

基于数据挖掘的净水器产品精准营销研究

赵磊

甘肃农业大学管理学院, 甘肃 兰州

收稿日期: 2025年10月13日; 录用日期: 2025年10月29日; 发布日期: 2025年11月28日

摘要

数据挖掘技术为净水器产品精准营销提供了有效的解决路径, 通过对消费者行为数据进行多维采集与预处理, 能够准确识别客户需求特征及行为模式。聚类分析实现了客户群体的科学细分, 预测模型提升了潜在客户的识别精度, 关联规则挖掘优化了产品组合推荐策略, 渠道匹配技术改善了营销触达效率, 个性化内容生成增强了客户响应率。实证研究表明, 基于数据挖掘的精准营销方法, 显著提高了净水器产品的市场转化率与投资回报率, 降低了营销成本, 为净水器企业实现数据驱动的营销决策提供了技术支撑与实践参考。

关键词

数据挖掘, 净水器产品, 精准营销, 客户识别

Research on Precision Marketing of Water Purifier Products Based on Data Mining

Lei Zhao

School of Management, Gansu Agricultural University, Lanzhou Gansu

Received: October 13, 2025; accepted: October 29, 2025; published: November 28, 2025

Abstract

Data mining technology provides an effective solution for precision marketing of water purifier products. Through multi-dimensional collection and preprocessing of consumer behavior data, it enables accurate identification of customer demand characteristics and behavioral patterns. Cluster analysis facilitates scientific segmentation of customer groups, prediction models enhance the accuracy of potential customer identification, association rule mining optimizes product portfolio recommendation strategies, channel matching technology improves marketing outreach efficiency, and personalized content generation increases customer response rates. Empirical research demonstrates that the precision marketing approach based on data mining significantly improves the market conversion rate

and return on investment for water purifier products, reduces marketing costs, and provides technical support and practical references for data-driven marketing decision-making in water purifier enterprises.

Keywords

Data Mining, Water Purifier Products, Precision Marketing, Customer Identification

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着健康饮水理念逐渐普及，净水器市场竞争变得异常激烈，传统营销模式难以满足消费者个性化需求，营销精准度迫切需要进行进一步改善，数据挖掘技术通过深度分析海量消费数据，能够精准识别客户群体、预测购买行为、优化营销策略，为企业提供数据驱动的营销决策支撑。然而净水器作为低频高价消费品，其消费数据呈现多源异构特征，客户行为复杂多变且决策周期较长，如何有效挖掘数据价值、实现精准营销成为关键问题。本文将数据挖掘技术系统性应用于净水器营销领域，构建从数据采集到效果评估的完整方法论，为企业降本增效提供实践指导。

2. 净水器消费数据挖掘与客户识别

2.1. 净水器消费者多维数据采集与预处理

净水器消费者数据有着多源异构的典型特点，数据采集覆盖线上电商平台交易记录、线下门店销售信息、社交媒体互动行为、搜索引擎查询日志以及客户服务系统反馈等多个渠道。电商平台数据包含产品浏览深度、加购行为、价格敏感度及评价内容等结构化信息，社交媒体数据涉及用户对水质安全话题关注度、品牌讨论热度及口碑传播路径等非结构化文本，数据预处理环节要针对不同来源数据的质量问题做系统化处理，缺失值通过均值填充或基于相似用户的协同填充方法来补全，异常值采用箱线图检测与孤立森林算法识别后剔除或修正，重复数据通过设定唯一标识符来进行去重处理[1]。针对非结构化文本数据，采用分词技术、情感分析模型提取消费者对净水器功能、价格、服务的态度倾向，并转化为可量化的特征变量。数据标准化处理消除不同维度数据的量纲差异，以确保各类数据在统一尺度下做后续分析与建模。

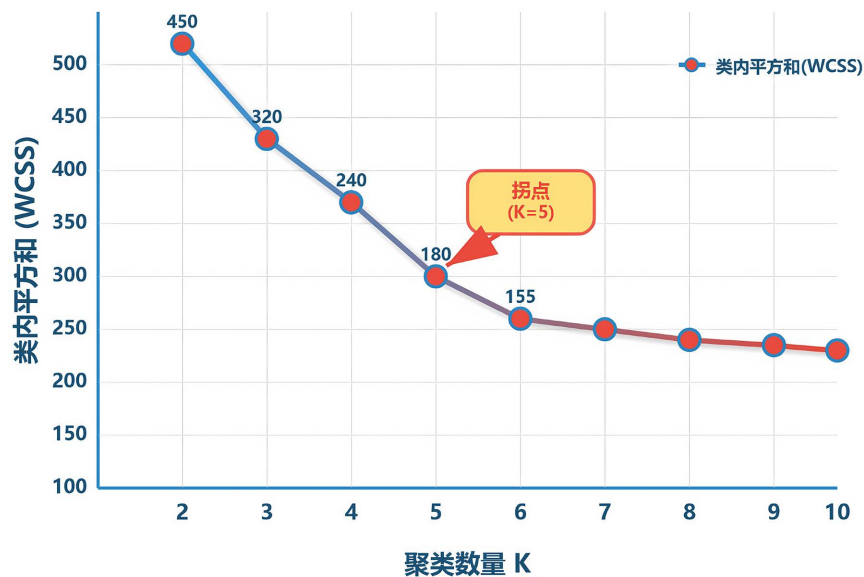
2.2. 客户行为分析与需求特征识别

客户行为分析聚焦于消费者在净水器购买决策过程中，展现出行为规律与需求偏好识别，通过构建用户行为序列模型，追踪消费者从信息搜集、产品比较、咨询互动到最终购买的完整路径，识别不同阶段的关键触点及影响因素。浏览时长与页面跳转路径以及咨询问题类型等行为特征，揭示了消费者对净水器过滤技术、滤芯更换成本、出水速度等功能属性的关注权重，需求特征识别从地域水质差异、家庭人口结构、消费能力水平、健康意识强度等维度构建特征变量体系，运用特征选择算法筛选出对购买决策影响显著的核心变量[2]。频繁购买净水器滤芯的用户群体显示出对产品维护便利性的高度敏感，而搜索“母婴用水”关键词的用户更注重水质安全等级，通过关联分析发现，咨询安装服务的用户转化率显著高于仅浏览产品参数的用户，这表明服务保障是影响购买决策的重要因素。这些行为模式与需求特征的量化表征，构成了客户群体划分与精准营销决策的核心依据。

3. 数据挖掘驱动的净水器精准营销实现

3.1. 客户群体聚类分析与市场细分

客户群体聚类分析借助 K-means 算法来构建多维度分群模型，最优聚类数量的确定遵循肘部法则与轮廓系数双重验证标准，肘部法则的实施流程是设定 K 值范围为 2~10，计算各个 K 值所对应的类内平方和，绘制 K 值与平方和变化率的曲线，识别下降幅度显著减缓的拐点位置以此确定最优 K 值(如图 1 所示)。轮廓系数用于验证聚类质量，通过计算样本到同类其他点与到最近异类点的平均距离比值，若系数值超过 0.5 的门槛则表明聚类效果达标，特征变量体系涵盖消费能力、购买频次、品牌偏好、功能关注度等多个维度，其权重分配依据主成分分析的维度贡献率结果[3]。净水器行业特有的地域水质差异作为关键辅助变量被纳入模型，从国家水质监测数据库提取 TDS 值、硬度等级、污染物浓度等指标，结合地理编码技术实现客户地址信息的精准匹配。算法运行参数配置为最大迭代次数设定为 500 轮，收敛阈值设定为 $1e-6$ 以确保算法充分收敛，初始中心点选择 K-means++ 策略来提升收敛稳定性。



注: 拐点位置(K=5)表示最优聚类数量, 此时类内平方和下降幅度显著减缓

Figure 1. Analysis diagram of K-means clustering elbow rule

图 1. K-means 聚类肘部法则分析图

3.2. 购买需求预测模型构建

购买需求预测模型构建遵循三阶段特征工程流程，确保入模变量的统计显著性与业务相关性，首阶段运用卡方检验筛选与目标变量显著相关的特征，显著性水平阈值 0.01，次阶段利用互信息评分度量特征与目标变量的非线性关联强度，保留评分排序前 80% 的候选变量，末阶段执行递归特征消除算法，结合交叉验证评估确定最优特征子集。经三轮筛选后最终入模的 18 个核心变量为：浏览深度、咨询频次、价格敏感度、决策周期天数、同类产品对比次数、家庭人口数、所在地区 TDS 值、历史家电购买频次、品牌搜索次数、详情页停留时长、客服咨询轮次、产品收藏行为、评价查看深度、促销活动参与度、安装服务咨询、滤芯更换咨询、智能功能关注度、售后服务评价。XGBoost 算法优化过程采用 5 折交叉验证网格搜索机制，参数候选区间包括树数量 100~500、最大深度 3~7、学习率 0.01~0.2，验证集 F_1 值最大化作为参数选择准则，其计算公式为：

$$F_1 = \frac{2 \cdot (\text{Precision} \cdot \text{Recall})}{(\text{Precision} + \text{Recall})}$$

(1)

式中，Precision 为精确率，表示预测为正例的样本中实际为正例的比例；Recall 为召回率，表示实际正例中被正确预测的比例， F_1 值作为精确率与召回率的调和平均数，能够综合评估模型在净水器购买需求预测任务中的性能表现，有效平衡了误判率与漏判率的影响[4]。新用户冷启动难题通过迁移学习策略解决，基于人口统计学属性构建初始预测能力。

3.3. 产品关联规则挖掘与组合推荐

产品关联规则挖掘整合改进 Apriori 算法与时间衰减机制，深入挖掘净水器产品的购买关联模式(如图 2 所示)，阈值参数设定充分考虑净水器市场特征，主流产品最小支持度设为 3%以兼顾关联发现与计算效率，长尾产品支持度降至 1%避免遗漏重要模式，最小置信度设为 40%确保关联规则的预测有效性，提升度门槛设为 1.2 筛选超越随机关联的强规则[5]。时间衰减模型专门针对净水器耗材更换周期进行设计，采用指数衰减函数并设定权重，衰减系数在 0.1~0.2 区间基于历史滤芯购买间隔分析来确定，关联规则质量评估引入兴趣度指标与卡方统计检验双重验证机制。算法执行包含四个核心环节，数据预处理阶段构建商品 - 交易二元矩阵，频繁项集挖掘阶段实施自底向上逐层扩展策略，关联规则生成阶段计算支持度与置信度指标，规则过滤阶段应用预设阈值剔除低质量规则。推荐引擎集成关联规则输出结果，构建商品关联知识图谱，客户产品浏览行为触发智能推荐算法，商品关联强度时间衰减机制保障推荐内容的时效性与精准度。

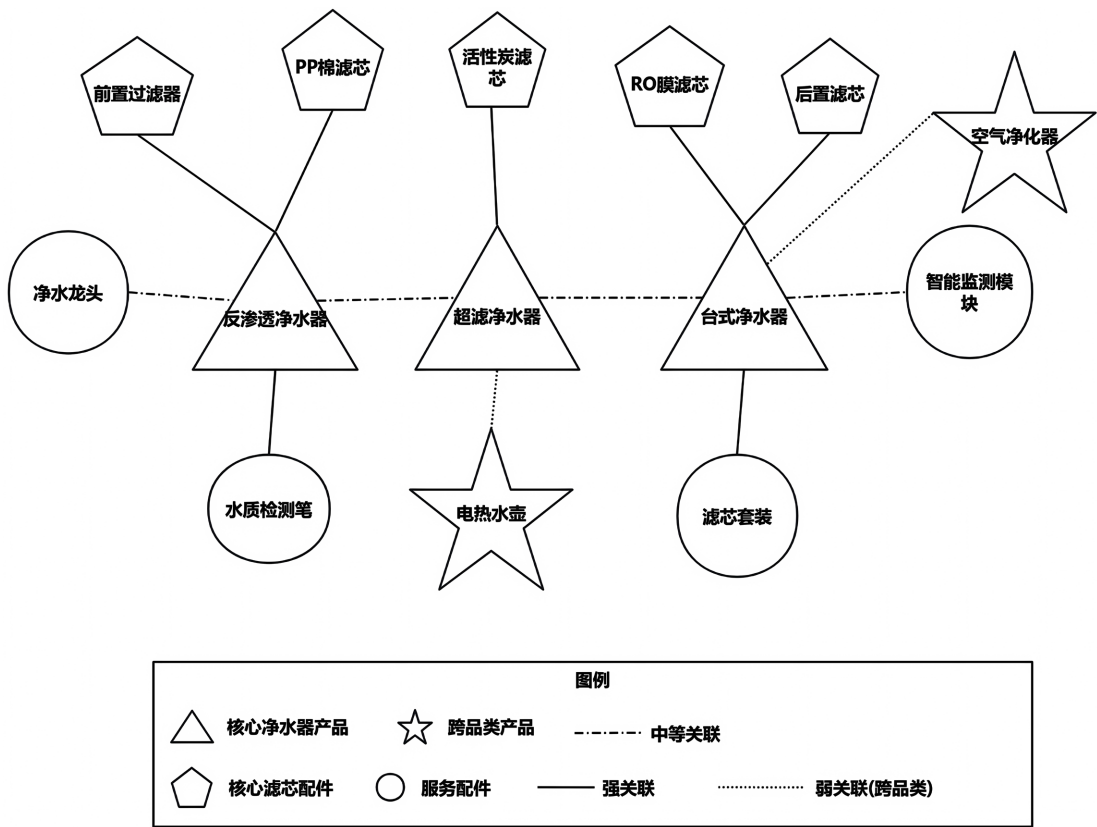


Figure 2. Network diagram of product association rule
图 2. 产品关联规则网络图

3.4. 营销渠道匹配与触达时机优化

营销渠道智能匹配机制依靠基于客户画像和渠道属性的多维相似性评估算法，渠道属性特征建模包含用户年龄结构、活跃时段分布、内容消费偏好、交互行为模式等关键维度，客户-渠道相似度计算采用引入的余弦相似度算法，客户特征向量整合人口统计学、行为偏好、消费水平三大子维度信息。渠道匹配评分权重体系由专家评估确定，其中年龄契合度占比 30%、兴趣偏好匹配度 25%、活跃时段重叠度 20%、消费层级匹配度 15%、内容形式偏好 10%，触达时机优化运用时间序列挖掘技术，设定 1 小时粒度时间窗口，通过连续 4 周行为数据观察识别客户活跃规律。客户生命周期阶段划分基于 RFMBI 扩展模型，整合近期购买时间 R、购买频次 F、购买金额 M、浏览深度 B、互动频次 I 多维指标，经聚类分析输出潜在期、兴趣期、考虑期、决策期、复购期五阶段分类。触达频次智能控制基于疲劳度函数设计，当信息接收次数超限时系统自动暂停推送避免损害用户体验，多渠道协同调度策略确保营销信息精准投放，以实现最优时机最适渠道的高效触达。

3.5. 营销内容个性化生成与精准投放

个性化内容生成平台将模板化生产和智能组合技术整合起来，构建出灵活且高效的内容创作架构，内容模板知识库设计成三层架构形式，主题分类层涵盖产品介绍、技术解析、使用指导、促销活动等内容类型，素材元素层包含文案模块、视觉素材、视频片段、交互组件等，组合规则层定义不同元素的搭配逻辑与展示策略。客户画像特征提取融合地域水质环境数据、家庭人口结构特征、历史消费偏好模式、购买决策阶段状态这四维信息，内容智能生成采用规则引擎与机器学习协同架构，规则引擎负责确定性业务逻辑处理，机器学习模型处理复杂特征交互与个性化推理。A/B 测试实验框架集成流量分配、效果监控、统计检验三大功能模块，多臂老虎机算法实现流量比例的动态优化调整，内容效果评估指标体系横跨曝光维度、互动维度、转化维度三个层面，加权评分机制综合衡量内容营销质量。个性化推荐算法融合协同过滤与内容过滤双重技术路径，加权融合机制输出最终推荐方案，内容优化迭代机制根据用户行为反馈数据持续改进模板库与生成策略，形成闭环优化体系。

4. 净水器精准营销效果评估与优化

4.1. 营销转化率与投资回报率测算

营销转化率和投资回报率的测算打造出精准营销效果量化评估体系，转化率指标会从潜在客户识别、营销触达、咨询互动、购买决策等多个环节来分层测算，各环节转化率连乘结果能反映营销漏斗整体效率，单一环节转化率异常波动可精准定位营销流程薄弱环节。投资回报率测算会综合考量营销活动直接成本与间接成本，直接成本包含广告投放费用、促销折扣损失、渠道佣金支出，间接成本涵盖数据处理、模型维护、人员培训等方面投入，客户生命周期价值引入让投资回报率测算从单次交易延伸至长期价值评估，复购率、客单价提升幅度、推荐转化贡献等指标被纳入综合收益计算框架。差异化营销策略在不同客户群体中投资回报表现有显著差异，高价值客户群体获客成本高但长期收益更可观，短周期评估侧重于即时转化效果，长周期评估关注客户留存与品牌价值提升，两种评估周期结合确保营销策略动态平衡。

4.2. 模型反馈优化与营销策略迭代

模型反馈优化和营销策略迭代构建起数据驱动的持续改进机制，营销活动产生的实际转化数据持续回流到预测模型之中，新增样本不断丰富训练数据集并修正模型相关参数，模型预测准确率伴随数据积累呈现动态提升的趋势。客户行为模式的演变要求模型具备自适应的能力，购买决策周期缩短、新兴渠

道崛起、产品功能偏好迁移等市场变化通过特征权重调整与变量更新机制被及时捕捉，A/B 测试结果为营销策略优化提供实验数据方面的支撑，不同营销内容版本、推送时机选择、渠道组合方案的效果差异通过对照实验被量化评估。模型的定期重训练用以应对市场环境 with 用户结构的动态变化，新兴用户特征或原有行为模式的漂移通过模型输出结果的对比分析被识别出来，关联规则的时效性监测，确保产品推荐策略的持续有效，关联强度的周期性评估与规则库的动态更新，维持推荐系统的准确性，数据驱动文化与跨部门协作机制支撑模型优化的持续开展。

5. 结语

数据挖掘技术将净水器营销从经验驱动转变为数据驱动，显著提升了客户识别精度与营销资源配置效率，通过构建完整的数据采集、模型构建、效果评估闭环体系，净水器企业能够实现客户需求的精准洞察与营销策略的动态优化。实践表明，数据挖掘方法在降低获客成本、提高转化效率、增强客户满意度方面展现出明显优势，然而，数据质量标准尚未统一、模型适用性存在局限、隐私保护压力持续增大等问题制约了技术深度应用。未来需要加强多源数据融合技术研究，探索算法模型的自适应优化机制，建立规范的数据安全与伦理体系，推动数据挖掘在净水器精准营销领域的纵深发展。

参考文献

- [1] 张宇, 陆文成, 赵锡成, 等. 基于深度学习的用户回访行为预测算法改进[J]. 计算机仿真, 2025, 42(9): 472-476.
- [2] 禚圆华. 人工智能驱动的种业电子商务平台精准营销策略分析[J]. 分子植物育种, 2025, 23(9): 3157-3162.
- [3] 朱逸, 朱瑞庭, 王鹏. 人工智能营销的机器学习算法谱系及实践域[J]. 管理现代化, 2022, 42(5): 152-160.
- [4] 危婷, 张宏海, 蔺小丽, 等. 云服务网站用户回访行为预测模型研究[J]. 数据与计算发展前沿, 2022, 4(3): 124-130.
- [5] 张方风, 倪然. 基于 SOMTE 平衡的在线用户消费预测模型研究[J]. 数学的实践与认识, 2020, 50(15): 49-59.