境电子商务出口交易额占跨境电子商务总交易额的比重分别为 83.33%、82.09%、78.16%、78.89%、76.48%，占比均达到八成左右，可以看出，在我国跨境电子商务交易中，出口型跨境电子商务占有主导地位，并持续拉动我国经济增长，

同时,该领域呈现出的良好态势也给国内出口型跨境电子商务平台带来发展机会^[15]。然而,尽管在跨境电子商务浪潮的持续推动下我国跨境电子商务平台数量规模得到迅速扩张,但实际上大多数企业致力于卡位市场,并未对企业自身定位与营销模式有清晰认知,导致各平台缺乏特色,服务大同小异,从而引发同质化竞争下的“商品促销战”,使用户的平台体验感大打折扣,增加其流失风险。基于此,本文以主要面向海外消费客群的 JDC 集团跨境电子商务平台为例,充分利用平台交易数据,从目标客群定位和用户购买行为预测模型构建两方面入手重点展开研究,旨在为该集团提供相应营销建议以提高其跨境电子商务平台知名度及企业核心竞争力。

1.1.2 研究目的

就国内跨境电子商务平台的发展现状来看,目前同步开展国内零售与跨境业务的平台相对较少,该现状于 JDC 集团发展而言是机遇也是挑战。因此,JDC 集团跨境电子商务平台如何背靠集团实力在众多跨境电子商务平台中打响名声,提高知名度,并成为跨境电子商务领域具有强劲实力的竞争者是集团亟需考虑的关键问题。基于此,本文研究目的主要有以下三点:

1、通过梳理 JDC 集团跨境电子商务平台发展现状,提出其各层面的现存问题,并基于分析结果为平台提供营销新思路。

2、利用数据深入探究用户特征,挖掘用户价值并洞察用户各层面消费偏好及需求,从而使 JDC 集团跨境电子商务平台能真正把握用户动向,不断提高平台用户与产品之间的匹配度,推动平台实现目标用户定位精准化、个性化页面推送高效率的营销效果^[16]。同时,随着大数据技术快速发展,还要不断助力企业在经营过程中由凭借“经验主义”进行决策向充分利用数据开展营销工作转变,进而减少平台运营成本,提升用户体验,累积发展优势。

3、在对目标用户进行精准定位的基础上,整合机器学习算法,构建用户购买行为预测模型,以便进一步把握用户购买意愿。最终依据分析结论为 JDC 集团跨境电子商务平台提出相应营销策略建议,助力平台获得持久竞争优势。

1.1.3 研究意义

1、理论意义

本文先是通过深入分析 JDC 集团跨境电子商务平台用户资料,准确把握平台用户偏

好、价值等消费特征，随后，又在深入分析平台用户资料的基础上，利用平台交易数据和机器学习算法，构建用户购买行为预测模型，进一步探究用户是否会产生购买行为，使平台侧全面了解用户购买意愿。这种通过平台用户定位与购买行为预测模型构建以掌握用户消费情况从而进行营销策略制定的研究方法对其他跨境电子商务平台具有理论意义和参考价值。

2、现实意义

本文研究的现实意义主要包括以下三个方面：

(1) 从用户角度出发，充分利用平台交易数据，并基于分析结果助力 JDC 集团跨境电子商务平台精准定位目标用户以有效提高用户与产品之间的匹配度，实现低成本、高效率营销。

(2) 在全面把握平台目标用户特征的基础上，进一步对其购买行为展开预测并向不同类型用户进行产品和信息的精准化推送，通过在平台和用户之间建立一种不间断联系，持续增强用户的平台信任感，提高其购买转化。

(3) 本文提出的对策建议对 JDC 集团跨境电子商务平台进一步提高对外营销水平能起到参考、借鉴、启发作用，同时对其他面向国际市场的跨境电子商务平台也具有参考和借鉴价值。

1.2 文献综述

1.2.1 国外研究现状

国外学者对跨境电子商务领域的研究起步较早，并将研究重点主要集中在以下两方面：一，电子商务与国际贸易之间的关系与相互影响。A.R.Kemal（1998）早在 1998 年便分析了电子商务在国际贸易中起到的作用，认为率先开启电子商务相关业务的国家在国际贸易中占据的份额将会增加，并以巴基斯坦为例，提出该国家出口停滞的主要原因为未能探索新市场、使出口业务多样化^[1]。Nuray Terzi（2011）分析了在国际贸易和就业层面电子商务产生的影响，并指出从长远的发展角度看，发展中国家相比于发达国家来说将受益更多^[8]。Erdal Aydın 等人（2014）探究了全球化与电子商务之间的关系，认为全球化在为电信和运输业带来蓬勃发展的同时，也创造了通过电子商务进入市场的新途径^[4]。Mesut Savrul（2014）认为经济全球化虽然使中小企业面临严峻考验，但其可以通过电子商务从新环境中获益，并分析中小企业在不断变化的商业环境下将面临的阻碍

及发展跨境电子商务的潜力,提出全球化的电子商务平台将会弥补中小企业规模小、机遇少等劣势的结论^[7]。二,跨境电子商务的影响因素及应用。Estrella Gomez-Herrera 等人(2014)以欧盟为例,探究跨境电子商务的驱动和影响因素,并通过研究发现跨境电子商务相比于线下交易而言降低了贸易距离成本,但增加了语言、物流运输及线上支付等新成本。最终,依据分析结果提出跨境支付系统的使用效率每提高 1%,跨境电子商务整体效率则增加 7%的结论^[5]。Guy H.Gessner 等人(2015)探讨了加拿大和美国这一世界上活跃度较高的双边贸易市场进行跨境电子商务交易的潜力,并指出跨境电子商务的快速发展为全球中小型企业进入国际市场提供契机,但不同国家在海关、关税制度等方面的差异以及贸易摩擦也会成为阻碍其发展的主要因素^[6]。Bin Hu 等人(2018)基于“互联网+”背景对跨境电子商务进行研究,并分析跨境电子商务对国际贸易交易成本的影响,提出“互联网+”背景下跨境电子商务发展的营销策略^[3]。Ángel Valarezo(2018)以西班牙为例,深入剖析基于个人行为的跨境电子商务驱动及阻碍因素^[2]。

1.2.2 国内研究现状

近年来,随着网络经济和进出口贸易的蓬勃发展,国家及地方政府对跨境电子商务行业的重视程度与扶持力度日益增加,并在一系列政策的出台下逐步实现了跨境电子商务监管到位、界定清晰的局面。与此同时,国内学者对跨境电子商务的相关研究也日渐丰富,本节将从跨境电子商务发展现状、跨境电子商务对国际贸易的影响、跨境电子商务营销角度三方面对国内研究现状进行综述。

1、跨境电子商务发展现状

廖蓁,王明宇(2014)从发展潜力、市场结构、运营模式三个视角对跨境电子商务发展现状展开剖析,并根据现状分析结果进一步归纳跨境电子商务在消费者习惯、交易安全、支付、物流层面存在的主要问题及相应解决措施,最终得出我国跨境电子商务虽处于初步发展阶段,但该对外贸易新趋势给电商行业发展带来新方向,有巨大发展前景等结论^[24]。鄂立彬,黄永稳(2014)首先分别对全球跨境电子商务和我国跨境电子商务发展现状进行概述,并得出全球跨境电子商务发展具有传统强国仍占据主导地位、新兴市场迅速扩张、跨境电子商务市场前景广阔等特征,而我国跨境电子商务行业则呈现出发展势头迅猛、与传统贸易对比鲜明、出口目标市场和企业地区分布集中等特点,接着对跨境电子商务与传统贸易模式展开对比研究,说明二者在交易环节、产业链、成本、产品等层面的差异性,最后又从政府和企业两视角入手,提出跨境电子商务发展的对策

建议^[10]。关怀庆（2016）在对近几年跨境电子商务发展趋势进行回顾的基础上，结合我国现下政治环境对该领域市场前景进行展望，并就交易规模、政策环境、发展机遇层面的现状分析引申出我国跨境电子商务行业面临的主要困境，进而提出相应对策建议^[14]。孙琪（2020）从市场规模、行业背景、市场领域拓展等角度分析我国跨境电子商务发展现状，并总结跨境电子商务在信息技术、产业基础、贸易服务等层面正面临的诸多挑战，最终分别从政治因素、服务水平、品牌建设、人才培养等方面提出我国跨境电子商务在未来发展的新思路^[27]。

2、跨境电子商务对国际贸易的影响

应小凡（2011）就电子商务在国际贸易中的应用展开研究，并提出电子商务能够为国际贸易带来高效率贸易模式、智能化管理、打破贸易时空限制、降低外贸成本、简化营销环节等发展优势^[34]。樊文静（2015）认为跨境电子商务正通过市场范围、主体结构、生产方式、贸易成本等方面的影响推动对外贸易模式转型，并从宏观、中观、微观三个层面展开跨境电子商务对我国贸易模式影响的机理分析^[12]。赵旭明，杨晓涵（2016）基于我国跨境电子商务研究背景及发展现状概述，从对外贸易增长、企业经营管理、服务和信息贸易发展三个角度阐述跨境电子商务对我国开展对外贸易产生的积极影响^[36]。李柏杏，潘开灵（2016）以跨境电子商务与对外贸易的相互影响及对比分析入手，构建古诺双寡头模型，对两者之间的关联程度展开深入探究，最终得出我国跨境电子商务与对外贸易之间存在长期正向影响及相互制衡关系的结论^[23]。

3、跨境电子商务营销角度

罗扬（2017）基于企业视角，根据 4Ps 营销组合理论，从价格、产品、促销、渠道四方面对跨境电子商务营销策略展开研究^[21]。吴娉娉（2016）以几大具有代表性的跨境电子商务平台为例，研究跨境电子商务企业的选品策略，并提出选品过程不仅要熟知平台规则及产品相关的法律法规为前提，同时还需通过多渠道展开调查，在试错中选品，逐步建立起稳定的供应商关系^[31]。方佳虹（2016）在总结跨境电子商务产品定价问题的基础上，又进一步提出跨境电子商务企业在价格制定上应充分考虑企业成本、市场需求、平台竞争、目标国家政策共四大因素，并以市场为导向，制定有策略、有创新的定价方案^[11]。刘永焕（2019）以“一带一路”战略为背景，分析跨境电子商务在定价、售后服务、出口通关三个层面出现的营销问题，并就这三方面问题提出优化定价策略、优化产品售后服务、优化出口通关、构建信息共享平台等营销策略^[22]。

1.2.3 文献述评

通过对国内外文献的梳理可知，国外研究者对跨境电子商务的研究起步较早，但多从电子商务与国际贸易之间的关系与相互影响、跨境电子商务的影响因素及应用两方面进行探究，在跨境电子商务营销策略方面的研究成果较少。而根据国内研究现状可知，由于我国跨境电子商务正以飞快的速度蓬勃发展，从而使国内有关跨境电子商务的研究愈加丰富。国内研究者主要从跨境电子商务发展现状、跨境电子商务对国际贸易的影响、跨境电子商务营销角度三方面展开研究，其中，对跨境电子商务营销策略的研究大多从跨境电子商务的现状分析、制约因素及对策建议展开或是以第三方平台为切入点，依据平台市场需求及定位总结其现存问题，并进一步提出相应营销策略。总体来看，学术界在跨境电子商务发展领域的研究重心大多向交易主体、贸易成本等方面倾斜，而较少涉及第三方平台的营销策略研究，并且对于第三方平台如何在对外贸易浪潮中基于平台交易数据开展外贸业务、壮大跨境业务交易规模的相关文献更是十分稀缺。因此，通过对国内外研究现状的综述与分析，本文研究思路可总结为：文章以数据库营销和跨境电子商务理论概述以及 JDC 集团跨境电子商务平台营销现状分析为基础，从以下两方面重点展开研究：一方面，对用户特征、用户价值、用户偏好深入剖析，精准定位平台目标客群；另一方面，充分利用平台交易数据，整合机器学习算法，构建购买行为预测模型，进一步把握用户购买意愿，最终从巩固平台已有用户及扩展新用户规模两个视角为 JDC 集团跨境电子商务平台提出营销策略建议，并为该方向研究略作贡献。

1.3 主要研究内容、方法及创新点

1.3.1 主要研究内容

1、具体研究内容

文章主要从七部分展开，具体内容如下：

第一章，绪论。该部分简要概述本文的研究背景、目的、意义、内容、方法、国内外研究现状及主要创新点，确定文章研究方向。

第二章，数据库营销概述及我国跨境电子商务发展概况。该部分内容从数据库营销和跨境电商两方面对本文相关理论进行阐述，夯实文章理论基础。

第三章，JDC 集团跨境电商营销现状概述。在本章，首先选取基础指标对 JDC 集

团跨境电子商务平台整体营销现状展开分析，并基于现状分析结果进一步从产品、渠道、客诉及商家管理、数据四个层面引申出 JDC 集团跨境电子商务平台现存的主要问题。

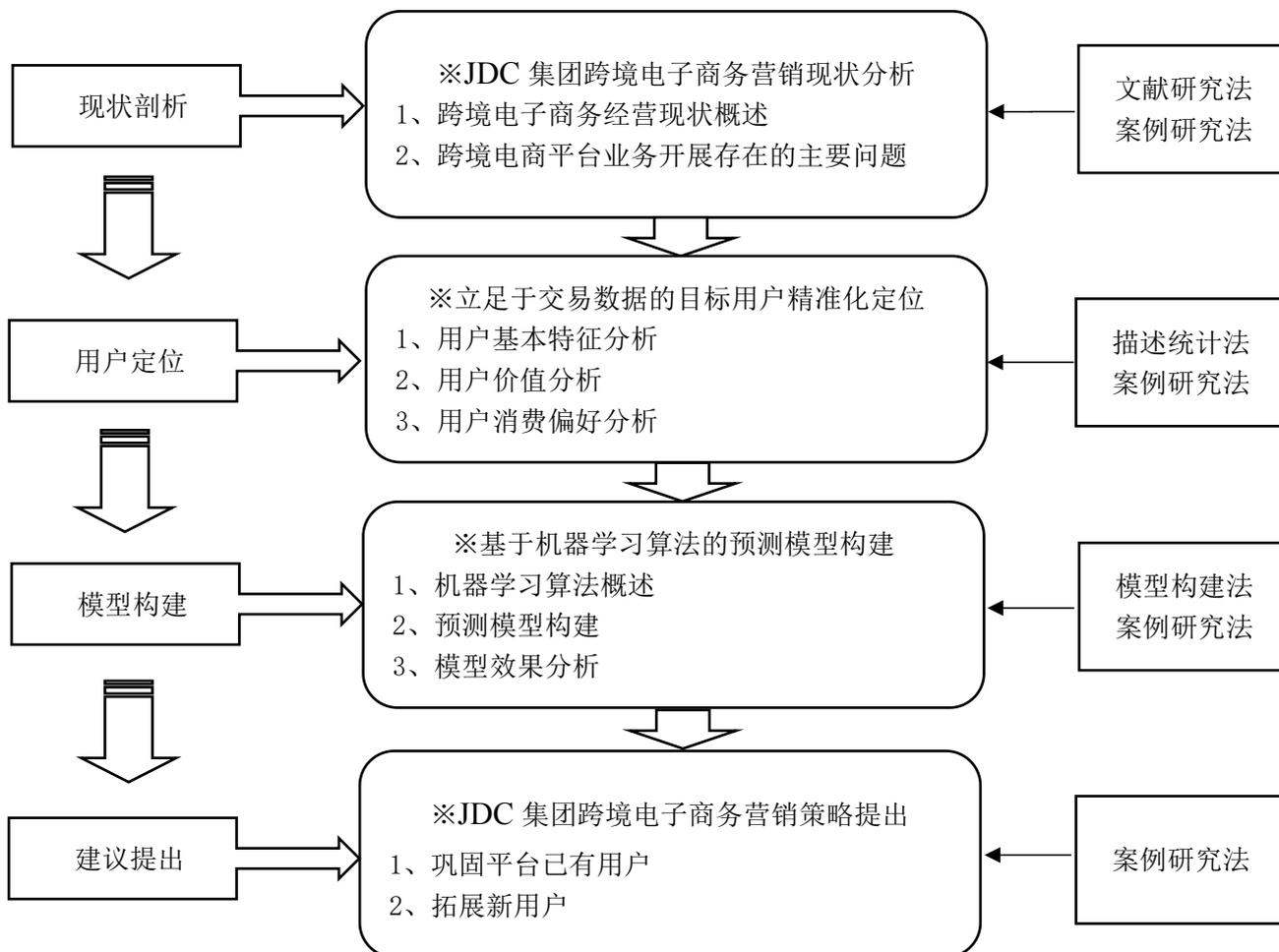
第四章，立足于交易数据的目标用户精准化定位。为了全方位洞察用户特征与购买需求以掌握平台用户消费情况，该部分内容从用户基本特征、用户价值以及用户消费偏好三方面展开研究，并归纳总结平台目标客群的主要消费特点。

第五章，基于机器学习算法的预测模型构建。该部分内容以 GBDT 机器学习算法概述及模型效果评价指标说明为基础，按照数据预处理——样本选取——特征变量的构造及选择——模型训练的流程构建用户购买行为预测模型。

第六章，JDC 集团跨境电商平台营销策略提出。该章节基于现状分析、目标用户定位以及预测模型构建的分析结果，从巩固平台已有用户和拓展新用户两个视角分别为 JDC 集团跨境电子商务平台提出相应营销策略建议。

第七章，总结与展望。该章节一方面对全文开展的工作进行梳理总结，另一方面对未来可以从哪些方面深入探究 JDC 集团跨境电子商务平台发展进行展望。

2、研究技术流程图



1.3.2 主要研究方法

1、文献研究法：通过查阅相关研究成果和文献资料，大致了解并把握文章要着手研究的主要问题。本文利用文献研究法通过对国内外研究成果的梳理与总结，确定文章研究方向及思路。

2、描述统计法：是指通过图表或表格的形式对数据资料展开剖析，并对数据的分布状态、数字特征以及随机变量之间关系进行估计和描述的统计方法。本文主要将描述性统计法运用于 JDC 集团跨境电子商务平台营销现状分析及目标用户定位两部分。

3、案例分析法：该方法是针对某一特定研究对象在较长时间内对某个研究主题的产生及发展全过程加以探究，并归纳总结目前有待解决的主要问题，最终得出相应可靠结论的研究方法。本文为探究 JDC 集团跨境电子商务平台营销策略，以集团跨境电子商务平台为特定研究对象，按照理论概述——现状分析——目标用户定位——购买行为预测模型构建——营销建议提出的思路展开研究。

4、模型构建法：该方法是指基于一定数学公式将某些事物的客观规律表示出来，并将其进行量化、模型化，从而解决相应实际问题。本文在第五章预测模型构建部分运用该方法，将选取的平台交易数据，基于 GBDT 机器学习算法，从数据预处理、样本选取、特征变量的构造与选择以及模型训练四个环节构建预测模型，进而根据分析结论为 JDC 集团提出相应对外营销建议。

1.3.3 主要创新点

本文的主要创新点在于：（1）从应用层面推动跨境电子商务的理论研究走向深化。目前，国内有关跨境电子商务的研究多是围绕跨境电子商务对我国对外贸易产生的影响以及跨境电子商务发展现状、主要困境和对策建议展开，但针对某个具体平台将目标用户定位与预测模型构建深度融合进行平台营销策略制定的个案研究却较为少见。本文选定 JDC 集团跨境电子商务这一特定研究对象，从平台用户出发，将目标用户精准化定位与用户购买行为预测模型构建相结合，并依据分析结论提出平台开展对外营销的策略建议。（2）将论文研究与实习工作深度融合。由于在实习单位工作期间不仅能更为全面且深入了解平台日常的业务开展，同时，还能为本文实证分析部分的研究提供数据支撑，确保研究成果兼具科学性、系统性和应用性。另外，本文的研究思路、方法及提出的营销策略建议也能为平台提供相应参考。

2 数据库营销概述及我国跨境电子商务发展概况

2.1 关于市场营销

2.1.1 市场营销的概念

市场营销是指以满足人民日益增长的美好生活需要为目的,通过市场变潜在交换为现实交换的一系列活动和过程,即企业为顾客创造价值并建立牢固的顾客关系,进而从顾客那里获得价值回报的过程^[13]。现代企业市场营销有别于销售,其营销活动包括市场营销研究、需求预测、新产品开发、定价、分销、物流、广告、公关、销售促进、售后服务等在内的一系列环节,销售仅是其中一个组成部分。市场营销全过程主要包括了解市场需求、设计用户驱动的营销策略、制定传递价值的营销方案、建立有收益回报的顾客关系、创造利润和顾客资产共五个步骤,其中,前四个环节是为顾客创造价值并与其建立关系的过程,最后一步则是企业从顾客身上获取价值回报的过程。另外,整个营销过程由供应商、企业(营销者)、竞争者、营销中介、最终用户共五个参与者组成,其具体流程为,首先,供应商为企业和竞争者提供生产性资源;接着,企业和竞争者再将产品信息直接或间接地传递给用户;最后,用户产生的价值回报又反馈给企业或竞争者。

2.1.2 市场营销理念演变过程

在所有营销理念中,人们习惯上将生产理念、产品理念和推销理念统称为传统营销理念。其中,生产理念是一种重生产、轻营销的商业理念;产品理念认为企业应重点关注附加值较高的产品生产,因而常将注意力放在产品和服务质量上,看不到市场需求变化,致使企业经营陷入困境;推销理念则认为企业须大力推销本企业产品以刺激消费者产生大量购买,该理念较为注重卖方需要,主要考虑如何变产品为现金^[29]。总之,传统营销理念以产品和企业利润为出发点,并没有充分考虑市场需求,在发展过程中逐渐暴露出诸多问题。而现代营销是一种以顾客需求为中心,且在目标市场设法发掘并满足顾客消费需求以实现企业目标的营销观念。实际上,与传统营销不同的是,现代营销理念更为注重买方需求,且将发展重心由产品转向消费者,并主要考虑如何向目标市场有效传送其所需产品或服务以满足市场及客户需要。从本质上说,现代营销理念是一种以顾客需要和欲望为导向的理念。综上所述,市场营销的演化路径及发展趋势可总结如下:

市场营销 1.0 是工业化时代以产品为导向的市场营销；2.0 是以消费者为导向的市场营销；3.0 是以价值观驱动、以人为本的市场营销；4.0 则是以大数据、社群、价值观营销为基础，企业将营销中心转移到与消费者积极互动、尊重消费者作为主体的价值观，让消费者更多地参与营销价创造的市场营销^[17]。

2.1.3 市场营销领域发展新趋势

随着全球化与逆全球化的角力，社会对企业营销道德的更高要求以及网络信息技术的愈发成熟，市场营销领域逐渐展露出发展新视野并随之产生新的营销策略，如，国际市场营销、绿色营销、社会营销、品牌营销、文化营销、数据库营销等。本节以全球化背景、来自营销道德和社会责任要求以及数字时代的市场营销领域发展新趋势为例展开详细介绍。

1、全球化背景下的市场营销

随着国际间越发频繁的合作往来以及改革开放的持续深入，我国企业营销发展逐渐走向国际化道路。全球化背景下的营销发展新趋势使得世界各地企业都或多或少参与到国际贸易及全球竞争中，并以全球化视角看待公司所属行业、竞争者及发展机会。而跨国公司延伸至世界各地的触角也给发展中国家的中小企业带来新机遇，使其除了面临海外竞争之外，也能抓住在本地获取国际订单的机会。

2、来自营销道德和社会责任要求的市场营销

来自营销道德和社会责任的要求使得营销者需要为社会和环境肩负起更大责任，为此，越来越多企业开始注重提升消费者福祉，而消费者福祉是指与消费品相关的处于消费生命周期不同阶段中主观幸福和客观幸福的一种理想状态。从消费者福祉视角入手，就要求企业除了要给消费者带来满意的购物体验之外，还要极力改善整个消费过程中各环节服务水平，真正为提高消费者生活水平、提升消费者购物体验、增进人类福祉做贡献。

3、数字时代的市场营销

新兴技术创造了新数字时代，计算机、信息、网络、交通和其他连接技术领域的快速发展对企业营销模式的优化升级产生了极大影响。而数字时代的到来也使得企业营销者实现了数字时代的营销，即：利用计算机工具及相关网络技术开发相应数据库系统，并利用数据精准定位每位顾客，为其提供专属的产品服务以满足其个性化购买需求，从

而达到高效率、低成本的营销效果。另外，新兴数字信息技术的蓬勃发展也将云营销、数据库营销、网络营销、社交媒体营销等全新营销模式呈现在营销者面前，并为其提供发展新思路。鉴于本文利用 JDC 集团跨境电子商务平台交易数据展开营销策略研究，故下文将对数字时代的市场营销领域发展新趋势——数据库营销展开详细介绍。

2.2 数据库营销概述

2.2.1 数据库营销概念

随着营销模式在数字时代的不断发展及各类型海量数据的有效积累，传统营销模式逐渐暴露出凭借经验开展决策、无法有效利用数据、营销过程的单向沟通致使个性化推荐效率低、推送不精准从而降低用户体验感等问题^[19]。总之，传统营销在未来必将被取代的主要原因是在该经营模式下，企业没有以市场需求为出发点，并且数据的潜在价值也得不到充分挖掘，加之买方市场不断壮大使得营销理论的研究中心也相应由产品转向用户，因而从用户入手，基于数据开展精准化营销十分关键，数据库营销也由此应运而生。

数据库营销是指企业通过搜集和积累大量用户信息，经过数据处理后预测消费者购买某种产品的可能性，进而借助数据精准定位该产品，并有针对性地传播营销信息以达到说服消费者购买产品和帮助企业制定精准化营销战略的目的^[28]。以活动推广为例，简要说明数据库营销的优势及作用。在往常，企业一般以高浏览量平台为首选推广渠道，采用“广撒网”的活动投放方式，但如今，大数据技术能够使企业了解受众群体偏好信息，即实现不同用户在关注同一媒体的相同界面时，推送内容会根据其不同偏好呈现出差异性，从而打破平台选择限制，达到高效率、精准化投放效果。应该说，数据库营销是对传统营销模式的补充和完善，它需要经过系统的用户数据资料收集、储存和有效分析以准确进行市场细分和定位，进而实施个性化营销。

2.2.2 数据库营销优点

数据库营销通过数据库的建立和分析能够帮助企业准确了解用户信息，定位企业目标客群并有针对性的开展营销工作，在降低营销成本的同时，提升营销效果。一般来说，数据库营销优点有以下四方面：（1）降低营销成本，提高营销效率。由于数据库营销实

现了对数据的充分利用,从而使平台能够精准识别具有相同偏好的目标客群,企业也因此可以进行有针对性的推广工作,避免低效宣传导致的资源浪费,进而达到降低营销成本、提高营销效率的目的。(2)增强顾客平台信任感,建立企业与顾客之间的不间断联系。通过数据库营销不仅能使企业与消费者保持频繁的双向沟通,而且还能强化两者之间的感情纽带,削弱外部竞争干扰。同时,企业在根据数据库信息全面剖析顾客需求及消费行为特征的基础上,还可以进一步对有着不同偏好、特征的用户开展个性化营销,从而提高用户与产品和页面推送之间的匹配度,提升用户体验。(3)营销效果能够被评估测量。在企业日常经营过程中,营销效果通常以用户反馈来评定。然而,传统营销模式的用户反馈信息通常很难被收集与直接测定。但在数据库营销中,能够通过收集平台用户评论、复购、分享行为等反馈信息以实现从数据层面量化营销效果。(4)开拓新市场,把握新机会。数据库营销要求营销者与用户之间建立不间断的紧密关系并频繁互动。因而企业营销者可以充分利用数据库系统中已有用户的交易数据,通过持续追踪和观察特定顾客的购买需求及偏好,建立有效的用户反应机制,并基于用户反应及时捕捉用户需求变动,为其提供新产品和服务,从而达到开拓新市场,抢占市场先机的目的。

2.2.3 数据库营销的运作程序

数据库营销的运作程序分为数据的采集、存储、处理、分析、使用、完善六个环节。其中,(1)数据采集。主要对用户的相关数据进行采集,比如用户个人信息、购买偏好、企业与用户之间的互动交易以及用户消费行为特征。同时,企业营销者还要尽可能多的搜集除用户以外的产品特征、竞争者信息、渠道流量等多层面数据以充分开发利用数据库资源。(2)数据存储。将搜集的数据信息准确无误地存储到相应数据库中。(3)数据处理。由于企业数据库不仅用于开展营销工作,还应用于研发、公共关系、产品等多部门,因而在利用大数据、人工智能等数据处理技术将无序、不规范以及不同属性的原始数据整合为条理清晰的数据库系统时,既要确保数据的一致性,还要实现数据的广泛应用。(4)数据分析。利用分类、聚类、推荐等算法对已处理的数据进行建模分析,并基于分析结果归纳用户消费特点,寻找、锁定理想用户。(5)使用数据库。数据库在建成之后,可从多方面使用数据。比如,依据数据全面洞察用户偏好及购买意愿等信息,并基于数据分析结果制定精准化营销策略。(6)完善数据库。随着多方面数据信息的不断累积增加,数据库也要随之更新完善,从而有助于企业营销者根据用户需求变化调整相

应营销战略，推动企业发展。

数据库营销以多平台的数据采集为依托，基于大数据技术的预测分析能力，能够给企业带来营销效率和收益效果的双提升。因此，在愈发成熟且被广泛应用的网络信息技术的推动下，数据库营销已逐渐成为电商行业未来发展的必然趋势。

2.3 我国跨境电子商务发展概况

2.3.1 跨境电子商务概念

跨境电子商务是指不同关境的交易主体在电子商务平台达成交易、进行支付结算，并通过跨境物流运输送达商品、完成交易的国际贸易活动^[20]。跨境电子商务作为依靠网络技术和国际物流运输的新型贸易模式，具有门槛低、环节少、成本低、周期短等优势，该发展模式不再将平台仅用于商品信息展示，而是使其直接面向终端市场，以满足用户的个性化需求。另外，依据不同划分准则可将跨境电子商务分为以下类型：（1）按照货物不同的跨境流向可将其分为进口型和出口型；（2）按照不同交易主体可将其分为企业间跨境电子商务交易（B2B）、企业对个人跨境零售电子商务交易（B2C）、个人对个人跨境零售业务（C2C）三种类型，并以发展 B2B、B2C 模式为主。

2.3.2 我国跨境电子商务发展现状

1、出口交易规模占据主导地位

随着我国“一带一路”倡议的带动、网络信息时代的蓬勃兴起以及全球经济贸易一体化的深度融合，我国跨境电子商务以良好的发展态势实现了交易额逐年增长。统计资料显示，2019 年我国跨境电子商务市场总规模为 105000 亿元，同比增长 16.66%，其中，在跨境电子商务总交易规模中，出口交易规模为 80300 亿元，进口交易规模为 24700 亿元，进出口比例结构为：出口占比 76.48%，进口占比 23.52%。由图 2.1 可知，近几年来我国出口交易规模占比均高于进口，在跨境电子商务发展中占据主导地位^[40]。尽管随着国内用户消费需求的不断增加，我国进口交易额占比逐渐展露出上升势头，但在跨境出口政策大力支撑及海外市场不断拓展的作用下，我国跨境出口贸易仍将以较快速度成规模发展，并持续拉动我国对外贸易增长，使“中国制造”深深扎根于世界各地。因而预计在未来很长一段时间内，我国跨境出口市场规模仍将占据较大比重。

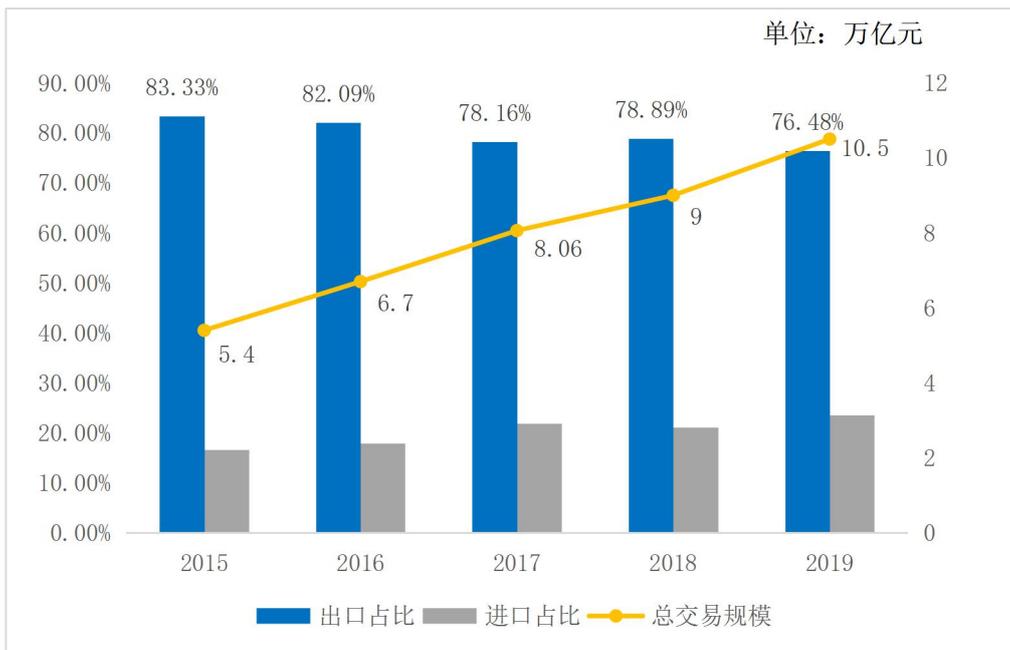


图 2.1 跨境电子商务进出口交易规模占比

数据来源：电子商务研究中心. 2019 年度中国跨境电子商务市场数据监测报告

2、跨境电子商务以 B2B 为主要交易模式

据上文分析可知，跨境电子商务根据不同交易主体可被划分为 B2B、B2C、C2C 三种模式，其中，B2C、C2C 又被并称为跨境网络零售。图 2.2 所示为近几年跨境电子商务 B2B 及跨境网络零售的交易规模占比情况，由图可知，B2B 交易规模占比远高于网络零售，在近五年均占据八成以上的比重，是跨境电子商务的主流交易模式，虽然其占比变动呈下降趋势，但 B2B 交易规模仍逐年增加。究其原因这是由于 B2B 模式是我国传统对外贸易在顺应网络信息发展基础上的贸易转型，因而以贸易企业为主体的 B2B 模式具有最完备、最成熟的贸易基础以及强大的贸易资源，故该模式占据的主导地位在短期内将难以被撼动。

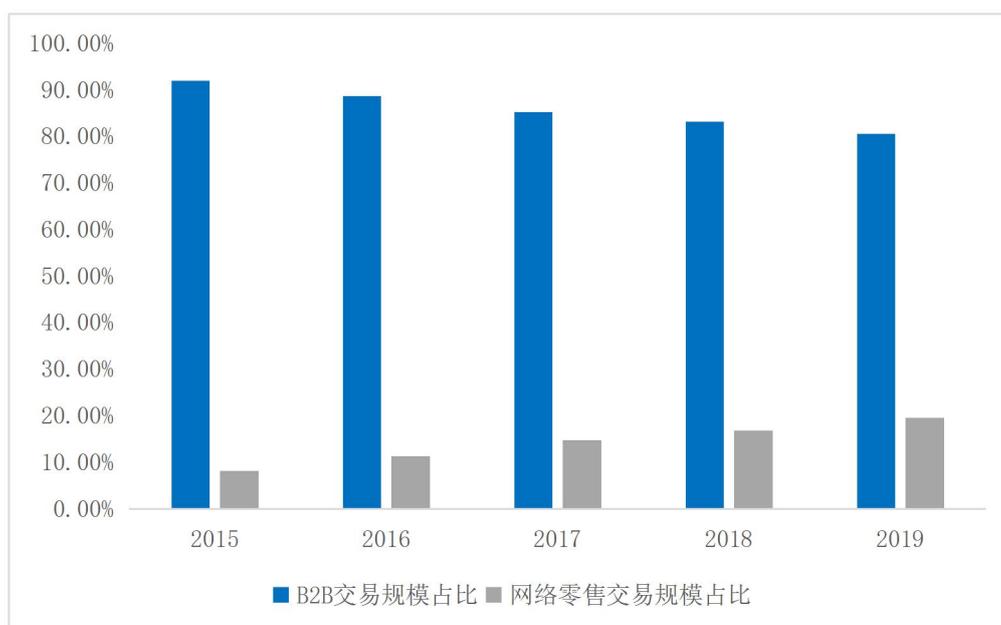


图 2.2 跨境电子商务 B2B、网络零售近几年交易规模占比情况

数据来源：电子商务研究中心. 2019 年度中国跨境电子商务市场数据监测报告

3、跨境电子商务目标市场不断扩展

目前，我国跨境电子商务的主要目标市场仍以欧美地区国家为主，但同时也不断向东南亚、南亚、南美等发展前景广阔的新兴市场扩展。图 2.3 所示为 2019 年我国跨境电子商务进出口目标市场分布情况，由图可知，2019 年我国跨境电子商务的主要目标市场分布在美国、德国、俄罗斯、英国等欧美地区国家，剩余则主要流向日韩、中国香港、中国台湾、澳洲、东南亚、南美、南亚等地区。值得注意的是，随着“一带一路”倡议的不断深入，我国与其沿线市场之间的贸易联系也愈发紧密，2019 年东南亚、南亚等市场在我国进出口目标市场交易额占比中均有显著上升趋势，展露出良好的发展势头。另外，我国出口型跨境电子商务企业在面对发达国家中低端市场以及东南亚、南美、非洲等新兴国家市场时，能充分发挥中国制造的高效率、低成本优势，保持其强劲的市场竞争力，加之网络技术的日趋成熟以及互联网的广泛普及使得新兴市场所属地区用户的在线消费习惯已逐渐形成并呈现出一个发展潜力巨大的跨境电子商务需求空间。因此，我国应紧跟国际发展趋势并不断拓展跨境电子商务进出口目标市场。

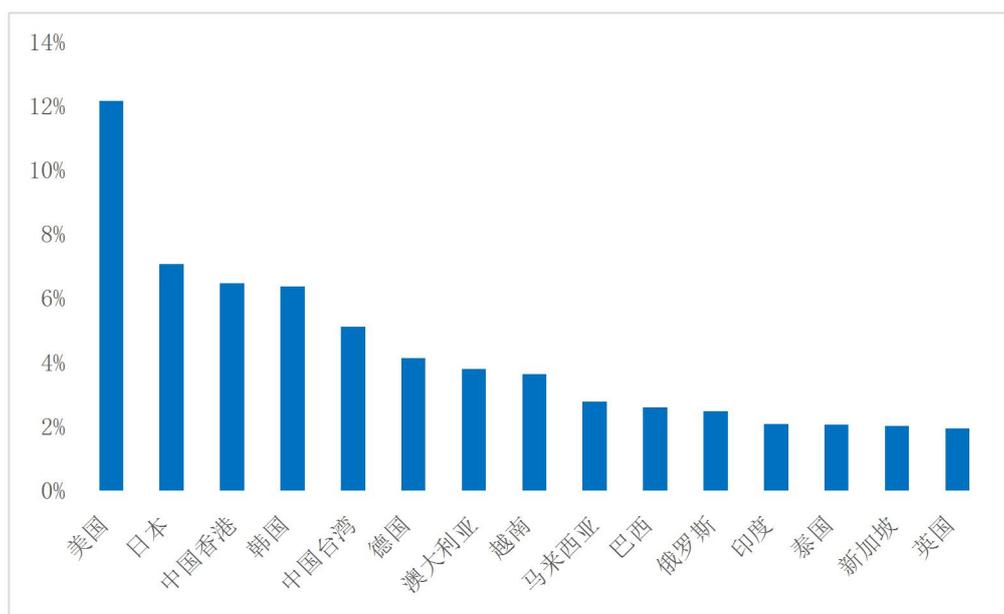


图 2.3 2019 年我国跨境电子商务目标市场进出口交易额 TOP15

数据来源：海关总署统计数据

2.3.3 跨境电子商务对我国对外贸易的影响

近年来，我国跨境电子商务在多方面打破传统贸易限制，正以迅猛的发展势头进入快速增长阶段，在成为我国进出口贸易新增长点的同时，还对我国对外贸易的开展产生了诸多积极影响。首先，跨境电子商务不仅是推动我国对外贸易模式转型的重要途径，而且还能充分发挥其简化中间交易环节、减少跨境交易成本、降低中小企业跨境贸易门槛、增加就业机会、促使电商企业发展升级等优势，给国内中小企业带来机遇，并持续拉动我国对外贸易增长。其次，跨境电子商务发展对我国开展进出口业务的企业经营管理模式也有着积极影响。由于在互联网、大数据技术的作用下，跨境电子商务发展模式在一方面使企业信息流、商流、物流得到高度融合，另一方面也进一步推动各企业逐步形成集信息展示、网络营销、在线支付、物流运输于一体的全新经营管理模式。最后，跨境电子商务还对我国服务贸易和信息贸易发展起到促进作用，使更多中国跨境电子商务企业重视网络信息技术，进而促使我国对外贸易向着知识密集型、技术密集型转变^[36]。

3 JDC 集团跨境电商营销现状概述

JDC 集团自创立以来,始终坚持用户至上的经营理念,持续创新营销模式,致力于为用户打造极致的购物体验,并伴随其零售业务的快速拓展逐步完成了电脑数码、手机通讯、消费品、生鲜等全品类覆盖,成为国内规模较大的零售电商企业。与此同时,集团又在此基础上进一步开拓海外货源以满足国内用户对海外商品的购买需求,并将集团的目标客群定位为国内消费者。但随着近年来在市场、政策等多方因素共同推动我国跨境电子商务发展的形势下,JDC 集团在仅面向中国大陆用户的基础上,进一步将商品售卖范围延伸至中国港澳台及全球 200 多个国家地区,创办跨境电子商务平台,旨在使全球用户都能购买 JDC 集团精选好物,享受优质购物体验。

JDC 集团跨境电子商务平台作为集团新兴起的业务板块,虽创办较晚,但随着近几年我国跨境贸易浪潮的大力推动以及集团的合理经营,使得平台业务发展已渐成规模。另外,由于 JDC 集团面向海外客群的跨境电子商务业务尚处于起步阶段,市场前景和发展空间均较为广阔,在集团未来战略规划中也占有重要地位。因此,为使 JDC 集团充分了解其跨境电子商务平台经营现状及存在的主要问题并做出相应改进,从而保持平台业务发展的良好态势,本章从 JDC 集团跨境电子商务平台的相关指标入手,在对其整体经营现状进行概述的基础上,进一步归纳总结平台目前存在的主要问题,为后文研究提供新思路。

3.1 JDC 集团跨境电商平台经营现状分析

目前,JDC 集团跨境电子商务平台发展仍处于起步阶段,且企业战略定位也以开展国内零售业务为主。但海外市场的不断拓展将促使海外客群持续释放消费潜力,进而推动平台逐步迈入快速发展轨道。基于此,本节从经营状况、用户运营、物流运输等多视角入手,对 JDC 集团跨境电子商务平台(以下简称为平台)的经营现状展开剖析。

首先,本节选取平台 2019.1—2020.6 运营期间的用户量、订单量、GMV 总值指标数据对平台经营情况进行分析。由图 3.1 可知,从 2019 年及 2020 上半年的整体发展趋势看,2020 上半年平台整体经营情况优于 2019 年,其各项指标数据与 2019 年相比均有一定幅度增长。据 2019 年的发展趋势来看,2019H1(上半年度)用户量、订单量、GMV 总值各项数据除了在 6 月“618 大促月”达到了同步上涨的最大值之外,其余月份的经营情况均较为平淡。但从 7 月开始,鉴于“618 大促”期间平台推广、活动宣传到位,

平台各项基础指标水平与 1-5 月相比均有了显著提高，并于 9 月达到全年第二个峰值。同时，由于受到十一月大促活动预热影响，用户会将购买行为集中在大促活动期间进行，故 10 月及 12 月的各项指标数据均呈现出回落趋势，而 11 月的平台各项指标在“双十一狂欢活动”的营销推动下达到全年最高水平。另外，从 2020H1 的发展趋势来看，基于 2019 年两次大促活动的宣传推广与用户累积，平台整体经营情况有了较大改善，用户消费处于持续增长状态。然而由于受到国内新型冠状病毒疫情影响，尽管 2020 年 1 月有“双旦”促销活动的引流支撑，但平台 1 月的各项指标数据与 2019 年 12 月相比均有所下降。接着到了 2 月、3 月，新冠疫情逐渐蔓延至世界各地。平台基于疫情严峻的发展趋势及时调整、补充口罩、防护服、医疗设备等防疫物资的产品上架，并售往世界各国地区，既拉动平台交易额增长，也使得平台各项指标数据于 3 月达到 2020H1 的第一个峰值。与此同时，在我国有效的防疫抗疫工作推动下，4 月、5 月国内疫情基本得到控制，因而平台经营逐渐恢复常态，并于“618 大促月”达到 2020H1 交易总额的最大峰值。以上一系列数据及趋势说明：（1）JDC 集团紧跟海内外发展动向，及时调整平台产品架构的经营策略起到了拉动交易额增长的作用；（2）JDC 集团在大促月活动期间用户活跃度最高、交易量最大，且目前的拉新引流、用户促活、平台知名度提升等相关推广、宣发工作也主要依靠每年两次大促活动展开，该发展现状也逐渐突显出由于平台用户的促销敏感特点而导致的平台用户粘性低、流失率高等问题；（3）目前平台用户在大促活动期间产生购买行为较多的主要原因除了有享受价格优惠之外，还有运费补贴。由于跨境物流运输的高成本使得买家须支付相对较高的运费金额，但高昂的运费往往也并不能给用户带来货品送达的及时性。因此，目前跨境物流运输仍是平台目前的发展短板，如何提高物流运输效率成为集团需重点考虑的关键问题。

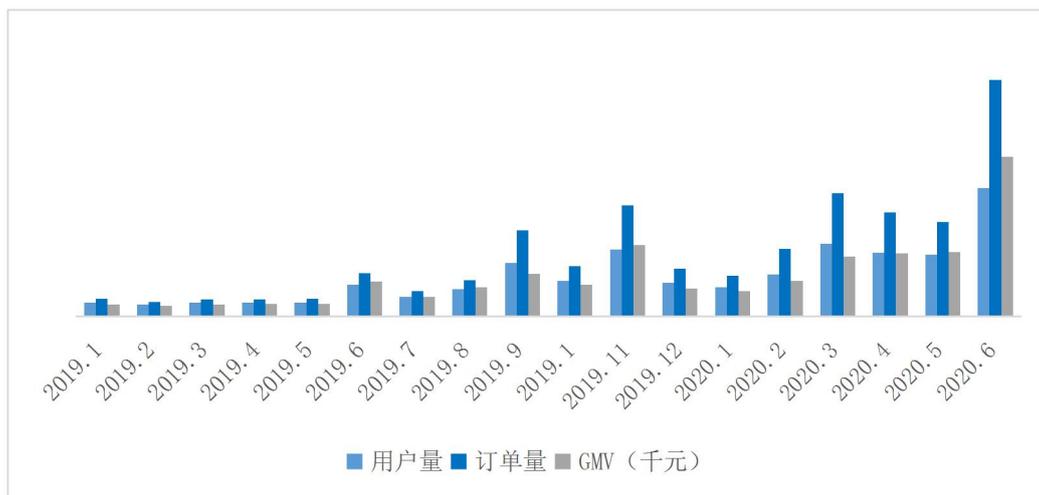


图 3.1 JDC 集团跨境电子商务平台 2019.1-2020.6 基础指标数据

3.2 JDC 集团跨境电商平台存在的主要问题

3.2.1 产品层面

从产品层面来看，各品类商品构成比例失衡、产品及页面推荐不精准、产品售后服务水平有待提高构成了产品层面的三大问题。其中，各品类商品构成比例如图 3.2 所示（数据统计周期为 2017.1.1—2020.6.1 期间平台品类交易的历史数据），由图可知，在商品数量排名前十的品类中，图书类目下的商品数量约占七成，而排在第二位的服饰内衣类目下的商品数量占比小于 10%，箱包皮具、手机通讯类目下的商品数量占比仅不到 1%。可以看出，平台各品类的商品数量差异较大，图书品类的商品数量远超出其他品类，品类的商品构成比例处于严重失衡状态，而该发展趋势则会导致平台消费客群受限，且无法满足不同用户日趋多样化的商品购买需求。同时，除了品类的商品构成比例失衡这一现实问题之外，目前平台还亟需解决产品及页面推荐不精准的问题，即由于无法准确定位用户需求从而导致用户与产品及页面推送之间的匹配度较低，从而影响用户平台体验，增加其流失风险。另外，各跨境电商平台之间的产品竞争不仅包括产品价格、品牌及品质，产品售后更是其重要组成，较为完备的售后服务一方面能延长产品生命周期，另一方面还会增加用户粘性，提高其平台忠诚度。JDC 集团跨境电子商务平台目前正面临的产品售后服务问题可分为两个层面，首先，于平台卖家而言，由于受到跨境运输成本、出口通关难等条件限制，商家在将产品售往世界各地后，通常很难保障跨境货物长期有效的售后服务，而产品售后服务的不完善则会导致平台用户部分流失、客诉率

增加等后果；其次，于平台买家而言，即便有买家存在退换货的售后服务需求，但退换货的跨境物流运输时长将直接影响其申请售后服务的意愿，而由于无法保障的售后服务也将使更多海外用户降低对中国跨境电子商务平台的信任感。

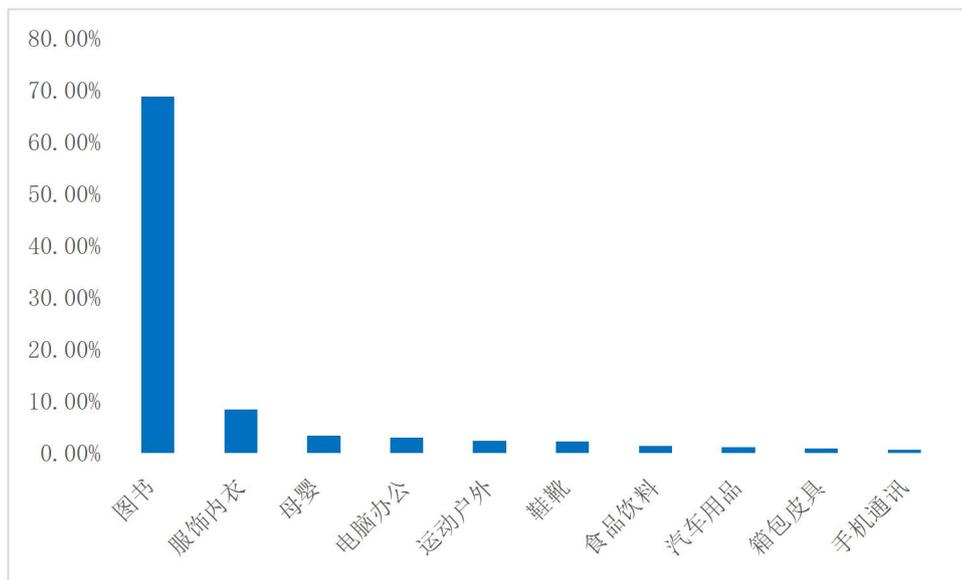


图 3.2 JDC 集团各品类商品数量占比情况

3.2.2 渠道层面

在渠道层面，跨境物流运输效率低及出口通关难度高成为制约 JDC 集团跨境电子商务平台业务开展的主要原因^[39]。其中，由于受到空间距离、配送时间成本、跨境物流产品规格、包裹丢件率等条件限制，JDC 集团的跨境物流运输发展仍存在短板。目前，JDC 集团跨境电子商务平台主要采用邮政小包、八达通、神州集运、UPS、马可达、跨境专线、国际快运等物流模式，在这几种常用物流模式中，只有中国港澳台有逐一对应的专线物流，因而相比于其他地区，中国港澳台地区的物流运输效率相对较高，其余地区的跨境运输效率还有待提升。另外，JDC 集团跨境电子商务平台在近几年的经营过程中，出口通关问题一直是阻碍其发展的重要原因，更是影响“中国制造”国际口碑建立的制约因素。由于 JDC 集团跨境电子商务平台主要开展 B2C 业务，故下文将针对 B2C 模式的通关问题展开剖析。B2C 物流运输模式呈现出的订单较为零碎、物流包裹规格差异较大等特点使得平台包裹时常由于产品出口通关不符合海关要求造成货物滞留海关，影响物流运输效率，进而延长买家收货时间，降低其平台购物体验。

3.2.3 客诉及商家管理层面

在客诉及商家管理层面存在的主要问题为用户投诉率高、商家存在违规行为。首先，平台客诉率较高的原因集中在物流运输、产品质量和售后服务三方面，其中物流运输时间过长是大部分海外用户产生投诉的主要原因。据前文分析可知，跨境物流运输受限较大，运输距离远、速度慢以及出口通关难度大都是造成买家收货时间过长的直接原因。另外，产品质量问题与售后服务缺失也使用户的平台购物体验大打折扣，故而转向投诉。其次，在商家管理层面，平台存在的商家管理漏洞使得部分商家会产生一些诸如刷单、虚拟发货等违规操作。而商家产生违规行为的主要原因是由于平台成立较晚，各项奖惩机制还未成体系且平台对于违规操作的惩罚力度较小，因而使部分商家产生“钻空子”心理。为此，JDC 集团跨境电子商务平台还应不断建立健全平台商家奖励惩罚机制，加大管理力度，以有效解决商家违规行为。

3.2.4 数据层面

据前文现状分析可知，目前 JDC 集团跨境电子商务平台由于其用户促销敏感特征使得平台每年 6 月、11 月两次大促活动的成交量及交易额较高，但非活动期间的经营现状仍有待改善。然而，仅靠促销活动吸引用户下单消费并不是跨境电子商务企业经营的长久之计，在非活动期间仍然能有大量用户产生购买行为才是推动企业交易额连年增长、规模不断扩大的必要保障。而该现状存在的主要原因在于平台日常经营并没有与交易数据紧密结合，各层面数据得不到充分利用，海量交易数据中蕴含的潜在价值无法被挖掘出来。另外，交易数据无法充分利用也会进一步引发平台目标用户定位不精确、产品及页面推送内容与用户需求之间匹配度低、用户平台购物体验不佳等一系列问题。因此，要想解决由用户促销敏感性以及数据利用不充分带来的一系列后续问题就要将交易数据深度融合到企业经营中去，并基于数据分析结果展开后续的企业决策。

由于数据不充分利用导致的后续问题将使得平台无法针对各类目标用户展开个性化推荐及精准化营销，而该问题则会直接影响平台用户消费、留存、复购、促活等经营情况，故数据层面存在的问题是阻碍 JDC 集团跨境电子商务平台发展的主要因素。因此，在下文分析中，将充分利用平台交易数据，从目标用户定位和用户购买行为预测模型构建两方面重点展开研究。

4 立足于交易数据的目标用户精准化定位

4.1 用户基本特征分析

据前文分析可知，第二章对数据库营销理论和我国跨境电子商务发展概况进行了详细说明，第三章基于相关理论内容，又进一步探究 JDC 集团跨境电子商务平台营销现状及存在的主要问题。因此，本章内容将依据上文分析结论，立足于平台交易数据，从用户基本特征、用户价值、用户偏好三方面入手，对平台目标用户进行精准化定位。本节首先从用户人口属性及用户消费习惯两方面对用户基本特征展开研究。

4.1.1 用户人口属性

用户群体的准确划分能助力企业基于不同用户需求开展个性化营销，进而使得企业营销的精准化程度不断提高^[26]。那么，要想实现用户的精细划分就需首先把握用户人口属性。基于此，本文选取 2020 年上半年（2020.1—2020.6）平台用户画像数据，从性别、年龄、婚姻育儿、学历水平、地区分布四方面对用户人口属性展开分析。

1、性别、年龄

首先，年龄是用户静态属性中极为重要的组成因素，由于随着用户处于不同年龄阶段，其购买偏好及需求也会发生相应改变，因而分析各年龄阶段用户有助于企业把握用户消费特征。依据平台用户资料划分标准可将用户年龄划分为六个阶段：15 岁以下、16-25 岁、26-35 岁、36-45 岁、46-55 岁、56 岁以上。其次，即使性别只有男性和女性两种分类，但该属性也会在用户价值、用户偏好层面体现出较大差异，因而性别也是用户人口特征的重要组成因素。由表 4.1 可知，处于 26-35 岁年龄阶段的用户占比最高，36-45 岁、16-25 岁次之，因而平台主要面向 16-45 岁的处于求学、初入职场、初为父母、养家育儿、二胎父母、资深中产等不同社会状态的消费客群。同时，根据表 4.1 中用户性别差异可知，16-45 岁的女性用户占比为 87%，男性用户占比为 86%，因而在主要消费客群处于的年龄阶段中，女性用户略多于男性用户。

另外，在对不同年龄用户进行研究时，除了关注用户量最大，消费最多的年龄阶段之外，还应考虑具有潜在消费实力的人群，如年龄在 56 岁以上的用户。该类用户既有一定经济自主性，又具备一些与中青年人相同的购物特质，从而形成了特点鲜明的中老年消费理念及生活方式。56 岁以上的用户可归为银发人群，该群体形成的规模性消费可

随之称为“银发经济”。“银发经济”看上去像是没有发展前景的“夕阳红经济”，但实则不然，是随着互联网、线上购物理念不断向中老年人群渗透以及我国老龄化进程不断加快而逐渐展露出巨大市场价值的“朝阳经济”，因而平台在后续经营中应重视该群体发展。

表 4.1 用户年龄、性别分布

年龄	性别	男	女
		15 岁以下	0%
16-25 岁		18%	8%
26-35 岁		40%	46%
36-45 岁		28%	33%
46-55 岁		7%	6%
56 岁以上		7%	7%

2、婚姻、育儿状况

据平台用户偏好分析可知，有孩子和没孩子的已婚用户所对应的购买偏好存在明显差异，大概率有孩子的已婚用户主要偏好于购买母婴、玩具乐器、童书及中小学教辅等品类，而小概率有孩子的已婚用户则偏向于购买居家日用、食品饮料、电脑办公及电子数码类产品。因此，把握用户婚姻育儿状况对企业开展不同品类活动时能否准确定位受众客群而言十分关键。表 4.2 所示即为平台用户婚姻育儿状况，由表可知，在未婚用户中有 61%的用户属于有孩子的概率低及较低，39%的用户属于有孩子的概率高及较高；在已婚用户中有 65%的用户属于有孩子的概率高及较高，有 35%的用户属于有孩子的概率低及较低，该数据说明在未婚群体中没有孩子的用户占比高，在已婚群体中有孩子的用户占比高，因而应将母婴类、育儿类专场活动更多的推荐给已婚群体。

表 4.2 用户婚姻、育儿情况分布

有孩子		婚姻状况	
		未婚	已婚
概率低		35%	19%
概率高		14%	14%
概率较低		26%	17%
概率较高		25%	51%

3、学历

据相关学者调查研究可知，学历水平高低将会对用户的产品认知、选择意愿以及购买行为产生影响，即用户受教育程度与其购买偏好之间有直接影响关系。一般来说，高学历用户与低学历用户在消费偏好、消费金额、消费频次上均存在明显差异。因此，分析用户学历不仅能精准划分用户类型，还能使平台针对不同受教育程度的受众群体开展品类或品牌的个性化推荐。由图 4.1 可知平台用户学历分布情况以大学本科为主，研究生及以上次之，大中专科、高中及以下的用户量占比较小。因此，平台用户学历在大学本科及以上的用户量占比为 93%，可以得出平台消费客群有着较高学历的结论。

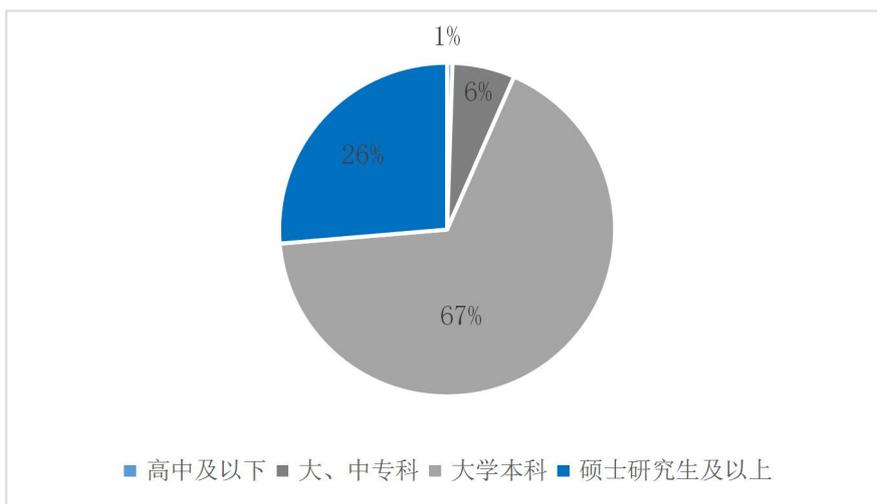


图 4.1 用户学历分布

4、用户收货地区分布

根据前文市场营销理念分析可知，现代营销战略已由产品导向转为用户导向，因而

研究用户收货地区分布能为企业未来区域规划提供相应参考。就目前发展现状来看, JDC 集团跨境电子商务平台已开通中国港澳台、欧美、东南亚等多个独立站, 并预计在未来开设更多站点, 实现各区域的相互连接, 逐步形成遍布世界的站点网, 以达成“中国制造, 销往全球”的目标。由平台数据可知用户主要分布在 104 个国家和地区, 但由于用户收货地大多集中在某些地区, 故本文选取用户量占比排名 TOP15 的地区展开分析。如图 4.2 所示, 平台用户收货地主要分布在中国港澳台及美国, 其中, 中国香港的用户量占比最高, 说明中国香港用户在平台产生了较多购买行为。另外, 从图中能够看出排在后几位国家的用户量占比较低, 均不到 1%。因此, 如何有效刺激用户消费较少的国家地区增加用户消费并扩大用户规模也是平台经营过程中仍需关注的重点问题。

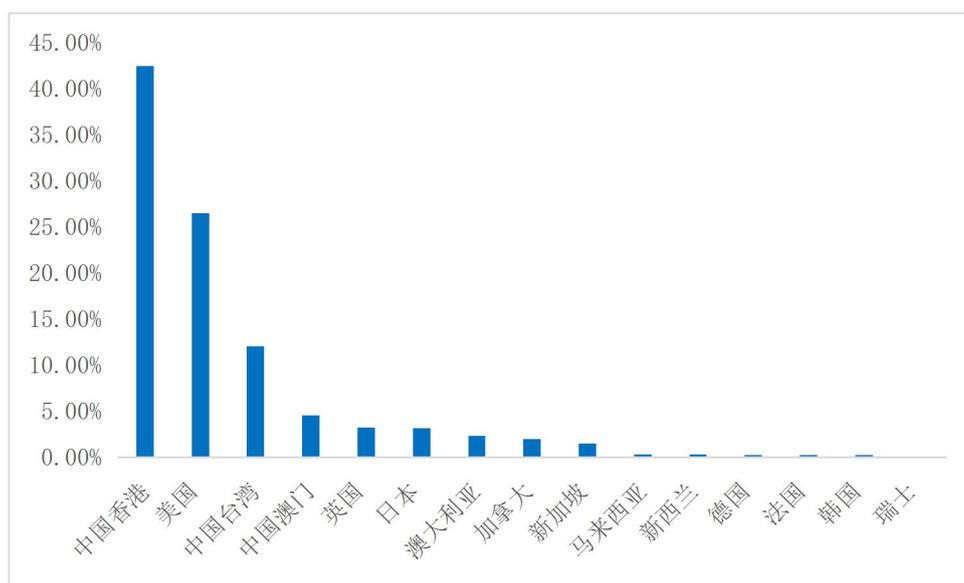


图 4.2 不同收货国家地区的用户量占比

4.1.2 用户消费习惯

通过上文对用户人口属性的分析可大致归纳出平台用户主要人口特征为: 用户年龄阶段分布在 26-35 岁、性别以女性用户为主、婚姻育儿情况为未婚无孩子和已婚有孩子、学历以大学本科及以上为主、收货地所属地区集中在中国港澳台地区和美国。基于上文研究结果, 本节选取 2018.1—2020.6 期间平台交易数据, 进一步从端口偏好、支付方式偏好以及使用优惠券情况三方面剖析平台用户的消费习惯。

1、端口偏好

JDC 集团跨境电子商务平台有 APP、M、PC、微信共四个登录端口, 其中, APP

端指应用程序端口，M 端指移动客户端（手机网页端），PC 端指个人计算机端口，微信端指微信客户端。由图 4.3 可知，通过 APP 端触达平台的用户量最高，是平台用户使用端口的首选，PC 端次之，使用 M 端及微信端的用户占比较小。因此，平台在应用程序的设计上不仅应充分考虑如何通过首页推送吸引用户浏览商品进而产生下单购买行为，同时，还要进一步结合用户页面点击渠道偏好以及品类品牌购买偏好，对不同用户群体实行个性化推送。

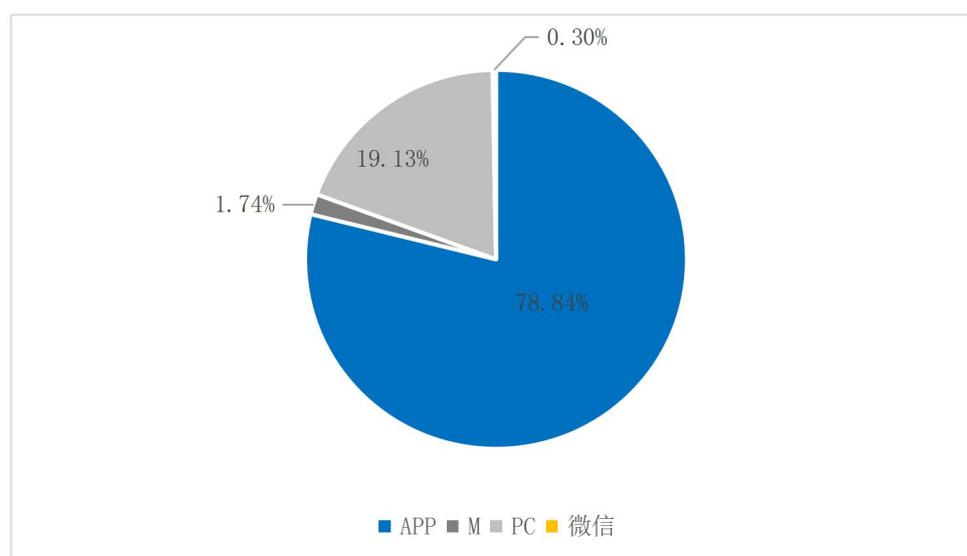


图 4.3 各端口用户量占比

2、支付方式偏好

JDC 集团跨境电子商务平台的用户支付方式主要有 Apple Pay、微信、银联、八达通、外卡、网银、其他支付共七类支付方式。由图 4.4 可知，使用微信线上支付的用户量最多，占比将近过半，外卡和网银支付次之，选择 Apple Pay、八达通及银联进行支付的用户较少。另外，从选择支付方式的用户收货地区分布来看，中国港澳台、美国、日本用户使用的支付方式最多，而平台用户收货地也集中分布于这几个地区，故可从侧面反映出以上五个地区用户活跃度较高。用户支付方式偏好的分析结果一方面能帮助企业基于用户支付偏好深入了解用户消费习惯，另一方面可促使平台根据用户支付偏好与相应支付渠道进行交易合作，以“支付优惠”刺激用户消费。

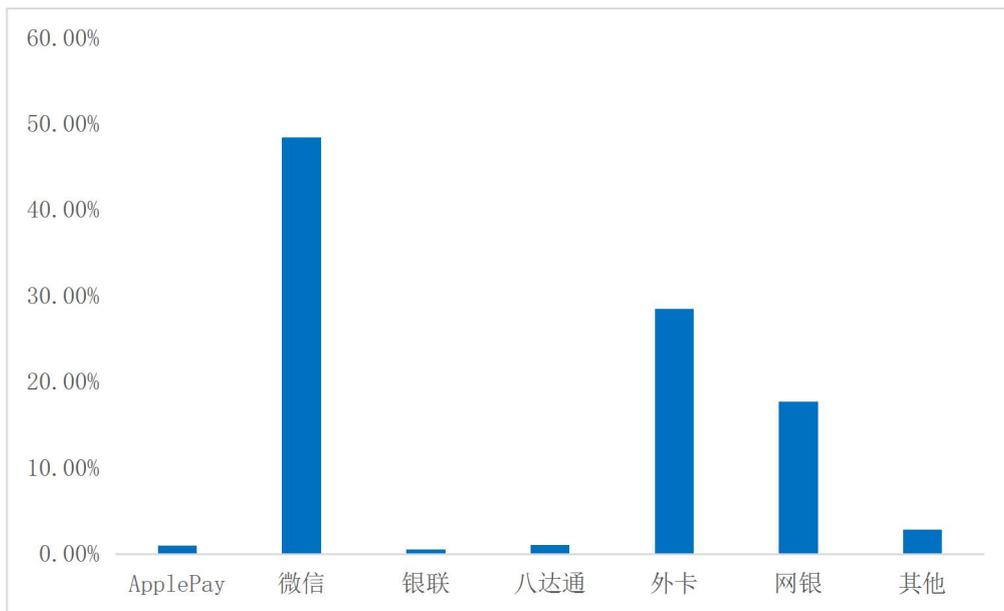


图 4.4 不同支付方式的用户量占比

3、优惠券使用情况

优惠券营销是电商企业解决用户拉新、促活、复购、提高客单价难题的有效营销方式，优惠券在一定程度上既能实现电商企业的利润最大化，又能达到刺激用户冲动消费的效果，这也是大多数商家在大促期间选择发放优惠券而不是直接降价的原因。JDC 集团跨境电子商务平台常见的优惠券活动类型主要有：特价、折扣、满减、满赠、秒杀、团购、拼单、低价换购、第二件半价、N 元任选几件（如 99 元任选 3 件）。其中，满减、满赠、低价换购、第二件半价、N 元任选几件类型的优惠活动对运营结果的直接影响为提高客单价；团购、拼单活动则是为了拉新引流，吸引新老用户组合下单，提升新用户增长率；特价、折扣、秒杀活动是为了促使近期无消费记录的沉默用户再次产生购买行为，将其转化为回流、活跃用户。由图 4.5 可知，平台使用优惠券的用户量占比高于不使用优惠券的用户，说明平台用户大多偏向于使用优惠券下单。因此，针对不同购买需求且对优惠活动有不同偏好的各类群体设计出相应优惠营销活动并对其进行精准化推送是平台日常运营的关键一步。

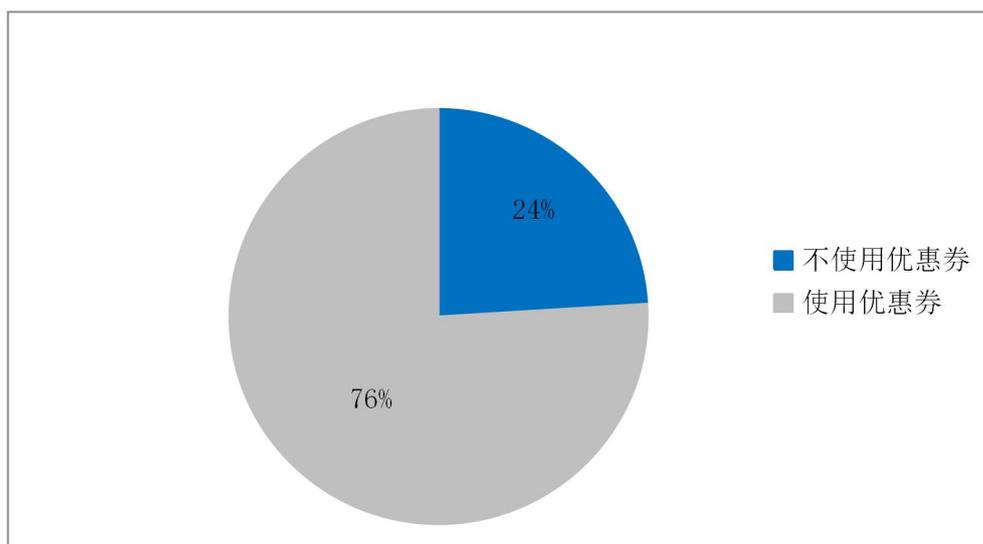


图 4.5 用户使用优惠券情况

4.2 用户价值分析

随着网络信息时代的到来，用户在拥有更多自主选择权的同时，也逐渐由满足自身消费需求转变为价值追求者，因而企业只有正确、全面认识用户价值内涵，确立以用户价值为中心的观念，最大限度满足用户需求，才能在跨境电子商务的激烈竞争中获得持久优势。为此，本节选取平台 2018.1—2020.6 期间交易数据，通过剖析用户忠诚度、用户等级以及用户留存、复购、新用户增长等情况进一步对 JDC 集团跨境电子商务平台用户价值展开研究。

4.2.1 用户忠诚度及消费能力

1、忠诚度分析

用户忠诚度是指用户内在态度与外在行为的统一，其中，内在态度是指用户对某商品、品类或品牌存在积极的态度、情感或偏爱，而外在行为则表现为用户的重复购买。一般可将用户忠诚划分为三个层次：情感忠诚、意识忠诚及行为忠诚，情感忠诚指用户对商家本身及其产品、品牌的积极态度，意识忠诚指用户未来产生购买意愿的可能性，而行为忠诚就是指用户实际的重复购买行为^[18]。另外，按照时间长短可将用户忠诚度分为长期忠诚和短期忠诚，前者在长时间内不会轻易改变，而后者会在各时期由于偏好的变化而随之改变。

JDC 集团跨境电子商务平台的用户忠诚度模型共分为高度、中度、低度三个程度，其中，低度忠诚型又可细分为普通、投机、偶然三种类型以及近期、远期两个维度。普通型是指用户对平台的忠诚度较低且有流失风险；投机型是指用户对平台偏好程度较低且通常只会在平台产生“薅羊毛”行为；偶然型是指用户对平台有一定忠诚度，但购买、访问频次较低，即在平台有下单倾向的沉默用户。忠诚度的具体模型由图 4.6 所示，平台各等级忠诚度的用户量占比由图 4.7 所示。据图 4.7 可知，平台低度忠诚的用户占比最高，高度忠诚次之，中度忠诚用户量占比最低。另外，在低度忠诚用户的细化分类中，偶然型用户占比最高，普通型次之，投机型用户占比最少，该数据说明平台目前发展趋势为：处于低度忠诚状态的平台用户最多，其中以偶然型和普通型用户为主，但中高度忠诚用户群体也渐成规模。要使平台用户忠诚度等级不断提高，就要求平台在运营过程中一方面需通过满足用户需求、提升用户购买体验以提高留存率，保留中高忠诚度用户，同时，还要尽可能唤醒沉默用户、挽回流失用户；另一方面还要深入探究用户下单频次低、存在流失风险的原因，并将高度忠诚用户与低度忠诚用户的各项指标进行对比分析从而找出差异，并进一步采取措施将低度忠诚用户逐渐转化为中高度忠诚度用户。

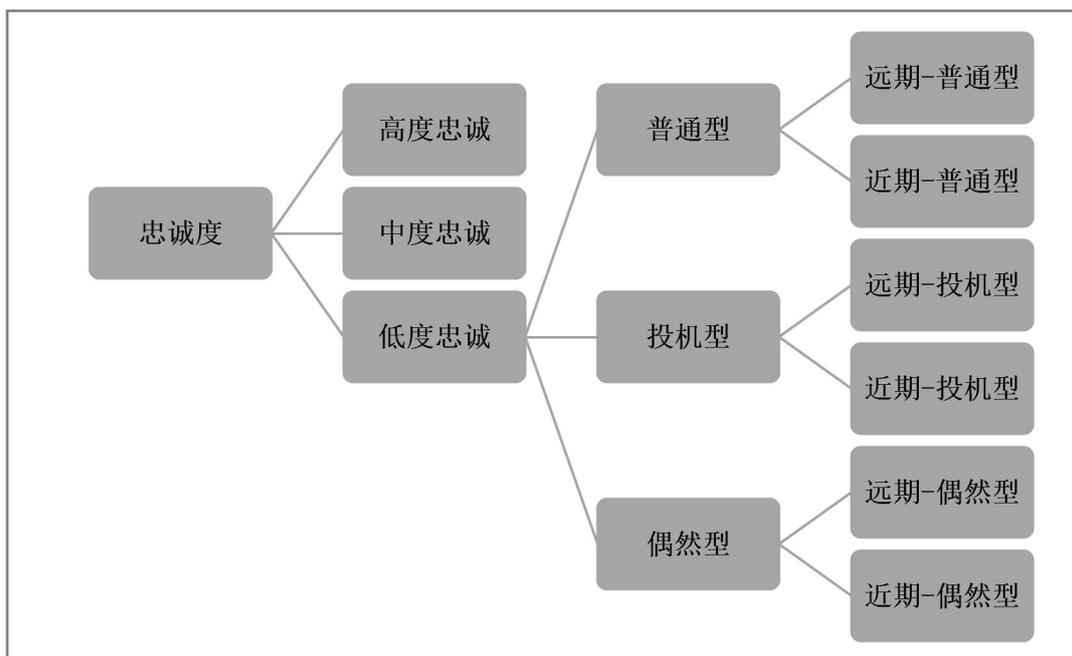


图 4.6 用户忠诚度模型

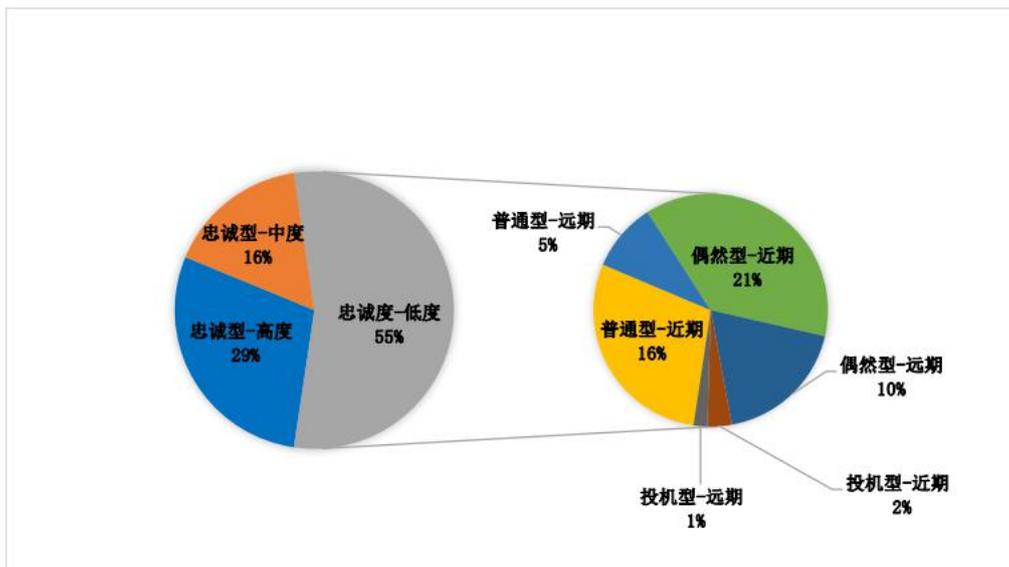


图 4.7 不同忠诚度的用户量占比

2、消费能力分析

消费能力是用户价值体系的重要组成，用户消费能力水平的高低将直接影响用户价值等级。JDC 集团跨境电子商务平台在进行用户消费分析时，常使用购买力等级、用户会员等级、价值分组这三个指标来衡量用户消费能力，下文将分别从这三个指标入手展开研究。

(1) 购买力等级

购买力是指用户通过支付一定货币购买商品或劳务的能力。平台将不同用户的购买力等级划分为五种程度，从 LV1 到 LV5 是购买力递增的过程，即 LV1 等级用户的购买力最弱，LV5 等级用户的购买力最强。由图 4.8 可知，购买力等级为 LV2 的用户量占比最高，占 74.60%，LV3 等级的用户量占比次之，占 21.88%，购买力较弱的 LV1 以及购买力较强的 LV4、LV5 用户量占比均较少，尤其是购买力较强的两个等级，占比均不到 1%。该数据说明平台用户购买力处于中下等水平，购买力强的用户群体规模仍有待扩展。因此，平台在进行用户规划时，不仅应深入思考如何扩大具备消费实力的高购买力人群，还要同步考虑促使中低购买力人群增加消费的营销模式。

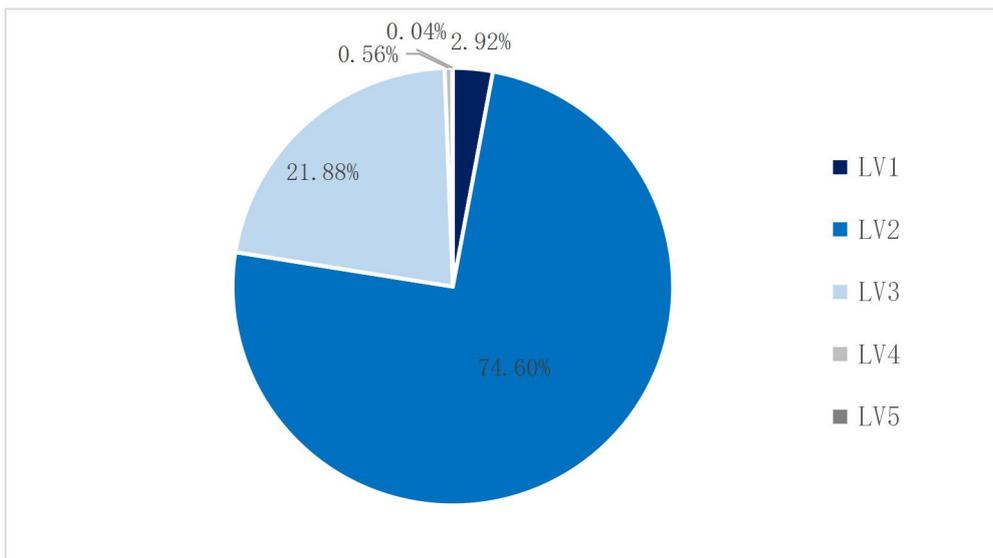


图 4.8 不同购买力等级的用户量占比

(2) 用户会员等级及价值分组

JDC 集团跨境电子商务平台的用户价值分组按照会员积分区间划分，平台将会员积分区间在 0-200 之间的用户归为风险用户、201-1000 之间的用户归为低价值用户、1001-5000 之间的用户归为一般价值用户、5001-10000 之间的用户归为优质用户、10001 以上的用户归为高价值用户。另外，按照会员积分区间的不同还将会员等级分为注册会员、铜牌会员、银牌会员、金牌会员、钻石会员五类，并与不同用户价值逐一对应，具体说明如表 4.3 所示。由表可知，平台用户价值分组及会员等级均取决于会员积分，而会员积分主要通过消费、活跃度、账户资料三种渠道获取，其中，消费指用户在平台的消费下单行为，消费次数越多、频率越高，用户会员积分增加值相应越高；活跃度是指用户在平台产生登录、浏览、分享、评价等行为，活跃度越高，会员积分增加越多；账户资料指用户在平台上的信息完善程度，用户资料信息越完整，将越有助于积分增长。

接下来重点分析平台用户价值分组的具体情况，图 4.9 所示为不同价值分组的用户量占比。由图可知，价值分组的用户量占比由高到低依次为：低价值用户、一般价值用户、优质用户、高价值用户、风险用户，其中低价值用户占比为 69.39%，是用户价值分组中的主要类型，该数据说明目前平台消费客群的会员积分不高，消费能力居中下水平。另外，高价值用户占比 6.61%，风险用户占比 2.42%的数据说明平台高价值用户量高于风险用户量，即目前平台有流失风险的用户少于频繁消费的用户，该数据从侧面反映出平台已形成一定规模忠诚度较高且经常产生复购行为的用户群体。

表 4.3 用户会员等级与价值分组

会员级别	用户价值	会员积分区间
注册会员	风险用户	0-200
铜牌会员	低价值用户	201-1000
银牌会员	一般价值用户	1001-5000
金牌会员	优质用户	5001-10000
钻石会员	高价值用户	10001 以上

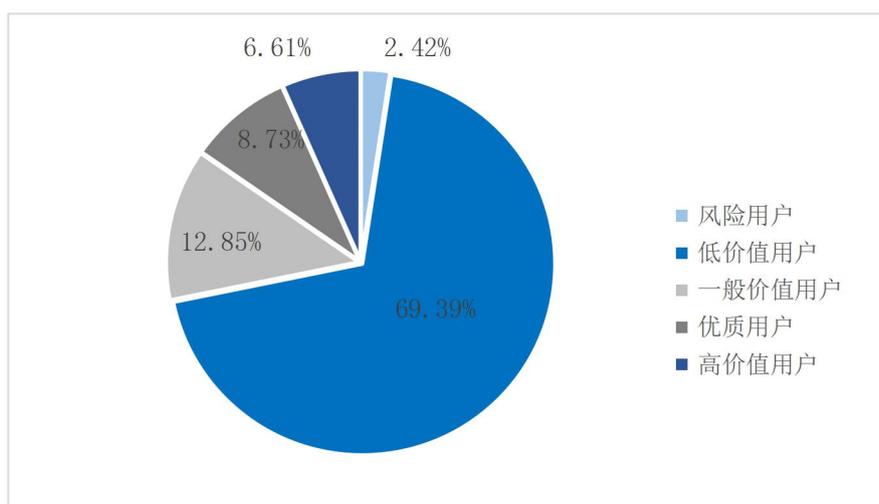


图 4.9 用户价值分组

4.2.2 用户复购、留存及新老用户占比

营销学大师菲利普·科特勒曾提出，据研究发现企业拉新成本比挽回流失用户要高 5 倍，而用户流失率每降低 5%，企业利润率将增长 25% 以上，且被成功挽留的用户将会为企业带来更多的销售额。另外，统计资料显示，一个满意度较高的用户可能会通过推荐分享为企业带来 8 笔潜在交易，而一个满意度较低的用户则可能会直接影响 25 个用户的购买意愿。加之近年来互联网广告投放渠道以及流量成本的不断增加，也导致各电商企业的拉新难度随之上升，因而，在平台的整个生命周期内，企业如何能够将平台用户成功保留并不断挖掘、提升其用户价值，使其持续产生购买转化是企业营销过程中需考虑的核心问题。为此，下文将重点分析 JDC 集团跨境电子商务平台目前用户复购、留存及新用户增长情况。

1、用户复购

复购是指用户对某一品牌商品或服务产生重复购买行为，一般以复购率来表示。复购率水平的高低不仅能直接反映各企业用户群体的稳定程度，也能说明用户忠诚度高低，即用户重复购买的次数越多，其平台忠诚度就越高，反之则越低。本节选取 JDC 集团跨境电子商务平台 2020 上半年的用户复购数据，其复购率变化趋势如图 4.10 所示，由图可知，2020 年 H1 平台用户复购率呈现出先上升后下降再上升的趋势，其中在三月份上涨至 10.67% 的第一个峰值后，又在五月下降至 7.74% 的最低点，随后到了六月大促月又提高至 16.6%，与五月相比有较大涨幅。该数据表明除了大促月之外，平台用户复购水平整体偏低，用户产生的重复购买行为较少，只有在优惠力度较大的活动期间才会增加重复消费。而三月复购率达到第一个峰值的原因在于 JDC 集团海外用户在三月大量、集中购买了口罩、防护服等防疫物资，因而拉动了用户复购率增长。而五月较低的复购率水平是由于当用户在五月有购买需求时，往往会等到活动力度较大的“618 大促”再进行购买，该消费心理使得平台五月的复购率水平在 2020 年 H1 中达到最低，该趋势也能从侧面印证平台用户的促销敏感特征。因此，为有效提高复购率水平，平台还需从产品质量到售后服务的各个环节改善用户体验从而促使其在大促月以外的时间增加重复购买行为以实现平台交易总额的持续快速增长。

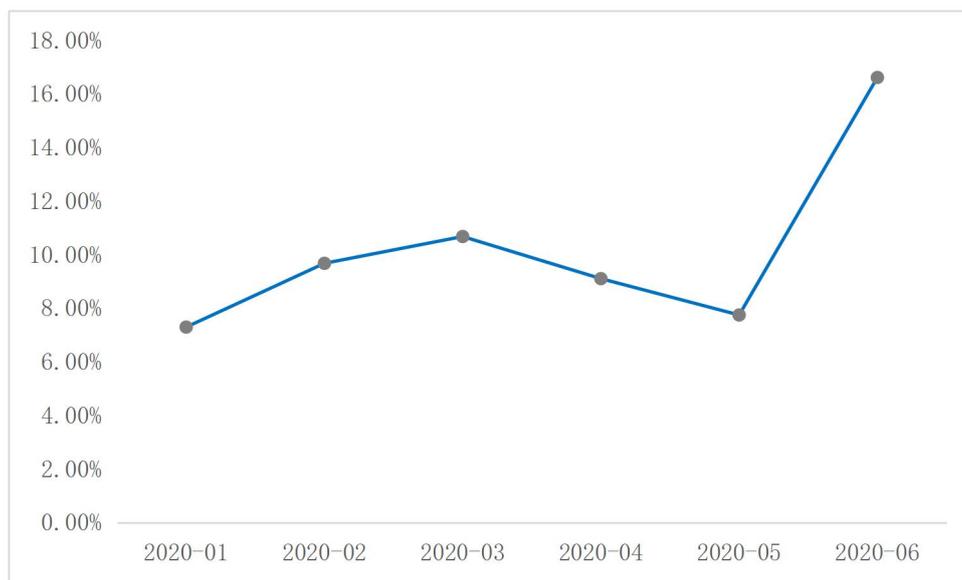


图 4.10 2020 年 H1 用户复购率变化趋势

2、用户留存

在电商平台发展过程中，若用户在某段时间内在某平台产生购买行为，且经过一段时间后，仍在使用该平台，那么该类用户即为留存用户，这部分用户与最初基础用户量的比值被称为留存率，而留存率又常被用于衡量平台质量和保留用户的能力。本部分内容选取 2020 年 1—6 月分别作为起始期与留存期来计算各月留存率，2020 上半年平台用户月度留存情况如图 4.11 所示，由图可知，当 1 月—5 月为起始期、6 月为留存期时，对应的用户留存水平相对较高，留存率分别为 26%、28%、27%、32%、41%。其中，以 5 月为起始期、6 月为留存期的留存率最高，该数据含义为 5 月的基础用户在 6 月仍然留下来产生下单行为的占比为 41%，究其原因是由于六月“618”大促活动宣传及预热吸引了较多在 5 月有消费行为的用户再次产生购买行为。另外，除 6 月以外，其他月份在作为留存期时，留存率大多处于 30%以下的水平，说明平台用户留存水平有待提高。

留存期	起始期					
	2020-01	2020-02	2020-03	2020-04	2020-05	2020-06
2020-01	100%					
2020-02	22%	100%				
2020-03	24%	33%	100%			
2020-04	20%	25%	26%	100%		
2020-05	19%	22%	21%	29%	100%	
2020-06	26%	28%	27%	32%	41%	100%

图 4.11 2020 上半年用户月度留存率

3、新老用户占比情况

根据目前电子商务发展现状来看，于各企业平台而言，规模扩展是第一要务。而扩大规模通常从获取平台流量和推动企业销售额增长两个维度协同推进。其中，获取平台流量是指新用户首次注册登录平台时，会增加大量访问、点击、浏览等各类型流量数据，因而平台流量的规模性增长通常需要对新用户进行抓取来获得。而推动企业销售额增长是指要促使平台销售额能够成规模增长，此时则需将目标对象转向获取成本低、购买转化高的平台已有用户，通过深度挖掘用户价值，拉动平台销售额增长。

JDC 集团跨境电子商务平台 2020 上半年新老用户占比及用户总量趋势如图 4.12 所示，由图可知，从新老用户占比整体趋势来看，老用户占比高于新用户，新老用户每月占比变化幅度较小。另外，用户总量变化动态呈现出的先上升后下降再上升的趋势与用

户复购率变化趋势在整体上保持一致，从而可进一步说明每月有消费记录的总用户量在很大程度上取决于老用户复购情况。结合以上数据及用户复购率发现，平台老用户占比虽高于新用户，但复购率水平较低，进而使得复购用户对拉动平台交易额的贡献度不够。另外，由于目前电商企业越发昂贵的推广成本导致企业发展新用户的难度也随之增加，新用户增长速率不断放缓。因此，平台在进行营销策略规划时，要做到新老用户两手抓，以新用户扩大流量规模，以老用户提升销售效果，做到规模与效果的统一。

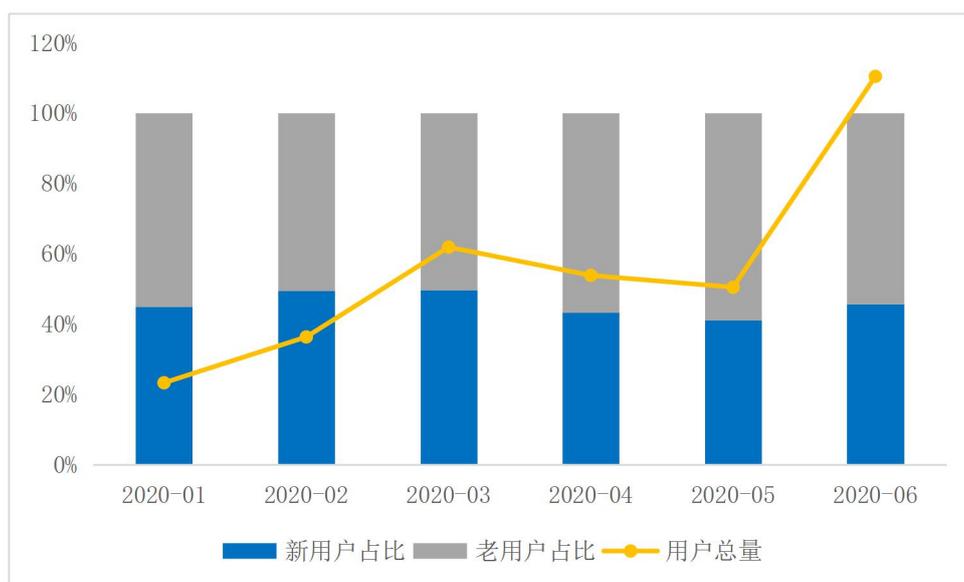


图 4.12 2020 上半年新用户量占比

4.3 用户消费偏好分析

在营销研究的学术领域内，用户行为、心理研究一直是极为重要的组成部分，不同用户会产生不同消费行为及消费态度，也就是说用户在消费偏好方面存在明显差异。消费偏好的概念可界定为：用户在交易过程中，对交易过程及结果的某种习惯性倾向^[9]。在企业经营的实践过程中，精确识别用户行为特征，并基于用户消费偏好进行商业决策是一个企业能否具备市场竞争力的关键。而且通过前两节对用户基本特征及用户价值的分析发现，用户消费行为由于其消费偏好的不同存在较大差异。因此，为进一步挖掘不同用户消费行为及其偏好的具体差异表现，本节将从商品品类、品牌及页面点击渠道三方面对 JDC 集团跨境电子商务平台的用户偏好展开深入剖析。

4.3.1 用户商品品类偏好

由于不同用户的品类需求及偏好存在较大差异，因而企业在运营过程中不仅要整体把握用户品类消费情况，还要对不同性别、年龄的用户进行消费倾向分析，并根据用户品类偏好分析结果进一步对用户偏好 TOP1 的品类进行更为详细的渗透率分析。基于此，本节选取 JDC 集团跨境电子商务平台 2018.1—2020.6 期间的用户画像数据对用户品类偏好进行分析。

1、用户品类偏好整体情况分析

由图 4.13 所示的用户品类偏好整体情况来看，根据各品类销量占比排名可得到平台用户的品类偏好 TOP10 分别为：图书、食品饮料、电脑/办公、母婴、厨具、手机通讯、数码、家庭清洁/纸品、家用电器、服饰内衣。其中，图书品类销量占比远大于其他品类的主要原因为：在该数据统计时段内，平台曾开展“图书超级品类日”、“618 图书节大促”、“双十一图书百亿补贴”、“购图书、贴邮费”等优惠力度较大的促销活动，吸引了大批用户产生购买行为。此外，据各品类 GMV 占比及客单价趋势能进一步发现，各品类客单价与 GMV 总值的发展趋势在整体上保持一致，如，图书品类销量虽排名第一，远高于电脑/办公，但就二者 GMV 值占比来看，电脑/办公的 GMV 占比却远超前图书品类，其原因在于电脑/办公的客单价较高，从而使该品类的交易总额较高。因此，为有效增加平台交易总额，除了首先要考虑消费者偏好的品类之外，还应充分考虑客单价偏高、对 GMV 总值有较大贡献的品类，并针对不同消费偏好的用户群体设计相应营销方案以吸引用户下单，拉动交易额增长。

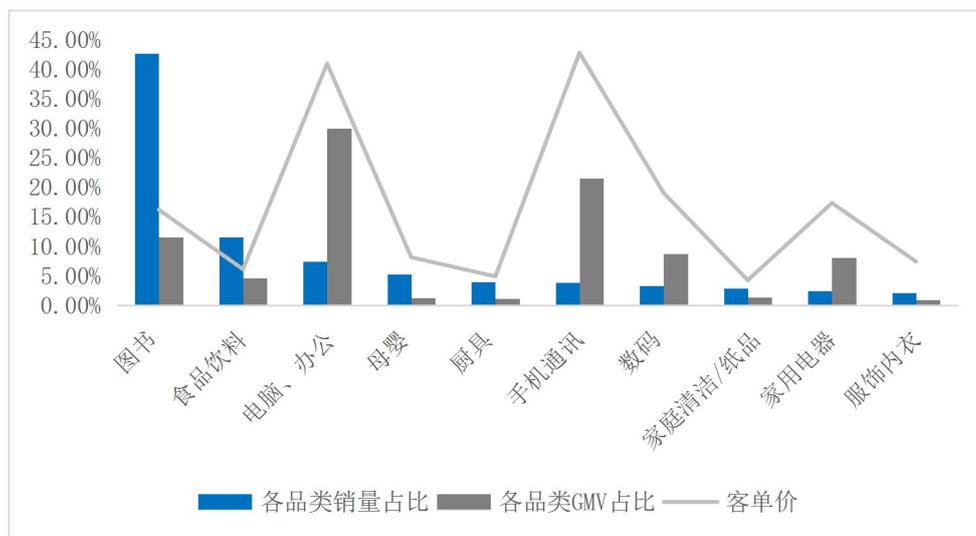


图 4.13 用户品类偏好 TOP10

2、不同性别的品类消费倾向分析

图 4.14 所示为不同性别用户的品类偏好差异，由图可知，除了大促活动多、优惠力度大的图书品类在男女用户偏好中均位居第一之外，可明显看出不同性别用户对品类偏好的较大差异。就除去图书品类的品类排行来看，男性偏好的主要品类为电脑/办公、食品饮料、手机通讯、母婴、数码，而女性偏好的主要品类为食品饮料、母婴、电脑/办公、厨具、家庭清洁/纸品。通过不同性别的品类购买偏好可看出男性与女性在生活中的关注侧重点不同，男性更关注电子数码类产品，而女性则更偏爱于购买居家日用类生活品。

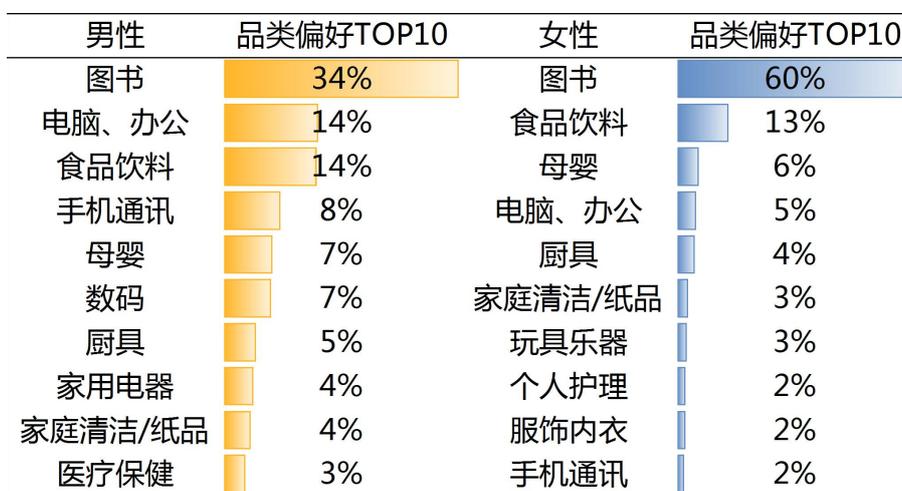


图 4.14 不同性别的用户品类偏好

3、不同性别、年龄组合的品类消费倾向分析

据前文分析可知，处于 26-35 岁年龄阶段的用户消费最多，36-45 岁、16-25 岁次之，故下文将重点分析 16-45 岁年龄范围用户的品类消费倾向。此外，由于图书品类销量受大促影响从而在各年龄段的品类销量占比中均排第一。因此，为更清晰直观分析其他品类在不同年龄、性别组合下的差异性，下文将对剔除图书的品类销量占比进行具体分析，并分别计算排在 TOP10 的品类在前十名品类总销量中的占比。图 4.15、4.16、4.17 分别显示 16-25 岁、26-35 岁、36-45 岁男性用户和女性用户的品类偏好。

(1) 男性用户在不同年龄的品类偏好纵向对比分析

由图 4.15、4.16、4.17 可知，16-25、26-35、36-45 岁这三个年龄段的男性用户仍然以电子数码类及食品饮料类产品为主要偏好，母婴、厨具类等居家用品次之。从三个不同年龄段的品类偏好变化来看，16-25 岁的男性用户在该年龄阶段正处于求学或初入社

会的状态，大多数购买行为是为了满足自身消费，因而在电子数码及食品饮料类的购买占较大比重，在居家类产品上的购买较少。接着到了 26-35 岁，此时大多数男性用户已进入养家、育儿等成家状态，家庭负担较重，此时用户的家庭消费属性逐渐凸显。由图 4.16 可知该年龄段男性用户的电子数码类产品购买占比与 16-25 岁相比有明显下降，同时，母婴、家用电器等居家生活类产品购买比重有所上升，不同品类的销量占比变动也反映出男性用户在该年龄段更为注重家庭，并减少了满足对自身喜好的产品购买。最后到了 36-45 岁，此时的男性用户可能处于事业稳定、养育儿女的状态。由图 4.17 可知，36-45 岁男性用户的品类销量占比变化为：电脑、办公品类的购买量持续下降，并在品类销量占比排行中不再居于榜首，母婴类销量持续增长，而手机通讯、数码类产品与 26-35 岁的男性用户相比购买比重有所上升，并新增运动户外品类的消费，由此可大致判断出 36-45 岁的男性用户在关注家庭生活的同时，也会适当满足个人消费需求。

（2）女性用户在不同年龄的品类偏好纵向对比分析

由图 4.15、4.16、4.17 可知，女性用户在 16-25、26-35、36-45 岁这三个年龄段内的品类偏好大多集中在食品饮料、居家生活、母婴及个人护理类。从三个年龄段的品类偏好变化趋势看，16-25 岁女性用户可能处于求学中或是初入社会没有过多家庭负担的状态，因而该年龄段女性用户的购买行为大多也是为了满足自身消费需求。由图 4.15 可知，16-25 岁女性用户在食品饮料品类的购买占比过半，其余的购买主要分布在个人护理护肤、电脑办公、母婴以及居家类产品上。到了 26-35 岁，女性用户的品类偏好发生明显变动。该阶段女性可能多数处于将生活重心放在家庭的状态，此时，该年龄段女性的购买行为不再仅是为了满足自身消费需求，而更多是为家庭消费。由图 4.16 可知，与 16-25 岁女性用户相比，该年龄段用户除了在母婴品类上的消费有大幅增加外，厨具、家庭清洁/纸品及家用电器的品类销量也有一定幅度增加，而个人护理类销量占比有所下降。另外，品类销量占比 TOP10 中新增了玩具乐器、服饰内衣两个品类，而家居日用、美妆护肤则跌出 TOP10 排名。各品类销售占比的一系列变动表明，26-35 岁的女性用户多数以家庭为重，不再将个人喜好放在首位。最后到了 36-45 岁，该年龄段用户可能处于同时兼顾事业和家庭的状态，此时的女性用户在购买居家生活用品之外，还会满足自己的工作消费需求。由图 4.17 可知，与 26-35 岁的女性用户相比，36-45 岁的女性用户在品类销量占比上变动最大的即为电脑、办公品类，其销量占比有明显增加，另外，在品类销量占比 TOP10 中还新增了医疗保健和家居日用。上述品类销量变动表明：该阶段女性用户不再将全部精力放在家庭上，而是将生活部分重心转移至工作中。

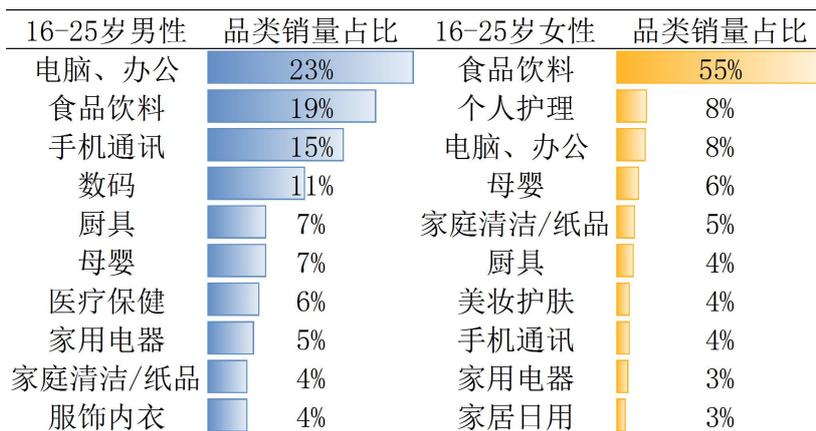


图 4.15 16-25 岁的男性、女性用户品类偏好

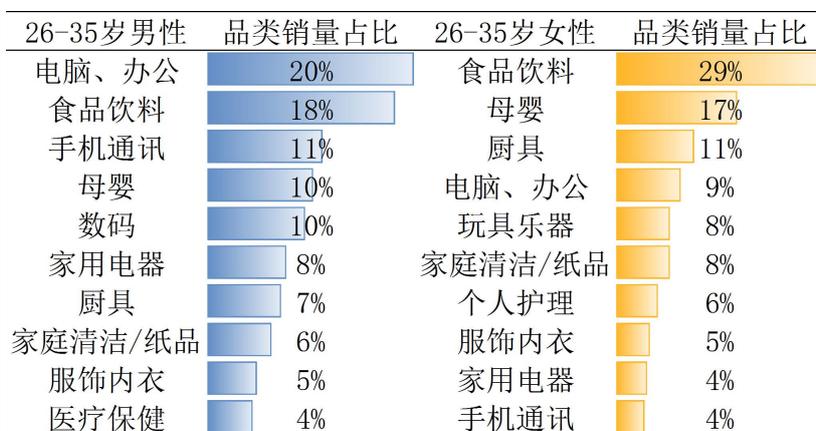


图 4.16 26-35 岁的男性、女性用户品类偏好

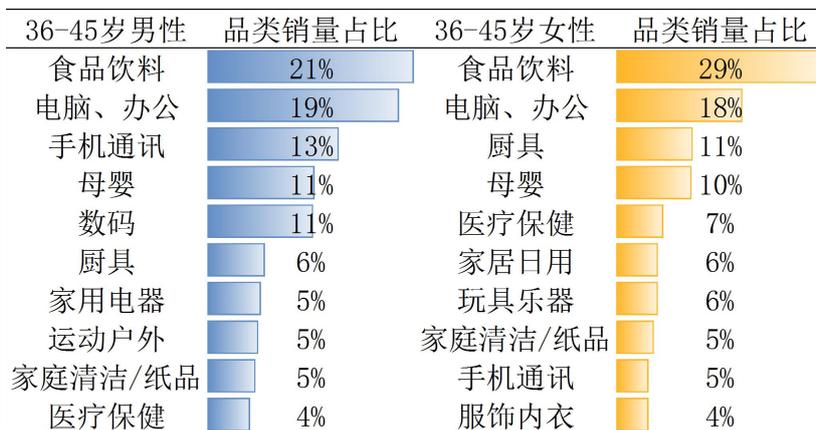


图 4.17 36-45 岁的男性、女性用户品类偏好

4、品类喜好 TOP1 细分类目渗透率分析

据上文对品类偏好的整体分析可知，图书类目销量最高，位居品类销量占比榜首。而图书类目又可细分为多个二级类目，若要探究图书类目高销量的真正原因，还需进一

步剖析图书各二级类目的销量占比。因此，为分析图书各二级类目的销售情况，现对其渗透率进行研究，渗透率计算公式： $\text{细分品类渗透率} = \text{细分品类用户消费人数} / \text{平台用户总人数}$ 。图 4.18 所示为图书类目下渗透率排名 TOP10 的各二级类目，由图 4.18 可知，童书类图书渗透率远高于其他各二级类目，居于第一，中小学教辅类次之，小说类排在第三，其余的各二级类目渗透率均较低。童书类及中小学教辅类图书的排名靠前说明平台购买图书的主体客群为“宝爸宝妈”，加之图书品类销量高的原因在于有过多次优惠力度较大的促销活动从而吸引了大量用户下单消费，由此可推断出“宝爸宝妈”这一消费客群对促销活动较为敏感。另外，据调研发现，JDC 集团跨境电子商务平台的用户交易主体并不是位于世界各国的外国人，而是海外华人及海外留学生，中小学教辅类图书的较高渗透率也可侧面印证该结论，这是由于大多数海外华人在对孩子的教育上仍保留较为传统的中式教育观念，因而会在教辅资料类书籍上有较多消费。

由上述分析可知，细化品类的渗透率分析能帮助企业更加直观了解特定群体的消费需求及偏好。因此，企业应根据用户消费偏好、特征，深挖各个群体在不同消费场景下的消费需求，从而及时调整品类策略以不断提高用户购买转化率。

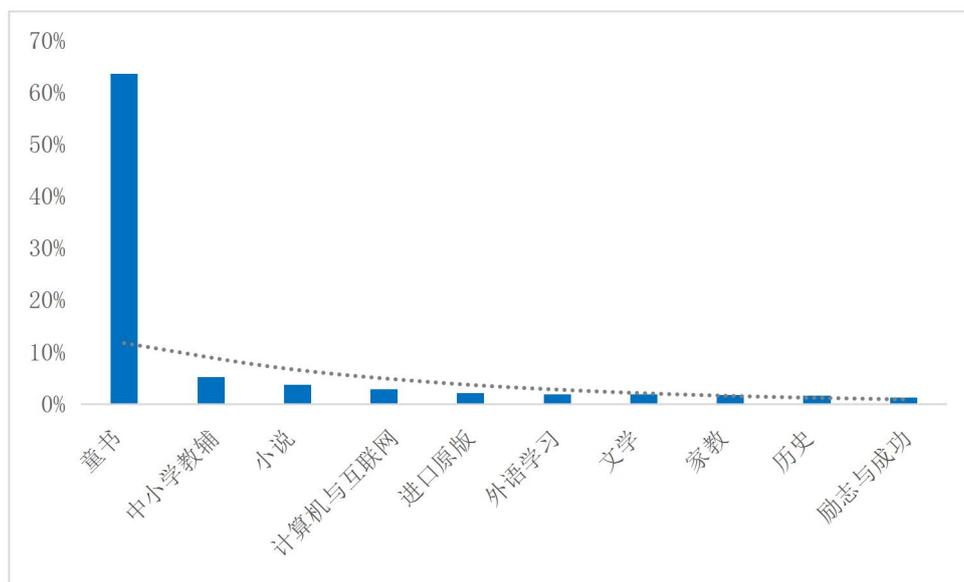


图 4.18 图书品类的细分类目渗透率

4.3.2 用户品牌偏好

近年来，随着人们生活水平不断提高，多数人在线上购物时已不再仅限于满足自身

商品购买需求，更多的是追求品牌，即除了首要关注产品性能以外，最看重品牌，甚至有部分消费群体会在网购过程中逐渐发展成为品牌追随者。品牌是一种标识、理念，是产品品质优异的体现。从当下市场竞争现状来看，品牌定位决定企业市场定位，品牌价值决定企业市场竞争力。因此，深入剖析用户品牌偏好不仅有助于企业把握不同用户品牌需求及品牌忠诚度，还能使企业对自身发展有清晰界定。为此，本节选取 JDC 集团跨境电商电子商务平台品牌消费记录的历史数据（2017.1.1—2020.6.1）从中国品牌与海外品牌的用户偏好差异以及各品牌购买地区分布两方面对用户品牌偏好展开分析。

1、中国品牌与海外品牌的偏好差异

据前文分析可知，平台消费客群以中国港澳台用户、海外华人、海外留学生为主，世界各地的外国人次之。图 4.19 所示为平台用户消费的中国品牌与海外品牌占比，由图可知，中国品牌的消费占比要远高于海外品牌，故平台用户更倾向于购买中国品牌。究其原因有以下两点：（1）国货意识。国货意识是指某个国家消费者出于对自己国家或民族的支持热爱以及对国外品牌可能会对本国利益造成一定伤害的顾虑，而对源自本国企业品牌的认同和偏好^[37]，国货意识换句话说也就是本国消费者的民族情感在消费观念上的映射，且该意识一旦形成，将不会轻易消失，并可能长期影响消费者对于本土品牌及外来品牌的态度及偏好。由于留学生及海外华人为平台主要消费群体，会受到国货意识及品牌消费习惯影响，因而多数华人在中国在线购物平台上会将中国品牌作为首选进行消费购买。（2）在华外国人、海外华人对中国品牌的传播影响作用^[38]。消费绝不是一种孤立行为，中国品牌的跨文化交流传播是中国企业与世界各国地区用户之间建立关系的过程，而在这两者之间搭建桥梁的中间人即为在华外国人及海外华人。这两类群体在接受中国背景下的各种中国品牌信息之后再向国外受众群体进行传播，将接收到的“信息”进行转换，并输出为对其他消费者的“影响”。在这两种“桥梁人群”中，在华外国人则是更为天然的品牌文化“解码者”。由于在中国品牌跨文化传播过程中，中国企业与外国受众群体并不处于同一文化体系。因此，两者之间不可避免的存在源于文化隔阂而导致的品牌文化排斥，而减少文化因素干扰的最佳方法是使传播者与受众者处于同一文化环境，如此一来，传播者的信息输出对受众者产生的影响也将更有说服力，中国品牌信息才能更容易被国外受众群体接收。

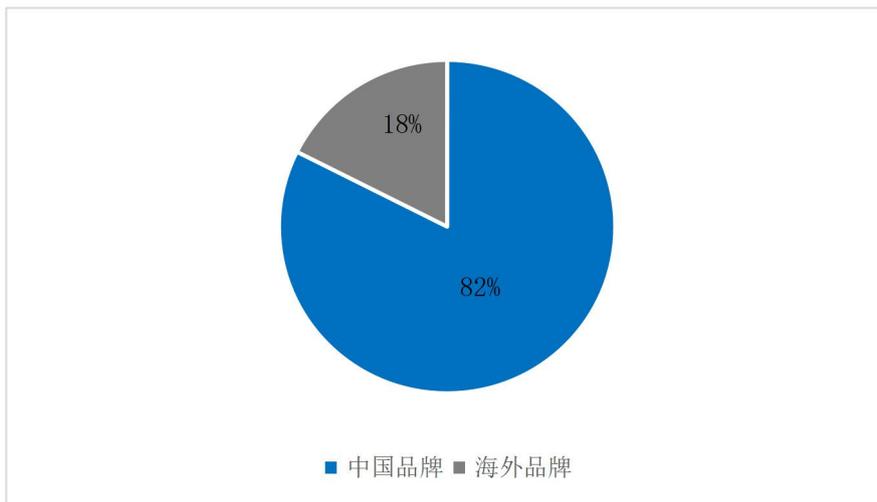


图 4.19 中国品牌与海外品牌的销量占比

2、品牌消费的地区分布

JDC 集团跨境电子商务平台用户的品牌消费地区主要分布在中国港澳台、美国、日本、澳大利亚、新加坡共 104 个国家和地区，且由于排名靠后的几个国家品牌消费量较少，故本文只选取品牌消费数量 TOP15 的国家和地区进行分析。由图 4.20 可知，中国香港的品牌消费数量在 104 个国家和地区中排名第一，销量占比最高，美国、中国台湾、中国澳门、日本次之，故下文将主要从中国香港、美国、中国台湾、中国澳门、日本五个国家和地区入手，具体分析不同地区用户的品牌偏好。据前文分析了解到，目前，平台已开通中国港澳台站、欧美站、东南亚站等多个独立站点，且处于起步发展阶段，因而深入剖析各个地区用户的品牌偏好也有助于平台各站点更加精准掌握其所在区域的用户消费偏好，从而基于用户品牌偏好制定相应品牌策略。

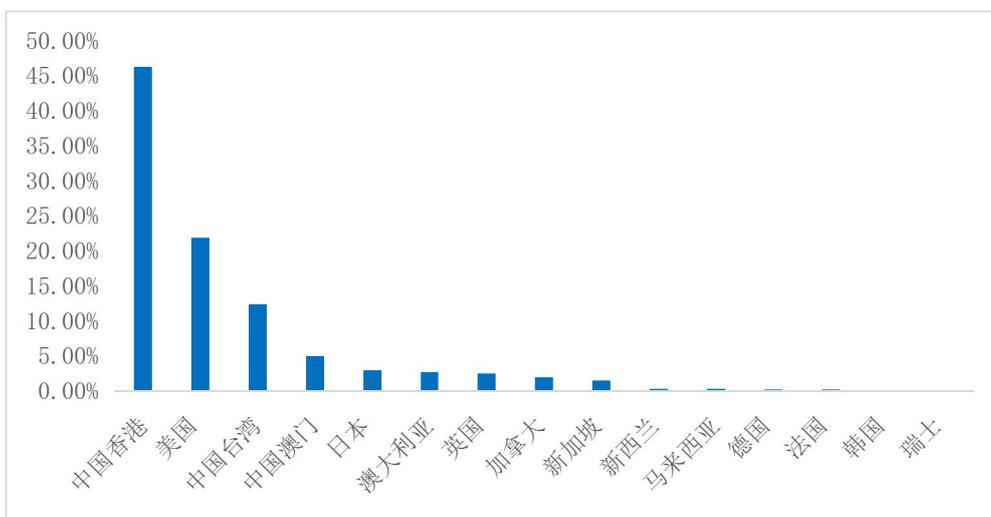


图 4.20 不同国家和地区品牌消费数量占比排名

由图 4.21 可知，中国香港地区用户的品牌购买偏好以食品饮料类为主，居家日用类和手机、电子数码类次之。因此，集团在进行中国香港站点的运营过程中，不仅要时常开展“零食节”、“美食大牌活动日”等与食品饮料类品牌相关的活动，手机数码、家用电器以及居家日用类品牌等专场活动也要随之进行。美国地区用户的主要品牌偏好为图书类品牌，因而平台可针对美国消费群体将营销侧重点放在图书类品牌的专场促销活动中，开展一些如“图书大牌日”、“图书节”、“童书小专场”之类的图书品牌活动，以促进该地区用户群体增加消费。中国台湾地区用户的品牌偏好集中在手机、电子数码类以及居家日用类品牌上，其中手机、电子数码类品牌为其主要偏好。为此，中国台湾站点的日常运营应经常开展“手机大促节”、“3C 小家电品牌日”等关于 3C 数码、家用电器品牌的促销活动。中国澳门地区及日本的站点运营同理，总之就是要设计出与用户品牌偏好相一致的促销活动，以促进该地区用户消费。

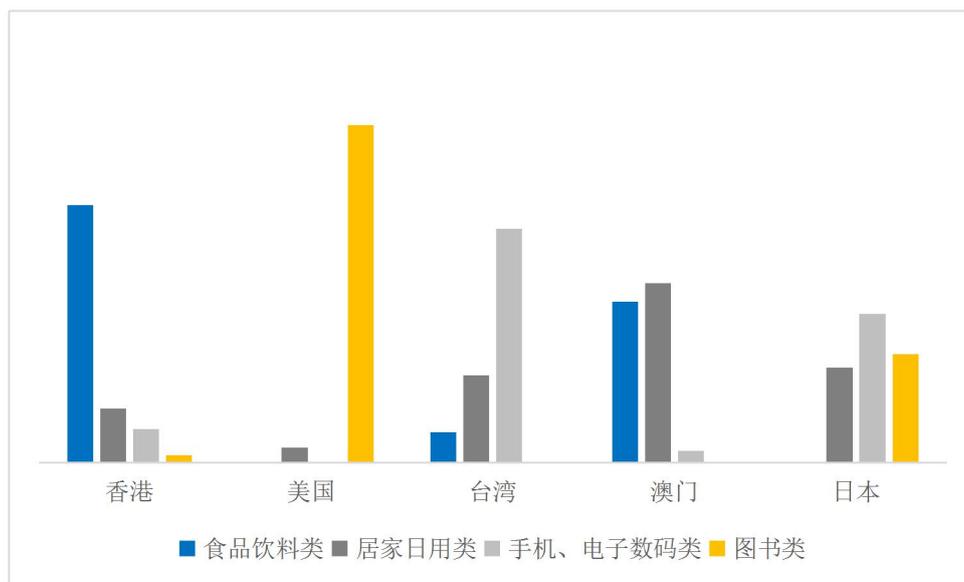


图 4.21 不同国家和地区的品牌偏好情况

4.3.3 用户页面点击渠道偏好

前两部分内容分别从品类及品牌入手分析用户消费偏好，得出的分析结论可以使平台对各类用户的商品购买偏好有大致了解。但随着跨境电子商务平台及网络技术的飞速发展，于商家而言，除了掌握用户特征并了解其商品购买偏好之外，研究用户通过哪种渠道触达到商品或活动页面同样至关重要，也就是说不仅要明确用户是谁，还要关注用

户产生的流量从哪来。因此，本节选取 2018.1—2020.6 期间平台流量数据，从页面点击渠道的访问次数占比以及页面来源的流量转化率两方面入手，归纳 JDC 集团跨境电子商务平台主要页面来源，并进一步分析用户渠道偏好。

1、页面点击渠道的访问次数占比

平台各页面点击渠道的流量来源包括搜索栏、直接流量、商品详情、下单界面、店铺、平台首页、公共频道等在内的共 38 个页面来源分类。本节内容选取平台页面来源历史数据，并将各页面来源划分为搜索栏、直接流量、频道活动型页面、商品购买型页面共四大类展开剖析。其中，（1）搜索栏。搜索栏是指用户在平台搜索栏输入商品关键词以检索某商品信息而产生的流量，偏好于该渠道的用户属于有特定商品需求且目的性较强的消费类型；（2）直接流量。直接流量主要包括用户在浏览器直接输入网址产生的流量、用户点击收藏网页产生的流量和丢失来源信息的流量三部分，一般是 PC 端用户最偏好的触达方式。（3）频道活动型页面。偏好于该类页面并访问较多的用户一般并无特定购物目的，而是通过访问频道页面以增加对某些商品的购买意愿。（4）商品购买型页面。经常访问、点击该类页面渠道的用户一般有较多下单行为，且通常会在浏览某个商品、店铺后再进行购买，与偏好于其他消费渠道的用户相比，该类用户从浏览商品到最终产生下单购买行为的转化率较高，因而在制定用户营销策略时，应格外重视该类购买转化率较高的用户类型。

图 4.22 所示即为平台主要页面来源分类的访问次数占比，由图可知，商品购买型页面访次占比最高，搜索栏、频道活动型页面占比次之，直接流量占比最低，该数据表明用户在进入平台页面后最偏向于点击商品购买型页面，说明目前 JDC 集团已具有一定规模的活跃用户倾向于在平台产生消费行为。另外，搜索栏访次占比高也能突显出平台用户购买目的较强的消费特征。而直接流量占比低则说明 PC 端用户相对较少，这也与前文端口偏好分析中得出的 APP 端是平台用户最为偏好的渠道结论相对应。

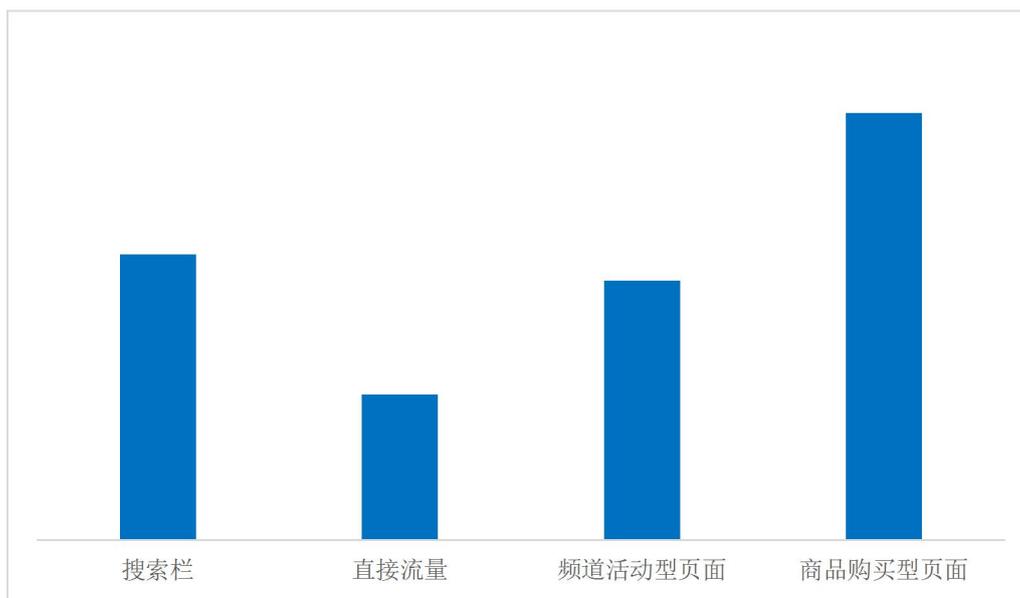


图 4.22 页面来源分类的访问次数占比

2、页面来源的流量转化率

通过上文对页面来源访问次数占比的分析可知平台用户的页面渠道访问偏好，因而本节内容将基于上文结论进一步剖析用户在各个主要点击渠道的访次流量转化率，以便更加直观了解用户从最初进入的平台界面与最后触达的订单提交页面之间各环节的流量转化。下面将从直接触达页面、搜索栏、商品详情、下单界面四个环节逐级分析流量转化率，即：直接触达页面——搜索栏、搜索栏——商品详情、商品详情——下单界面、直接触达页面——下单界面。其中，直接触达页面主要包括平台首页等频道活动型页面及直接流量，商品详情和下单界面即为商品购买型页面中的商品详情页和订单提交页，具体数据如图 4.23 的转化率漏斗图所示。由图可知，直接触达页面——搜索栏的转化率为 71.82%，该数据说明用户在进入平台各频道首页后偏向于使用搜索栏进行商品搜索，因而直接触达页面到搜索栏的访次转化率较高；搜索栏——商品详情的转化率为 42.74%，该数据表明用户在搜索栏产生商品搜索行为后仅有不到一半的比率会点击商品详情页查看该商品。因此，如何使搜索结果页呈现出的商品推荐更能与用户购买需求相匹配就显得尤为重要；商品详情——订单提交页的转化率为 27.13%，该数据含义为用户在访问商品详情页，对产品有进一步了解后仅有不到 30%的比率会下单购买，而此环节偏低的转化率也说明商详页信息在影响用户最终产生下单行为时缺乏说服力，因而在设计规划商详页时应将主图、细节图、对比图、产品规格尺码表、适当的买家秀图、售后保障等内容考虑在内，以增加促使用户下单的说服力；最后，直接触达页面——订单

提交页的转化率仅为 8.33%，即用户从一开始访问平台页面到最后只有 8.33% 的比率产生了下单行为。而较低的转化率数据也反映出在直接触达页面与下单界面之间的中间环节有较多用户流失，因而如何提高用户从直接触达页面到订单提交页的购买转化率，并不断减少中间环节的用户访问流失对于平台发展而言至关重要。

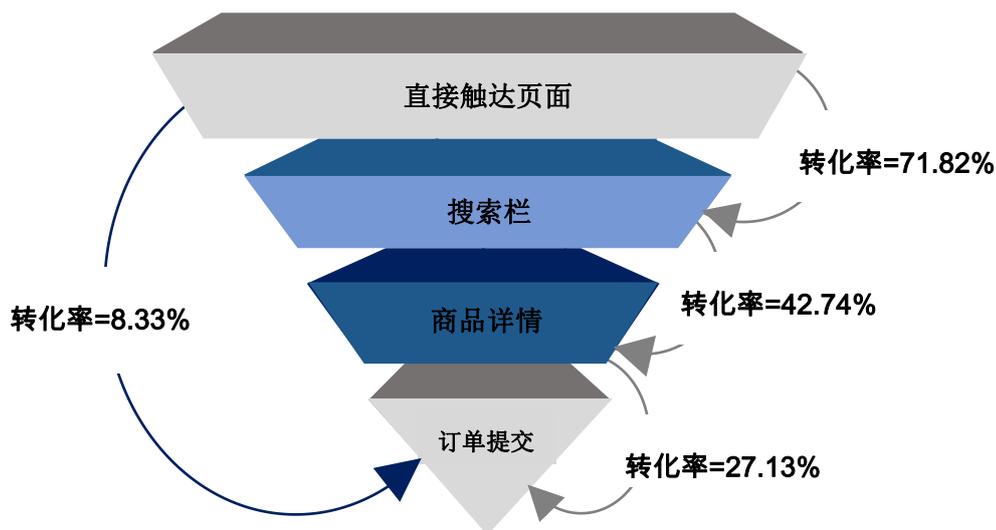


图 4.23 页面来源转化率漏斗图

4.4 本章小结

本章主要从用户基本特征、用户价值、用户偏好三方面对 JDC 集团跨境电子商务平台目标客群消费特征展开深入剖析。首先，在用户基本特征分析中归纳总结出平台海外客群的消费特征分别为：女性、26-35 岁、已婚、大学本科、中国香港、偏好使用 APP 端、常使用微信端支付、偏向使用优惠券。接着，文章又从用户忠诚度及消费能力和用户复购、留存、新老用户占比情况两方面入手，深挖平台用户价值，得出目前平台用户忠诚度偏低、消费能力处于中下等水平、低价值用户占比大，并存在平台用户活跃度低、用户留存现状有待改善等现实问题。最后，又分别对用户品类、品牌以及页面点击渠道偏好展开重点研究。其中，在品类偏好分析中，发现不同性别、年龄组合的用户品类偏好差异较大；在品牌偏好分析中，得出平台用户更倾向于中国品牌，且各地区用户品牌偏好也存在较大差异；在页面点击渠道偏好分析中，得出商品购买型页面点击率最高，但从平台最初触达页面到最终下单页面之间存在较多用户流失的结论。

5 基于机器学习算法的预测模型构建

据上文研究可知,本文基于各层面交易数据对用户基本特征、用户价值、用户偏好等多方面消费特点展开分析,从而对不同用户群体的个性化需求已有初步了解,并对平台目标客群也有了较为精准的定位。因此,为更进一步将上述研究成果落实,本节将从用户购买行为入手构建预测模型,将用户消费特征定位与用户购买行为预测模型相结合,构成一套基于数据且完备、专业的数据库营销体系,以达到有效吸引新用户消费,促使老用户复购,不断提高用户留存率的营销效果。另外,本文在机器学习算法模型上选择 GBDT 算法,选择该算法的原因有如下三点:(1) GBDT 算法不仅可以处理包括离散型和连续型数据在内的多种数据类型,且调参时间短;(2)由于本文模型构建部分使用的原始数据是集团平台真实交易数据,故该数据集异常值较多。而 GBDT 算法有多种损失函数,对异常值的鲁棒性较强,使得预测精度也相对较高;(3) GBDT 算法能够处理非线性数据,既能做回归也能做分类,并且对分类变量的包容性更强。因此,基于以上三点理由,本文选取 GBDT 机器学习算法进行用户购买行为预测的模型构建。

5.1 GBDT 模型

5.1.1 GBDT 算法概述

GBDT 算法利用加法模型,通过不断拟合模型训练过程中产生的残差来对数据集进行回归或分类。由于 GBDT 算法是以决策树为基础的集成学习算法,故本节首先对决策树算法进行说明。

1、决策树算法

决策树是一种常见的监督式机器学习算法,是一种树形结构,其中每个内部节点表示对一个属性的判断,每个分支表示一个判断结果的输出,最后每个叶节点代表一种分类结果^[44]。决策树既可以用于分类问题,也能处理回归问题,并主要包括 ID3、C4.5、CART 三种经典算法,本文主要介绍 CART 算法。

CART 算法根据具体研究问题的不同既可以用于构建分类树做分类,也可以用于构建回归树做拟合,而 CART 决策树构建的重点是如何选择各分支的分支点,从而通过对分支属性的选择提高数据纯度^[30]。若用于生成分类树时,通常计算 GINI 值衡量节点纯度,若用于生成回归树时,则采用样本方差衡量节点纯度,具体计算公式如下:

当生成分类树时，GINI 值表达式如式 (5.1) 所示：

$$GINI(D) = 1 - \sum_{i=1}^M P_i^2 \quad (5.1)$$

其中， M 为类标签数目， P_i 为类别 i 在样本 D 中出现的概率，GINI 值越大代表节点纯度越低，GINI 值越小则代表节点纯度越高。

当生成回归树时，回归方差表达式如式 (5.2) 所示。若生成回归树的方差越大，说明节点数据越分散，方差越小，则说明节点数据越集中。

$$\sigma = \sqrt{\sum_{i \in I} (x_i - \mu)^2} = \sqrt{\sum_{i \in I} x_i^2 - n\mu^2} \quad (5.2)$$

因此，为有效降低误差，CART 算法还需选择使节点 GINI 值最小或回归方差最小的属性分别作为分类树和回归树的分支属性，故进一步引入两者的最小化公式：

当构建分类树时，最小化公式如式 (5.3) 所示：

$$Gain = \sum_{i \in I} p_i \cdot GINI_i \quad (5.3)$$

当构建回归树时，最小化公式如式 (5.4) 所示：

$$Gain = \sum_{i \in I} \sigma_i \quad (5.4)$$

在选择好分支属性后，需进一步分割数据样本。针对连续型属性，CART 算法只需利用上述公式寻找最优属性并将其直接一分为二即可。而对于离散型属性，则是先将其中某个离散值单独作为一个节点，其他离散值作为另一节点，并通过计算每种划分方法的 GINI 值或方差来确定最优化分。

2、GBDT 算法

GBDT 算法又称为梯度提升决策树算法，其实质是由一系列决策树构成，每颗决策树之间通过迭代建立，即下一棵决策树是对前面所有树的结论及残差的学习并通过拟合得到的，利用加法模型将之前所有生成的树进行累加就是最终建立的 GBDT 模型。

GBDT 模型构建过程可由式 (5.5) 表示：

$$F_m(x) = F_{m-1}(x) + \arg \min_h \sum_{i=1}^n L(y_i, F_{m-1}(x_i) + h(x_i)) \quad (5.5)$$

其中， $F_m(x)$ 表示第 m 次学习模型， $F_{m-1}(x)$ 表示第 $m-1$ 次学习模型， $L(y_i, F_{m-1}(x_i))$ 为第 $m-1$ 次的损失函数， $h(x_i)$ 为第 m 次的学习器， $L(y_i, F_{m-1}(x_i) + h(x_i))$ 为第 m 次的损失函数， y_i 表示预测结果， x_i 表示训练数据集，每轮迭代的目的是找到一个学习器 $h(x_i)$

使得每轮的损失函数不断减小^[30]。

GBDT 算法在整个建模流程中共涉及三个关键点：（1）GBDT 用到的弱学习器为 CART 回归树；（2）GBDT 从数据角度看，运用了 Boosting 集成学习思想，即通过将弱学习器提升为强学习器的集成方法以提高预测精度；（3）GBDT 的弱学习器为决策树模型，是沿着之前模型损失函数的负梯度方向建立的，每个决策树模型的建立都是为了使损失函数不断减小以达到提高模型精度的目的。

5.1.2 GBDT 模型的参数说明

一般来说，常将 GBDT 模型的重要参数分为 Boosting 框架和基学习器两部分，Boosting 框架下包括 `n_estimators`、`learning_rate`、`subsample`、`init`、`loss` 等参数，基学习器即决策树则包括 `max_depth`、`min_samples_split`、`min_samples_leaf`、`max_features`、`max_leaf_nodes` 等参数。而 GBDT 调参过程主要与 `n_estimators`、`learning_rate`、`max_depth`、`loss` 紧密相关，其中，`n_estimators` 为弱学习器最大迭代次数，该参数值设定过大容易导致过拟合，设定过小容易导致欠拟合，因而需根据实际问题取较为适中的值，其调参过程常与 `learning_rate` 结合考虑；`learning_rate` 为每个弱学习器的权重缩减系数，又可称为步长，在 0 到 1 之间取值，该参数是为了防止过拟合而添加的正则项系数；`max_depth` 为决策树最大深度，默认可以不输入，在特征或数据量较少的情况下可忽略该指标，即决策树在建立子树时不限制子树深度，但在特征或数据量较多的情况下要对该指标进行设置调节，常在 [10,100] 区间取值；`loss` 为损失函数，GBDT 分类模型与回归模型的差异便主要体现在损失函数选择上，分类模型一般有对数似然损失函数（`deviance`）和指数损失函数（`exponential`），而回归模型则有均方差（`ls`）、绝对损失（`lad`）、Huber 损失（`huber`）和分位数损失（`quantile`）。

5.2 模型效果分析指标

5.2.1 混淆矩阵

对于本文涉及到的用户购买二分类预测问题，模型最终会输出两个结果，即 $x=0$ 表示用户未购买， $x=1$ 表示用户购买，因而可以根据模型的真实及预测值将预测结果分为四部分：真正例（TP, True Positive）、真负例（TN, True Negative）、假正例（FP, False

Positive)、假负例 (FN, False Negative), 其中, 真正例 TP 代表模型预测为正样本, 真实类别为正样本; 真负例 TN 代表模型预测为负样本, 真实类别为负样本; 假正例 FP 代表模型预测为正样本, 真实类别为负样本; 假负例 FN 代表模型预测为负样本, 真实类别为正样本^[32]。

混淆矩阵能够直观表示模型输出结果与模型预测效果的好坏, 其具体构成如表 5.1 所示。

表 5.1 混淆矩阵

真实类别 \ 预测类别	预测正类	预测负类
正类样本	TP (True Positive)	FN (False Negative)
负类样本	FP (False Positive)	TN (True Negative)

此外, 根据混淆矩阵中 TP、TN、FP、FN 四个值, 还能得到衡量模型分类准确性的指标 Accuracy, 如式 (5.6) 所示:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (5.6)$$

该指标称为分类准确性, 用来表示分类器整体分类性能好坏。但在实际应用中, 利用 Accuracy 评价分类准确性仅适用于类别均衡情况, 若出现类别不均衡情况, 该指标值对分类结果的判断并无意义, 如正样本过多, 负样本过少, 预测结果全部为正样本, Accuracy 值很大, 但精度却不高。因此, 还需结合其他评估指标对分类器准确性进行评价。

基于 TP、FP、TN、FN 这四个概念, 可进一步引入准确率 P (Precision)、召回率 R (Recall)、F1 (F1-Score) 这三个指标将混淆矩阵数字化, 其计算公式分别为:

$$P = \frac{TP}{TP + FP} \quad (5.7)$$

$$R = \frac{TP}{TP + FN} \quad (5.8)$$

$$F1 = \frac{2 * P * R}{P + R} \quad (5.9)$$

其中，准确率表示预测正确的正例在所有预测正例中的占比，即在预测为正例的样本中实际有多少真的正样本。召回率表示预测正确的正例在所有真正正例中的占比，即在原本所有正样本中有多少预测正确。由上述公式可知，若一个模型既有高准确率，也有高召回率，那么该模型分类预测效果较好。然而，在实际应用中，准确率和召回率此消彼长、相互制约，若要使其中一个值变大，就必然要损失另一个值。因此，同时满足高准确率和召回率往往无法实现。但在模型效果评估时，准确率和召回率二者都要兼顾，故使用调和平均的方法得到 F1-score 指标对模型分类效果进行综合评估，该指标既包含对模型准确性的衡量，又能评价样本找回效果，涵盖面更全，效果更好。因此，本文选取 F1-score 值来评价基于机器学习算法用户购买行为的模型预测效果。

5.2.2 ROC 曲线、AUC 指标

在引入 ROC 曲线、AUC 指标之前，首先对真阳性率（TPR）、假阳性率（FPR）两个指标进行说明，具体公式如下：

$$TPR = \frac{TP}{TP + FN} \quad (5.10)$$

$$FPR = \frac{FP}{FP + TN} \quad (5.11)$$

其中，TPR 表示正样本被正确预测的占比，FPR 表示负样本被预测错误的概率。而 ROC 曲线即是以 TPR 为纵坐标，FPR 为横坐标，将各个点连接起来形成的曲线，曲线越接近左上角说明模型预测效果越好。但曲线图往往只能看到大致趋势，并不能量化模型评估效果，因而，通常将 ROC 曲线下方围成的面积称为 AUC 值，用该指标对分类器性能的好坏进行数值化表示。AUC 值与模型分类性能成正比关系，AUC 的值越大，越接近 1，说明模型分类性能越好。

5.3 模型构建流程

5.3.1 数据来源及预处理

1、数据来源

本文基于机器学习算法，选取 JDC 集团跨境电子商务平台 2019 年“618 大促”活动预热期至 2020 年“6.1 开门红”活动期间（2019.5.20—2020.6.1）的真实交易数据构建用户购买行为预测模型。其中，数据类型主要包括：用户编号（user_id）、商品编号（item_id）、品类编号（cate_id）、用户的五种行为类型（加入购物车、收藏、浏览、下单、购买行为）共四个字段。

2、数据预处理

由于原始数据存在数据缺失、重复、异常值等情况不能直接用于建模，而数据质量高低将直接影响模型效果。因此，在对数据进行训练前，数据清洗必不可少。

（1）缺失值处理

缺失值处理即对原始数据字段中的缺失数据进行填补。对于缺失值填补，常有以下两种解决方法：一是若某字段数据缺失过多，则直接考虑删除该字段；二是以均值、中位数、众数来补齐缺失值，本文主要以均值填补数据缺失值。

（2）重复值处理

重复值处理即删除原始数据中的重复值。本文选取原始数据中的主要重复值类型为数据值完全相同，对该类数据直接删除即可。

（3）异常值处理

异常值处理即对原始数据中的异常用户订单进行筛选、剔除。异常用户筛选通常需结合该用户的历史行为来判断，如某用户有较多笔交易记录，但几乎没有浏览、点击、加购行为，那么可将该用户归为“刷单用户”类别。又或是某用户有大量浏览点击记录，但从未购买，那么该用户大概率情况下为“爬虫用户”，针对该类异常用户数据，可将其直接删除。本文将从以下三方面剔除异常值：（1）浏览次数大于 100、购买次数为 0 的爬虫用户订单；（2）加购次数为 0、收藏次数为 0、浏览次数为 1 且有购买记录的刷单用户订单；（3）系统异常值数据。由于 JDC 集团平台交易的下单和购买属于两个独立行为，即：下单但并未支付成功则只算下单行为，下单且支付成功才算购买行为，若存在购买行为，则必然先有下单行为，因而将下单次数为 0、购买次数不为 0 的数据归为系

统异常值，并剔除该部分数据。

5.3.2 样本选取

本文以数据统计周期内用户是否有购买行为作为正负样本的选取标准，将 2019.5.20—2020.6.1 期间有购买行为的用户标记为正样本，无购买行为的用户标记为负样本。对经过预处理的数据集进行划分共得到 41405 个正样本、82192 个负样本，正负样本比例约为 1:2。

5.3.3 特征变量构建及选择

1、特征变量构建

鉴于本文获取的原始数据类型及企业常规用户消费行为分析、经营模式分析等实践归纳出的业务经验，可将特征类型划分为用户维度、商品维度、品类维度、用户-商品维度、用户-品类维度、商品-品类维度六大模块，并从基础特征构建、组合特征构建、衍生特征构建三个角度出发，构建特征变量共计 54 个。

(1) 基础特征构建

由于原始数据有 user_id（用户编号）、item_id（商品编号）、cate_id（品类编号）三个具有标识的字段和加入购物车、收藏、浏览、下单、购买共五类用户历史购物行为，因而基础特征变量的构建将从用户、商品、品类三个维度与五类历史购物行为入手。第一，用户维度，构建的特征分别为每个用户在原始数据统计时段内的总加购次数、总收藏次数、总浏览次数、总下单次数、总购买次数以及五类历史购物行为的次数总和，该特征的构建目的在于对每个用户在统计时段内的平台活跃度进行考量；第二，商品维度，构建的特征分别为每件商品在统计时段内的总加购次数、总收藏次数、总浏览次数、总下单次数、总购买次数以及五类历史购物行为的次数总和，该特征的构建目的在于考察每件商品在统计时间内的热度；第三，品类维度，构建的特征分别为每种品类在统计时段内的总加购次数、总收藏次数、总浏览次数、总下单次数、总购买次数以及五类历史购物行为的次数总和，该特征的构建目的在于对每种品类在统计时间内的热度进行衡量，再加上用户-商品-品类的 6 个原始数据特征，基础特征共构建 24 个，具体构建结果如表 5.2 所示。

表 5.2 基础特征构建说明

维度	特征说明
用户-商品-品类	总加购次数、总收藏次数、总浏览次数、总下单次数、总购买次数以及五类历史购物行为的次数总和，共计 6 个
用户	每个用户在整个数据周期内的总加购次数、总收藏次数、总浏览次数、总下单次数、总购买次数以及五类历史购物行为的次数总和，共计 6 个
商品	每件商品在整个数据周期内的总加购次数、总收藏次数、总浏览次数、总下单次数、总购买次数以及五类历史购物行为的次数总和，共计 6 个
品类	每种品类在整个数据周期内的总加购次数、总收藏次数、总浏览次数、总下单次数、总购买次数以及五类历史购物行为的次数总和，共计 6 个

(2) 组合特征构建

基础特征构建完成后，将从用户-商品、用户-品类、商品-品类三个维度入手，进行组合特征构建。第一，用户-商品维度，构建的特征分别为每位用户在统计时段内加购、收藏、浏览、下单、购买每件商品的总次数，该特征的构建目的在于研究每个用户对每件商品的偏好及喜爱程度。第二，用户-品类维度，构建的特征分别为每位用户在统计时段内加购、收藏、浏览、下单、购买每种品类的总次数，该特征的构建目的在于研究每个用户对每种品类的偏好及喜爱程度。第三，商品-品类维度，构建的特征分别为每种品类下对应的每件商品在统计时段内加购、收藏、浏览、下单、购买总次数，该特征的构建目的在于研究每种品类下各件商品的热度。组合特征共计 18 个，具体构建结果如表 5.3 所示。

表 5.3 组合特征构建说明

维度	特征
用户-商品	每位用户在整个数据周期内的加购、收藏、浏览、下单、购买每件商品的总次数以及五类历史购物行为的次数总和，共计 6 个
用户-品类	每位用户在整个数据周期内的加购、收藏、浏览、下单、购买每种品类的总次数以及五类历史购物行为的次数总和，共计 6 个
商品-品类	每种品类对应的每件商品在整个数据周期内的加购、收藏、浏览、下单、购买总次数以及五类历史购物行为的次数总和，共计 6 个

(3) 衍生特征构建

在对用户、商品、品类两两组合进行特征构建后，还能就不同原始特征、组合特征进一步生成衍生特征。本文将从比值特征及转化率特征两方面构建衍生特征。

比值特征是指利用两个已有特征的比值来构建新特征，以便从侧面挖掘用户的潜在行为特点。本文将构建以下四种比值特征，第一，商品购买次数与该商品对应品类的总购买次数的比值，该特征用于考量各商品在其对应品类下的用户购买偏好程度；第二，用户购买总次数与用户五种行为加总次数的比值，该特征用于考量用户购物谨慎度，即在产生购买行为之前，其他几种行为的发生频次；第三，用户-商品维度购买次数与用户维度购买次数的比值，该特征用于研究每位用户对其购买过商品的偏好程度；第四，用户-品类维度购买次数与用户维度购买次数的比值，该特征用于研究每位用户对其购买过商品类型的偏好程度。

转化率是电商运营的核心指标，各大电商平台在对业务指标进行深度数据剖析时，尤其看重转化率指标的变化趋势，追求重点也在于提高各环节转化率，如商品或品类加购/收藏/浏览/下单-购买转化率，即用户对某件商品或某种品类加购/收藏/浏览/下单次数与购买次数的比值，该指标含义为每个用户发生加购、收藏、浏览、下单行为后最终转化为购买行为的可能性。衍生特征共计 12 个，具体构建结果如表 5.4 所示。

表 5.4 衍生特征构建说明

视角	特征
比值	$\frac{\text{商品购买次数}}{\text{该商品对应品类的总购买次数}}$
	$\frac{\text{用户购买总次数}}{\text{用户四种行为总计次数}}$
	$\frac{\text{用户 - 商品维度的购买次数}}{\text{用户维度的购买总次数}}$
	$\frac{\text{用户 - 品类维度的购买次数}}{\text{用户维度的购买总次数}}$
转化率	商品转化率：每件商品加购/收藏/浏览/下单-购买转化率，共计 4 个 品类转化率：每种品类加购/收藏/浏览/下单-购买转化率，共计 4 个

2、特征变量选择

随着物联网、大数据等信息技术的飞速发展，大量高维数据能够被简单快捷的获取存储。然而，包含全面信息的高维数据在给使用者带来充足信息量的同时也严重影响了机器学习算法的学习效率和性能，一方面，高维数据会降低机器学习算法的执行效率，另一方面，大量冗余和不相关特征也将给机器学习算法性能带来负面影响，存在机器无法捕捉关键信息的可能。因此，在利用某些机器学习算法对数据样本进行训练之前，筛选有用特征、剔除冗余特征对降低数据维度以及提升算法执行效率来说至关重要。而特征选择作为改善算法性能、完成数据降维的方法，可以有效剔除不相关特征，精准选择出与研究目标较为匹配的有用特征。

特征选择是指从一组特征中找出一定数量能够反映数据变化的最重要属性以达到降低特征空间维数目的的过程。在预测研究领域中，特征选择是十分关键的环节，特征变量选择的效果好坏也将直接影响分类器性能及模型预测准确性。常见的特征选择类型主要有三种：（1）过滤式，过滤式特征选择与后续学习器无关，是按照发散性或相关性对各个特征进行评分，并通过设定阈值选择特征；（2）封装式，封装式特征选择就是把特征选择当作一个特征子集搜索问题，根据最优模型结果选择最优特征子集；（3）嵌入式，嵌入式特征选择利用正则化思想，根据模型来分析特征重要性，该方法需将特征选择过程作为模型训练的一部分融入模型中^[25]。本文选取随机森林法进行特征选择。

随机森林是一种以决策树为基学习器的集成学习算法，该算法利用随机重采样技术和节点随机分裂技术构建多棵决策树，通过投票即“少数服从多数”原则来决定最终结果^[35]。随机森林为 Bagging 集成学习思想的代表性算法，其中，随机指随机选取样本以及随机选取特征。随机森林之所以能被广泛用于特征选择，是因为该算法在完成模型训练后，可对各特征进行重要性度量及排序，即筛选出相对重要的特征，以便进一步展开模型训练。随机森林特征重要性评估思想实际上就是观察每个特征在每棵决策树上做出的贡献值大小，将每个特征值在每棵决策树的贡献取平均值后，再根据该贡献值大小进行重要性比较，通常使用基尼指数（GINI index）及袋外数据（OOB）错误率两种指标来衡量不同特征的贡献值大小^[32]。本文基于随机森林算法，通过计算各特征基尼指数进行特征选择，并将阈值大于 0.005 的特征选择出来进行排序，最终共选取 23 个特征，其重要性排序如表 5.5 所示：

表 5.5 特征选择结果

特征	重要性排序
基础维度-总下单次数	1
比值特征 1-每个用户购买行为在五种行为加总次数占比	2
基础维度-五种行为加总次数	3
比值特征 3-用户商品购买次数与用户购买次数比值	4
用户商品维度-总购买次数	5
用户维度-总购买次数	6
比值特征 4-用户品类与用户比值	7
用户品类维度-总购买次数	8
用户商品维度-总下单次数	9
用户品类维度-总下单次数	10
用户维度-总下单次数	11
基础维度-总下单次数	12
商品维度转化率：加购-购买转化率	13
用户维度-总浏览次数	14
用户维度-五种行为加总次数	15
用户维度-总加购次数	16
用户品类组合维度-五种行为加总次数	17
商品维度转化率：浏览-购买转化率	18
商品维度转化率：下单-购买转化率	19
用户商品维度-五种行为加总次数	20
品类维度转化率：下单-购买转化率	21
用户品类维度-总浏览次数	22
基础维度-总加购次数	23

5.3.4 预测模型构建

1、预测模型训练

经过数据预处理及特征选择后，共选取有效样本数 123539 个，其中，训练集和测试集数据按照 8:2 的比例进行划分，正负样本比例为 1:2，特征变量共选出 23 个。由上文提到的 GBDT 算法参数设置可知，要提高模型预测准确率还需对 `n_estimators`、`learning_rate`、`max_depth`、`loss` 这四个参数进行设置。本文利用通过遍历所有给定参数

组合来优化模型表现的网格搜索法进行调参，具体过程如下：一般来说，模型调参首先从 $n_estimators$ 和 $learning_rate$ 入手，先选定一个较小步长，通过网格搜索得到最优迭代次数。本文将步长初始值设置为 0.1，对迭代次数进行网格搜索得到 $n_estimators=200$ ，接着在确定迭代次数后，又进一步对决策树最大深度进行网格搜索得到 $max_depth=6$ 。另外，由于用户购买行为预测属于 GBDT 算法的二分类问题，故选择对数似然(deviance)作为损失函数^[33]。综上，本文选定 $n_estimators=200$ 、 $learning_rate=0.1$ 、 $max_depth=6$ 、 $loss=deviance$ 这组参数用于模型训练。

2、模型效果评价

(1) 混淆矩阵

混淆矩阵输出结果如表 5.6 所示，根据混淆矩阵可进一步计算准确率、召回率、F1 值，计算结果如下：

表 5.6 混淆矩阵输出结果

真实 \ 预测	购买	未购买	总计
购买	15257	1244	16501
未购买	1360	6847	8207
总计	16617	8091	24708

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} = 89.5\%$$

$$P = \frac{TP}{TP + FP} = 91.8\%$$

$$R = \frac{TP}{TP + FN} = 92.5\%$$

$$F1 = \frac{2 * P * R}{P + R} = 92.1\%$$

(2) ROC 曲线

经过 GBDT 模型训练，基于上文混淆矩阵结果可进一步输出如下图所示的 ROC 曲线，并且该曲线围成的面积，即 AUC 值为 0.965，结合上文由混淆矩阵计算得出 F1 值为 92.1%可知该模型预测精度较高，拟合效果较好。

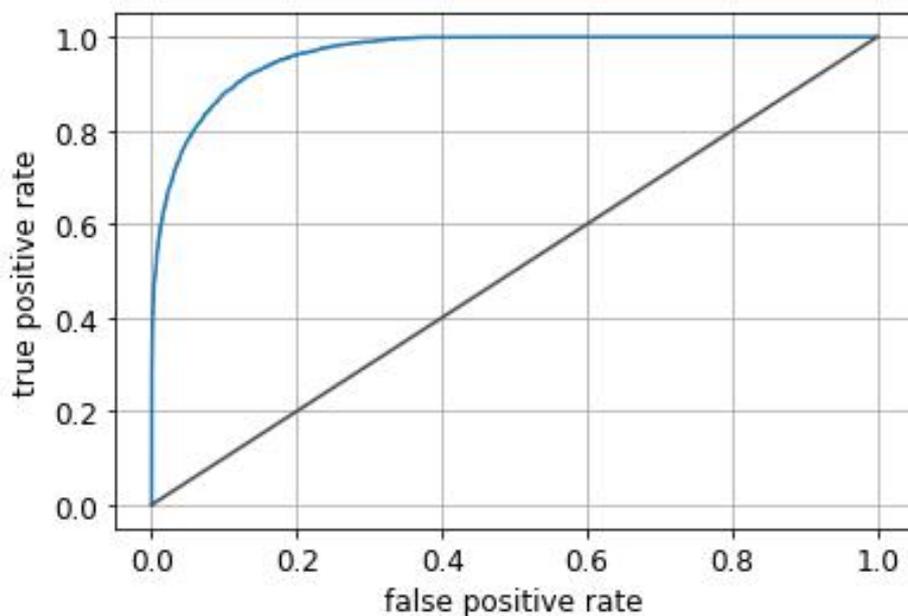


图 5.1 ROC 曲线

5.4 本章小结

本章基于决策树与 GBDT 算法概述以及模型效果分析的相关指标说明,阐述如何利用机器学习算法构建用户购买行为的预测模型,其中共包括数据预处理、样本选取、特征变量的构造与选择、模型训练在内的四个环节。具体构建过程可分为以下三步:首先,在确定数据来源后,对选取的平台交易数据进行缺失值、重复值、异常值的数据预处理,并设定、划分正负样本数据集;其次,从基础特征、组合特征、衍生特征三方面构造共计 54 个特征变量,并利用随机森林算法进行特征选择,将阈值大于 0.005 的特征选择出来进行重要性排序,最终共选择出 23 个特征变量;最后,基于 GBDT 机器学习算法构建预测模型,并输出混淆矩阵、ROC 曲线等模型效果评价指标。其中, F1 值为 92.1%、AUC 的值为 0.965,该数据表明模型预测效果较好。本文构建的预测模型主要通过用户的历史消费行为数据来量化并推测用户在短期内的购买意愿,从而为平台后续的个性化推荐奠定基础,同时,根据构建的预测模型也能倒推出对用户购买影响最大的特征变量,进而助力平台开展更有针对性的营销工作。

6 JDC 集团跨境电子商务营销策略提出

在前文分析中，第四章主要对 JDC 集团跨境电子商务平台目标用户进行精准化定位，第五章又基于机器学习算法，对用户在线购买行为展开预测，结合这两章内容能够使平台在充分利用交易数据的前提下，全面把握平台已有用户的消费行为。然而，平台日常运营除了要巩固老用户以避免用户流失之外，全方位扩大新用户规模也极为重要。因此，本章将从巩固平台已有用户及拓展新用户两视角为 JDC 集团跨境电子商务平台提出相应营销建议。

6.1 巩固平台已有用户视角

1、品类的产品构成优化、品牌推广层面

(1) 品类的产品构成优化

目前，JDC 集团跨境电子商务平台品类体系发展虽较为全面，但各品类对应的商品数量存在较大差异，品类的产品构成还有待优化。据前文分析可知，图书品类对应商品数量远超于其他品类，该现状也是平台上架商品分布严重失衡的直接原因。然而，随着 JDC 集团海外市场不断拓展以及跨境电子商务行业市场竞争日益加剧，为有着不同消费需求的用户提供多样化商品服务成为提升平台核心竞争力的关键举措。因此，本节从以下三方面提出推动产品构成优化的对策建议：一，充分利用平台交易、搜索、流量数据，基于各类型用户的商品需求分析，构建预测模型进一步把握其购买意愿程度，从而及时更新、补充和上架用户购买意愿较强的商品并针对不同用户分别进行精准的个性化产品推送；二，通过深度剖析海外市场各区域独立站点交易数据并针对各站点定期开展用户调研，总结用户购买需求，确定各站点平台发展定位，进而根据各区域用户需求变化以调整对应站点的上架产品；三，时刻关注对外贸易发展动态，探查、紧跟跨境电子商务行业发展新趋势，并根据行业变动及时调整产品策略，抢占市场先机。

(2) 品牌推广

“中国制造”在多年的成功探索及推广下已走上国际化道路，在全球市场上享有一定知名度，但由于用户需求的多变性，产品的低价及优质特点已不再成为吸引海外用户产生购买行为的最主要原因，而我国各跨境电子商务企业之间的竞争也必将从产品的“价格、质量战”转向“品牌战”。因此，JDC 集团要想具备较强的海外市场竞争力，就要紧跟国际市场发展趋势，打破产品同质化限制，突出集团发展特色，提升自营品牌

及平台入驻品牌的国际化地位，将平台各品牌逐步扎根于海外市场并使其成为集团在面向海外客群营销时的竞争优势，推动“中国制造”向“中国创造”升级转型^[41]。

2、用户标签系统构建层面

目前，JDC 集团跨境电子商务平台的用户分析一般从用户浏览记录、用户特征、用户价值以及用户偏好等方面入手，并没有具体划分用户类型分别展开研究，如此一来，尽管通过用户分析能够使平台对用户需求和偏好有大致把握，但对用户进行产品和页面的逐一推送会造成平台信息推送效率低、平台资源浪费、营销成本增加等后果。因此，平台应逐步建立起完备的用户标签系统，依据标签特点将不同用户归纳分类，并针对同一标签用户进行集中的信息推送。用户标签系统是指基于不同用户基本特征、消费偏好、购买习惯等画像数据通过重复迭代及聚类分析绘制出不同种类的标签，其优势在于标签系统能够使平台根据用户标签信息定位不同类别用户的购买需求，从而实现精准且高效的信息推送。用户标签系统的构建思路为：用户预分类——关键词提取——各类型用户标签设定——针对不同用户标签分别展开剖析，首先，对收集到的平台用户交易数据进行聚类分析，深度提炼各类用户关键词，同时基于关键词信息进一步生成不同用户标签，如“都市白领”、“新晋奶爸”、“小镇青年”、“资深中产”等，最后根据各标签人群分别从用户基本特征、消费水平、品牌品类偏好、产品关注点、在线评论等方面深入挖掘用户数据的各层面信息。

另外，在构建标签系统的基础上，JDC 集团还需结合大数据网络技术，不断创新营销模式，根据跨境电子商务的各业务线分别建立数据库信息系统，以便收集、整合平台用户数据及不同渠道流量数据从而深度探索用户信息。同时，还应充分利用平台交易数据，持续追踪用户行为，并“投其所好”的为其定期推送特定产品信息及活动页面，做到用数据指导营销工作开展。

3、跨境物流运输层面

跨境物流运输由于受到时间、距离、通关等诸多限制，存在运输效率低、客诉率较高、售后服务不完善等问题，而该类问题将会严重影响和制约 JDC 集团跨境电子商务平台的经营发展。为此，本节将从海外仓建设、整合现有物流模式并实现模式创新两方面提出改善跨境物流运输的相关建议。

（1）海外仓建设

海外仓是指海外仓库，即在海外用户分布最多的主要目标地区预先建立或租赁海外仓库，并根据平台业务量，提前将货物批量运输到海外仓中，从而能够通过海外仓实现

仓储、分拣、包装、配送等一站式服务^[36]。由于 JDC 集团的跨境物流运输具有批量小、批次多等特点，因而进行大批量、规模化的货物运输不仅能够有效降低物流运输、清关、商检等环节的频次和成本以及在运输各环节中可能存在的风险，而且还能在一定程度上缓解出口通关难的现状。另外，在海外建立仓库还能实现在买家所在地进行本土发货，缩短订单反应周期及物流配送周期，在达到提高物流配送效率目的的同时，还能使买家正常享受退换货等售后服务，提升其购物体验，进而降低用户客诉率。目前，JDC 集团海外仓建设还处于起步阶段，因而在集团后续发展中应持续推动各个目标国家的海外仓建设以不断开拓海外市场。

（2）整合现有物流模式并实现模式创新

从物流服务角度出发，可以将邮政小包、国际快运、海外仓、跨境专线等现有的物流配送模式进行有效整合，并将不同物流模式与有着不同消费特点的用户相结合，形成一套系统、完善的跨境物流运输体系。比如，对于收货时间没有要求且偏好低价的用户，可以采用国际快运与海外仓两者结合的物流运输模式，既能通过国际快运的物流模式满足其低价偏好，又能通过海外仓对其售后服务进行保障。而对于不愿将过多时间成本花费在跨境运输上且对低价并无明显偏好的用户，则可以采用海外仓物流配送与跨境物流专线服务相结合的模式，一方面，海外仓的预先备货机制能够满足用户对货品的时效性需求，另一方面，跨境物流专线服务还能为客户定制个性化专业服务^[21]。

6.2 拓展新用户视角

由 JDC 集团跨境电子商务平台的新老用户占比情况可知平台老用户总量高于新用户，故平台用户运营工作通常围绕老用户以唤醒沉默用户，避免用户流失等方式展开。然而，若要使企业以良好态势长久发展，除了维护平台老用户之外，发展新用户、拓展用户规模也是关键一步。新用户拓展实际上是引导新用户在下载注册平台软件的基础上，进一步产生下单购买行为，故本节将从产品、市场、运营三方面重点研究用户拉新层面的营销策略。

1、产品策略

在产品策略方面，首先要从用户底层需求出发，通过观察或调研目标用户群体行为从而总结用户特征及其相关的用户场景。通常，目标用户特征和用户场景背后产生的底层原因即为用户的真正需求，而挖掘出用户需求就能为用户拉新提供持续不断的原动

力。

（1）平台新人福利

平台新人福利是指针对首次注册下单的新用户发放特定福利，根据用户偏好不同设置“买就返”现金红包、大额优惠抵扣券、首单“0元购”等福利供其选择。但通过发放平台新人福利吸引新用户注册下单的拉新方式容易使用户产生“薅羊毛”行为，因而在用户进行首单消费后，平台需按照巩固老用户的相关营销策略进一步维护该用户后续的平台交互行为以避免用户流失。

（2）激励性拉新

激励性拉新是指给予平台已有老用户相应回报方式，刺激该类群体通过分享指定页面带动新用户的注册转化。该拉新方式的基本思路是将企业原本计划用于广告投放的推广成本直接发放给老用户，从而在运营成本相同的情况下，通过吸引老用户获取奖励来实现新用户的有效增长，如邀请好友赢红包、赚赏金等形式。另外，平台在开展激励性拉新工作时，还需把握老用户消费特征及偏好以精准定位活动受众群体，提高拉新效率，并进一步引导新用户注册成功后完成首单消费，提升其购买转化。

2、市场策略

（1）广告投放

广告投放是企业经营的重要一环也是平台用户拉新的常用手段，一般的广告投放形式有搜索引擎营销、粉丝通、电视广告、电梯广告、贴片广告等。在进行广告投放时，要根据不同投放渠道的客群特点分别进行广告宣传设计，提高营销推广精度。但随着近年来广告管理成本及时间成本的不断增加，加之广告对用户吸引力的持续下降，通过广告投放获取新用户的难度也日益增加。因此，为给平台带来推广效率和效果的双提升，程序化广告应被逐渐采用。程序化广告是一种以人工智能为基础，根据定位算法购买、投放广告的过程，该类型与传统广告之间的根本区别在于程序化广告是从受众群体角度出发，将广告内容精准曝光给相应目标用户，从而减少投放成本，提升投放效率。

（2）线下活动推广

线下活动推广常与拉新工作相结合，一般来说，企业在进行线下地推活动时会同步开展拉新工作，并通常以礼品、促销活动等方式吸引现场参与者下载平台软件并完成注册。通过线下活动推广虽然能使平台用户量在短期内快速增加，但该方法获取到的新用户存在较高流失风险，故平台应将线下活动推广的拉新工作与提升用户购买转化结合考虑并同步推进。

3、运营策略

(1) 应用商店推广

应用商店是指用户下载或购买所需 APP 的线上商店，是用户获取各大平台应用软件的主要渠道。因此，各大企业平台都会将 APP 在应用商店上架供用户下载。但由于各系统应用商店的推广位有限，致使无法满足所有企业的软件推广需求，因而应用商店也推出相应付费推广及应用商店优化服务，即提高软件推广精度或提升某平台 APP 在应用商店排行榜和搜索结果排名。因此，平台一方面可以使用应用商店付费推广服务实现圈定平台目标用户进行广告的精准投放，提高拉新引流效率，另一方面还能开通应用商店优化服务以提高平台软件的搜索排名，增加软件曝光度和用户下载量。

(2) 裂变活动推广

裂变活动是拉新营销的一种方式，与上述需求方主动接触用户等一系列推广方式不同的是，裂变属于被动拉新，目标是通过创意活动策划引发平台已有用户自主传播、分享给新用户，最终达成活动拉新指标。裂变活动推广的最大优势为平台在经过获取用户、激发活跃、提高留存、增加收入、传播推荐五个环节后，随着用户规模不断扩大，单个用户获取成本将不断降低^[43]。常见的裂变活动有新老用户拼团购、新用户砍价助力老用户、帮拆红包、“集字领卡”等。

7 总结与展望

7.1 总结

随着区域化经济合作持续深入以及大数据技术愈益成熟，跨境电子商务作为对外贸易新趋势及经济增长新的驱动力，在以迅猛的势头持续拉动我国对外贸易增长的同时，还逐渐发展成为推动对外贸易升级转型的重要抓手和进出口贸易的重要渠道。而立应用充实完善跨境电子商务理论研究，对我国跨境电子商务健康、持续发展能起到积极的助推作用。因此，本文以 JDC 集团跨境电子商务平台为例，从用户出发，探究在跨境电子商务行业竞争日益激烈的市场背景下，企业如何基于交易数据精准定位平台目标用户并进一步把握其购买意愿以开展精准化营销的发展途径。总体来说，本文主要完成了以下几项工作：

(1) 对市场营销和数据库营销以及跨境电子商务相关理论进行梳理。其中，市场营销的理论概述主要从市场营销的概念、演变过程以及发展新趋势三方面展开，数据库营销理论则主要介绍数据库营销的概念、优点及运作程序，而我国跨境电子商务发展概况的研究主要从跨境电子商务概念、我国跨境电子商务发展现状以及跨境电子商务对我国对外贸易的影响三方面入手，为后文分析奠定理论基础。

(2) 平台经营现状分析。选取 2019.1—2020.6 期间平台基础指标数据对 JDC 集团跨境电子商务平台经营现状展开剖析，并根据现状分析结果分别从产品、渠道、客诉及商家管理、数据四个层面总结 JDC 集团跨境电子商务平台存在的主要问题。

(3) 目标用户精准化定位。从用户基本特征、用户价值、用户消费偏好三方面入手精准定位平台目标用户，并基于分析结果进一步归纳目标客群消费特征。

(4) 用户购买行为预测模型构建。基于 GBDT 机器学习算法，按照数据预处理—样本选取—特征变量构造及选择—模型训练的步骤对用户购买行为构建预测模型，进而根据前文分析结论从巩固平台已有用户及拓展新用户两个视角分别为 JDC 集团跨境电子商务平台提供营销策略建议。

7.2 展望

本文基于 JDC 集团跨境电子商务平台交易数据，利用机器学习算法对平台营销策略进行探究。由于该研究领域涉及多学科的理论交叉以及各类型研究数据，且受限于软件

技术、数据获取、论文篇幅等条件，文章还存在许多不足，需在后续实践中不断完善，因而未来可以结合 JDC 集团跨境电子商务平台实际运营情况，从以下几方面进一步展开研究：

1、用户精准化定位

用户精准化定位是企业开展精准营销与实现个性化推送的必要前提，由于受到无法获取全量数据的条件限制，本文在用户层面的分析仅从用户基本特征、用户价值、用户消费偏好三方面展开。因此，为使 JDC 集团能对其平台用户特征有更为确切的把握，在后续研究中，应从多角度出发，结合平台标签系统及数据库系统对平台用户定位进行更深层次的探究。

2、实时预测消费行为

目前，JDC 集团跨境电子商务平台已实现用户消费数据的实时监控，但伴随跨境电子商务行业竞争加剧、互联网广泛普及与大数据技术日趋成熟，消费情况的实时预测才是未来跨境电子商务研究领域的重要内容，而各跨境电子商务企业也必然要将用户消费的实时预测作为各平台未来开展运营的首选路径。因此，后续研究可基于本文研究结论进一步对消费者购买行为的实时预测进行探索。

3、海外仓建设

据前文分析可知，目前 JDC 集团跨境电子商务平台的发展短板之一是跨境物流运输问题，而物流层面的问题会进一步影响平台售后服务及用户体验。因此，从集团长远发展的角度来看，租赁或建设海外仓库是跨境物流运输的最佳模式。在第六章提出的 JDC 集团跨境电子商务平台营销策略建议中已经对海外仓建设提供了大致发展思路，并对其主要优势进行概述，但对如何根据平台情况开展海外仓选址、具体模式规划、海外物流合作等工作以及需要解决的建设成本、库存周转、配送售后等一系列问题并没有展开讨论^[42]，因而在 JDC 集团海外仓建设的未来研究中，可从海外仓选址、模式规划、物流合作及面临的主要问题等方面入手，对现有研究成果进行丰富拓展。

参考文献

- [1]A.R.Kemal. Electronic Commerce and International Trade of Pakistan [J]. The Pakistan Development Review, 1998, 37(4):849-859.
- [2]Ángel Valarezo, Teodosio Pérez-Amaral, Teresa Garín-Muñoz, et al. Drivers and barriers to cross-border e-commerce: Evidence from Spanish individual behavior[J].Telecommunications Policy, 2018, 42(6):464-473.
- [3]Bin Hu, Qianqian Luo. Cross-border E-commerce Mode Based on Internet+ [J]. IOP Conference Series: Materials Science and Engineering, 2018, 394(5)
- [4]Erdal Aydın, Burcu Kılınç Savrul. The Relationship between Globalization and E-commerce: Turkish Case[J].Procedia-Social and Behavioral Sciences,2014, 150:1267-1276.
- [5]Estrella Gomez-Herrera, Bertin Martens, Geomina Turlea. The drivers and impediments for cross-border e-commerce in the EU[J]. Information Economics and Policy, 2014, 28:83-96.
- [6]Guy H. Gessner, Coral R. Snodgrass. Designing e-commerce cross-border distribution networks for small and medium-size enterprises incorporating Canadian and U.S. trade incentive programs[J]. Research in Transportation Business & Management, 2015, 16:84-94.
- [7]Mesut Savrul, Ahmet Incekara, Sefer Sener. The Potential of E-commerce for SMEs in a Globalizing Business Environment[J]. Procedia - Social and Behavioral Sciences, 2014, 150:35-45.
- [8]Nuray Terzi. The impact of e-commerce on international trade and employment[J]. Procedia - Social and Behavioral Sciences, 2011, 24:745-753.
- [9]崔艳武,苏秦,李钊.基于电子商务环境的顾客消费偏好研究[J].软科学,2007(06):19-23.
- [10]鄂立彬,黄永稳.国际贸易新方式:跨境电子商务的最新研究[J].东北财经大学学报,2014(02):22-31.
- [11]方佳虹.跨境电子商务模式下出口定价新思路[J].现代商业,2016(04):49-50.
- [12]樊文静.跨境电子商务发展与我国对外贸易模式转型[J].对外经贸,2015(01):4-7.
- [13]郭国庆.市场营销学通论[M].中国人民大学出版社:2020(08):6-34.
- [14]关怀庆.我国跨境电子商务的发展现状趋势及对策研究[J].中国商论,2016(01):50-52.

- [15]郭四维,张明昂,王庆,朱贤强.新常态下的“外贸新引擎”:我国跨境电子商务发展与传统外贸转型升级[J].经济学家,2018(08):42-49.
- [16]郝胜宇,陈静仁.大数据时代用户画像助力企业实现精准化营销[J].中国集体经济,2016(04):61-62.
- [17]卡特勒,卡塔加雅,塞蒂亚万.营销革命 4.0: 从传统到数字[M].机械工业出版社:2018:35-43.
- [18]刘洪程.顾客忠诚度与客户关系管理的整合[J].当代经济,2004(08):26-27.
- [19]李晓龙,冯俊文.大数据环境下电商精准营销策略研究[J].价值工程,2016,35(03):31-33.
- [20]来有为,王开前.中国跨境电子商务发展形态、障碍性因素及其下一步[J].改革,2014(05):68-74.
- [21]罗扬.我国跨境电子商务企业出口营销策略研究[D].吉林大学,2017.
- [22]刘永焕.“一带一路”战略下跨境电子商务出口营销策略的优化[J].对外经贸实务,2019(02):36-39.
- [23]李柏杏,潘开灵.跨境电子商务对我国进出口贸易影响的实证分析[J].商业经济研究,2016(23):132-134.
- [24]廖蓁,王明宇.跨境电子商务现状分析及趋势探讨[J].电子商务,2014(02):9-10.
- [25]刘潇蔓.基于特征选择和模型融合的网络购买行为预测研究[D].北京交通大学,2017.
- [26]刘海,卢慧,阮金花,田丙强,胡守忠.基于“用户画像”挖掘的精准营销细分模型研究[J].丝绸,2015,52(12):37-42+47.
- [27]孙琪.我国跨境电子商务发展现状与前景分析[J].商业经济研究,2020(01):113-115.
- [28]宋宝香.数据库营销:大数据时代引发的企业市场营销变革[J].价值工程,2012,31(30):132-134.
- [29]檀辉霞.浅析市场营销观念的演变和发展新趋势[J].经济师,2011(04):272-273.
- [30]王立.基于机器学习的共享单车需求量预测模型研究[D].长安大学,2019.
- [31]吴娉娉.浅析跨境电子商务企业出口选品策略[J].时代金融,2016(30):235-236.
- [32]王天华.基于改进的 GBDT 算法的乘客出行预测研究[D].大连理工大学,2016.
- [33]温博文,董文瀚,解武杰,马骏.基于改进网格搜索算法的随机森林参数优化[J].计算机工程与应用,2018,54(10):154-157.
- [34]应小凡.电子商务对国际贸易的积极影响与存在问题分析[J].商业时代,2011(11):53-54.

- [35]姚登举,杨静,詹晓娟.基于随机森林的特征选择算法[J].吉林大学学报(工学版),2014,44(01):137-141.
- [36]赵旭明,杨晓涵.跨境电子商务发展对我国对外贸易模式转型影响分析[J].商业经济研究,2016(08):75-77.
- [37]庄贵军,周南,周连喜.国货意识、品牌特性与消费者本土品牌偏好——一个跨行业产品的实证检验[J].管理世界,2006(07):85-94+114+172.
- [38]赵云泽,滕沐颖,赵菡婷,孟蝶,刘璇.“桥梁人群”对中国品牌的跨文化传播的影响研究[J].国际新闻界,2015,37(10):65-78.
- [39]张夏恒,马天山.中国跨境电子商务物流困境及对策建议[J].当代经济管理,2015,37(05):51-54.
- [40]周记顺,郭昕宇.中国跨境电子商务发展面临的困境及对策分析[J].对外经贸实务,2019(07):25-28.
- [41]朱恺.我国跨境电商发展现状分析及相关建议[J].现代商业,2014(18):65-66.
- [42]张晓燕.我国跨境物流海外仓发展存在的问题及完善对策[J].对外经贸实务,2017(01):84-87.
- [43]周劲波,位何君.基于 AARRR 模型的用户增长策略研究——以拼多多为例[J].山西经济管理干部学院学报,2020,28(01):11-16.
- [44]张棫,曹健.面向大数据分析的决策树算法[J].计算机科学,2016,43(S1):374-379+383.

后记

弹指一瞬间三年的研究生生涯已经接近尾声，在这三年充实的学习生活中，我不仅掌握了统计学的相关专业知识，同时还参与了实习工作，理论与实践的结合一方面开阔了我的眼界，另一方面也使我对专业课知识有了更加深刻的理解。通过研究生三年的学习与钻研，我的能力和思想得到了飞跃提升，而这三年时光也成为我学习生涯中收获最多的三年。至此，在毕业论文即将完成之际，我特别向所有帮助、关心我的人表示最真诚的感谢。

首先，我要特别感谢我的硕士导师牛胜强老师。本篇论文是在导师的悉心指导下完成的，在我遇到困难、没有自信的时候，牛老师给予我鼓励与支持，并对我的毕业论文进行了十分细致的指导。牛老师丰富的理论知识，严谨的科研作风以及对学生认真负责的态度令我十分敬佩。研究生就读期间，牛老师不仅在科研与学习方面指导我，同时还鼓励我积极参与实践活动，使我在科研学习和生活上都收获很多。本篇论文从选题到最终定稿也花费了牛老师大量的精力和时间，在此期间老师为我提供的论文指导意见也使我能较为高效、方向明确的进行毕业论文撰写，在此，我要向牛老师表达最衷心的感谢！

另外，还要感谢研究生就读期间宿舍室友以及师门的同级伙伴、师兄师姐和师妹。感谢室友在研究生三年，特别是毕业论文撰写期间的陪伴与督促，每一个早起的清晨和每一个熬到的深夜都将使我铭记于心，不会忘记写论文期间那段简单、纯粹又充实的日子。感谢师门伙伴在科研项目研究中的相互探讨以及师兄师姐对我毕业论文的帮助与指导，正是有了和他们的交流才能使我在读研期间不断进步。

最后要感谢我的父母。感谢他们在任何时刻都是我最坚实的后盾，尊重我的每一个选择，陪伴我度过人生的每个阶段。读研的这三年也正是有了他们的鼓励和支持，才能使我毫无顾虑的勇往直前，实现自己定下的一个又一个目标。今后，我将再接再厉，用实际行动来回馈他们，并努力做一个对国家、社会有用的人，报效祖国。