

多行为序列建模下的电商用户购买行为与关键影响因素分析

吴浩然

上海理工大学管理学院, 上海

收稿日期: 2025年11月22日; 录用日期: 2025年11月25日; 发布日期: 2025年12月24日

摘要

在电商领域持续扩张的背景下, 用户在平台上的多样化互动为构建推荐模型和开展购买意愿预测提供了充足的数据基础。本研究利用阿里天池发布的电商行为数据作为分析对象, 构建融合浏览、关注等多类行为特征的序列化模型, 并采用XGBoost算法对用户的实际购买动作进行预测。研究中进一步引入SHAP方法, 以增强模型的解释框架, 从量化视角刻画不同特征的重要程度及其潜在影响路径。与Logistic Regression等传统对照模型相比, XGBoost在Accuracy、AUC与F1等核心指标上均表现更为突出, 体现出较强的预测能力。通过SHAP的解释结果可以观察到, 总浏览量等长期且累积性的行为变量对购买决策具有显著影响; 而短期行为以及收藏类动作, 对用户最终是否下单的贡献度相对较弱。

关键词

电商用户行为, XGBoost, SHAP, 多行为序列, 购买预测

Analysis of E-Commerce User Purchase Behavior and Key Influencing Factors under Multi-Behavior Sequence Modeling

Haoran Wu

Business School, University of Shanghai for Science and Technology, Shanghai

Received: November 22, 2025; accepted: November 25, 2025; published: December 24, 2025

Abstract

In the context of the rapid development of the e-commerce industry, the diverse and abundant user

文章引用: 吴浩然. 多行为序列建模下的电商用户购买行为与关键影响因素分析[J]. 电子商务评论, 2025, 14(12): 4220-4228. DOI: 10.12677/eci.2025.14124362

behaviors have laid a solid data foundation for building recommendation systems and predicting purchasing behavior. This study utilizes e-commerce user behavior data provided by Alibaba Tianchi as the research material and constructs a sequential model that integrates multiple behavioral features such as browsing and collecting. The XGBoost algorithm is employed to predict purchasing behavior, while the SHAP method is introduced to enhance model interpretability and quantitatively analyze the importance and impact of each feature. The results show that compared with baseline algorithms such as Logistic Regression, XGBoost demonstrates superior performance across multiple evaluation metrics—including Accuracy, AUC, and F1-score—indicating strong predictive capability. According to the SHAP analysis, factors such as the total number of browsing actions are key determinants of purchasing behavior, whereas short-term behaviors and collection actions contribute relatively less to explaining purchasing decisions.

Keywords

E-Commerce User Behavior, XGBoost, SHAP, Multi-Behavior Sequence, Purchase Prediction

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

当前，电子商务持续高速扩张，用户在平台上的互动方式也愈发多元，包括浏览商品、收藏心仪物品、添加购物车以及完成下单等常见行为[1]。这些行为不仅反映了用户的兴趣偏好与潜在消费意愿，也为预测购买决策提供了关键的数据基础。然而，既往有关用户购买预测的研究，多聚焦于单一行为特征，或主要依赖人口统计属性进行分析，往往忽视了多种行为按时间顺序发生时的动态关联及其复杂的交互结构。这种局限使得模型预测精度难以进一步提升，从而限制了其在个性化推荐和精细化营销中的应用价值[2]。近年来，机器学习技术在用户行为研究领域得到广泛应用。其中，XGBoost作为一种高效的梯度提升树模型，因其在处理非线性特征关系、捕捉关键变量的重要性以及应对复杂数据模式方面具有显著优势，已成为用户购买可能性预测中被频繁采用的算法之一[3]。与此同时，SHAP方法为模型解释提供了统一且严谨的理论框架，可用于量化每个特征对预测结果的边际贡献，有助于进一步提升模型的透明度与可解释性，从而加深对用户行为机制的理解[4]。从多行为序列的视角出发，若将浏览、收藏、加购和购买等行为按照时间顺序构建特征结构，能够更系统地刻画用户的行为路径，并揭示兴趣随时间演化的特征模式。借助XGBoost对这些顺序化特征进行学习，并结合SHAP分析其解释效应，可以识别对购买决策影响最显著的行为模式，并进一步量化各关键特征的重要程度。此类分析不仅有助于电商平台优化营销策略，也为实现更精准的个性化推荐提供了数据驱动的依据[5]。本研究所得结果不仅在理论层面深化了对用户购买行为的理解，也在实践层面提出了具有可操作性的策略建议，可为智能推荐系统的构建和平台转化率的提升提供可靠的数据支撑。

2. 文献综述

用户在电商平台中何时产生购买行为，一直是学术界与产业界关注的核心议题。早期研究主要依赖传统统计方法，通过分析用户的基本属性，如年龄、性别等人口统计特征，以及其以往的浏览与购买记录，对未来的购买可能性进行推断[4]。虽然此类方法在数据规模较小的情境下具备一定解释力，但在面对高维度、非线性结构的数据，以及用户行为之间具有明确时间依赖的情形时，其预测性能往往显得不

足。随着数据量的急剧增长以及计算能力的提升,机器学习逐渐成为用户购买行为分析的主要技术路径。决策树、随机森林及支持向量机等方法被广泛应用于复杂特征关系的建模,在面对多维输入与非线性模式时具有良好表现[5]。然而,这些模型通常未能充分考虑用户行为在时间序列层面的先后关系,亦缺乏对不同行为组合可能产生联动效应的系统刻画。为应对上述问题,多行为序列建模方法逐渐受到关注。该方法将用户在平台上的浏览、收藏、加购及购买等行为按照时间顺序组织为序列特征,使得用户兴趣的动态变化及其决策路径得以更完整地呈现[3]。研究表明,将行为顺序特征纳入模型能够显著提升购买预测的精度,同时为个性化推荐提供更具针对性的参考依据[6]。近年出现的深度学习模型,如 LSTM 与 Transformer,也被用于捕获长距离依赖与复杂交互关系[7],但其可解释性不足仍是其在实际应用中的主要限制。在提升模型透明度方面,SHAP 提供了一套统一的特征贡献度评估框架,可用于量化单个特征对预测结果的边际影响。当其与 XGBoost 等树模型结合使用时,能够更清晰地揭示各类行为特征在用户购买决策中的作用机制,为理解模型预测逻辑和解释用户行为提供了可靠依据[2]。在电商领域,识别关键影响因素、优化推荐策略以及制定精细化营销方案等任务,均可从该方法中获益[8]。综上,现有研究在用户购买预测方面取得了显著进展,但仍存在两方面不足:其一,多行为序列特征的潜在价值尚未被充分挖掘;其二,在模型预测能力与可解释性之间仍缺乏兼顾的有效方案。基于 XGBoost 的多行为序列建模并结合 SHAP 的解释框架,为同时提升预测精度与模型透明度提供了一条可行路径。

3. 理论方法

3.1. 方法构建

本研究旨在基于用户历史行为数据构建用户购买行为的预测模型。预测任务被形式化为一个二分类问题:对于每一名用户,根据其过去一段时间内的行为序列及行为统计特征,预测其是否发生过购买行为。标签变量定义如下:

正样本(1): 该用户在观察窗口内累计出现至少一次购买行为;

负样本(0): 该用户在观察窗口内未出现任何购买行为。

在样本构造过程中,以用户为单位进行特征聚合。所有用户均按照该规则进行标注,训练集与测试集按 7:3 的比例划分,并通过 stratify = y 的方式保证正负样本比例在训练集和测试集之间保持一致。本文所使用的数据来源于阿里巴巴天池大数据竞赛公开数据集

(<https://tianchi.aliyun.com/competition/entrance/1/informatio>),涵盖 20,000 名用户在某年 11 月 18 日至 12 月 18 日期间的完整行为记录。每条购物行为记录包含四个字段:用户 ID、品牌 ID、用户与商品的交互行为类型及行为时间。用户与商品的交互行为包括“浏览”“收藏”“加入购物车”和“购买”四类。

本研究所构建的特征主要分为两类:一类为统计特征,其详细描述见表 1;另一类为时间特征,其详细说明见表 2。

Table 1. Description of statistical features
表 1. 统计特征描述

编号	特征名称	特征描述
1	总浏览次数	用户在整个观察周期内对商品页面的浏览总次数
2	总收藏次数	用户累计将商品加入收藏夹的次数
3	总加购物车次数	用户累计将商品加入购物车的次数
4	访问商品数	用户在观测期内访问过的不同商品数量
5	活跃天数	用户在观测期内有行为(浏览、收藏、加购、购买)的天数

Table 2. Description of time-related features**表 2.** 时间特征描述

编号	特征名称	特征描述
1	7 天浏览次数	7 天内用户的商品浏览总次数
2	7 天收藏次数	7 天内用户收藏商品的次数
3	7 天加购物车次数	7 天内用户将商品加入购物车的次数

3.2. XGBoost 模型构建

本研究所采用的极端梯度提升树模型(XGBoost)属于典型的基于梯度提升框架的集成学习算法。该模型具有计算效率高、稳定性强以及相对易于解释等优势。其基本思想是通过迭代方式构建多个弱学习器，并依据预设权重将其集成，从而形成一个性能更优的强预测模型。整个训练过程以最小化目标函数为核心，而目标函数由损失项与正则化项共同构成：前者确保模型能够充分反映数据特征，后者用于抑制模型复杂度，减少过拟合风险。在参数配置方面，本研究将学习率(learning rate)设定为 0.05。此数值可在训练稳定性与模型泛化能力之间取得较为理想的平衡。最大深度(max_depth)设置为 6，用以控制单棵树的结构复杂度，防止模型过度拟合局部特征。子采样比例(subsample)设为 0.8，有助于增加模型的鲁棒性与稳定性。此外，为获得更优的超参数组合，本研究采用五折交叉验证(5-fold cross-validation)对模型进行系统性调参。在训练阶段引入早停机制(early stopping)，即当验证集 AUC 指标在连续 50 轮迭代中未出现显著提升时，训练将自动中断。此策略能够有效防止模型在后期训练过程中出现过拟合现象，从而进一步提高预测性能的可靠性。

3.3. 多行为序列输入设计与特征编码方法

本研究采用阿里巴巴公开的电商用户行为数据，对用户行为序列进行特征构建与建模分析。研究聚焦于四类核心行为：浏览(type=1)、收藏(type=2)、加入购物车(type=3)以及购买(type=4)。在数据预处理阶段，首先以用户 ID 为索引，将所有行为日志按照时间顺序进行排序，从而构建反映用户行为演化过程的序列数据。在序列结构构建完成后，本研究基于 7 天的时间窗口提取多种行为统计特征，包括各类行为的发生频次(如浏览次数、收藏次数等)、行为转化指标(如从加入购物车到购买的转化率)、用户的活跃天数，以及用户在该时间段内访问的不同商品数量等，以反映用户在此期间的参与度与消费潜力。为了进一步捕捉行为在时间维度上的动态变化，本研究引入滑动窗口统计方法，对不同时间区间的短期与中期行为趋势进行提取。该方法能够反映用户行为在近几日或近数周内的变化轨迹，使模型更容易识别兴趣波动及行为周期性特征。此外，为提升特征分布的稳定性，加快模型训练过程的收敛速度，并增强模型在新数据上的泛化能力，本研究对所有序列特征进行了归一化处理，使不同特征被映射到统一区间内，从而降低尺度差异带来的影响。

3.4. SHAP 解释方法

为提升模型预测结果的可解释性，本研究引入 Shapley Additive Explanations (SHAP)方法用于解析 XGBoost 的决策过程。SHAP 基于合作博弈论思想，其核心目的在于衡量各特征对模型输出的边际贡献，从而量化特征在预测结果中的重要程度。在具体实施步骤中，首先采用 TreeExplainer 计算每个样本的 SHAP 值分布，以识别对购买行为预测影响最显著的特征。该步骤能够系统性地揭示特征在样本层面的贡献模式，为后续分析奠定基础。随后，通过特征重要性条形图(feature importance bar plot)呈现整体层面上影响力较大的关键特征，明确各特征在全局预测中的相对作用。此外，为进一步刻画主要特征与购买

概率之间的关联关系，本研究绘制 SHAP 依赖图(SHAP dependence plot)，从而观察特征取值变化对模型预测的具体影响。例如，在分析浏览次数与购买概率的关系时，可能出现类似“阈值效应”的模式：当浏览频次较低时，购买概率提升较为缓慢；而当浏览次数超过某一关键区间后，用户的购买倾向会呈现更明显的上升趋势。该类分析有助于揭示行为特征对决策过程的非线性影响，为理解模型机理提供更坚实的依据。

3.5. 模型评估指标

为全面评估模型性能，在本项研究中，选取了准确率(Accuracy)、AUC (即 ROC 曲线下的面积，Area Under the ROC Curve)以及 F1-score 这几个指标作为主要评价指标。准确率主要用于衡量模型整体预测的正确程度，通过预测正确的数量除以总预测数量得出，该比例即为准确率。例如，若模型预测 100 次，其中 80 次正确，则准确率为 80%。

接着看 AUC，它能够反映模型区分正样本与负样本的能力。以疾病诊断模型为例，正样本为患病个体，负样本为未患病个体，AUC 值越高，表明模型越能精准地区分患病与未患病的人群。

至于 F1-score 指标，它综合考虑了查准率(Precision)与查全率(Recall)。查准率指的是模型预测为正的样本中，实际真正为正样本的比例；查全率是指在所有实际正样本中，被模型正确预测的比例。F1-score 在样本分布不均衡时尤为重要，例如在某个数据集中，正样本数量极少，负样本数量众多，此时 F1 score 能更全面地评估模型性能。

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{1}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

其中，TP、TN、FP、FN 分别表示真正例、真负例、假正例与假负例数量。通过综合比较三项指标，本研究能够从准确性、稳定性与实用性三个维度评估模型在电商用户购买预测任务中的表现。

4. 实验结果与分析

4.1. 模型比较

为验证模型的有效性，本文将 XGBoost 模型与 Logistic Regression、Random Forest 以及 LightGBM 等基线算法展开对比，从 Accuracy、AUC 和 F1 这三项指标对其进行综合评估(详见表 3)。结果表明，

Table 3. Algorithm results
表 3. 算法结果

模型	Accuracy	AUC	F1
Logistic Regression	0.8832	0.8334	0.9379
Random Forest	0.8888	0.8261	0.9395
LightGBM	0.8827	0.8360	0.9377
XGBoost	0.8915	0.8502	0.9407

XGBoost 模型在各项指标上均展现出最优性能, 综合性能较其他模型更为出色。

从该表数据不难发现, XGBoost 模型的 AUC 值高达 0.8502, 显著高于其他模型。这表明在正负样本区分方面, XGBoost 模型具备极强的判别能力。以判断用户是否会购买商品为例, 正样本表示会购买, 负样本表示不会购买, XGBoost 模型能够更为精准地将这两类样本区分开来。此外, XGBoost 模型的 Accuracy 和 F1-score 同样位居榜首。XGBoost 模型在预测方面便呈现出这种可靠状态。这一结果进一步印证了基于梯度提升框架的模型, 在处理非简单线性关系的数据以及应对复杂交互特征时, 的确具有显著优势。并且, 这也说明通过集成学习方法整合多行为序列特征, 能够更为精确地反映用户真实的购买意愿。

4.2. 特征解释分析

从图 2 可以看出, 模型识别出的前四项核心特征分别为“总浏览次数”“总加购物车次数”“访问商品数量”以及“活跃天数”。这些特征的平均 SHAP 值显著高于其他变量, 表明它们在用户购买行为的预测任务中具有最突出的贡献。其中, “总浏览次数”对模型输出的影响最为明显, 说明浏览行为在购买行为形成的过程中具有基础且关键的作用。用户浏览频次越高, 意味着其对商品的关注度越强, 对平台信息的获取也更为充分, 从而更可能产生购买行为。

“总加购物车次数”同样呈现出显著的正向影响。加购行为通常代表用户从兴趣阶段向购买决策阶段的转变, 是购买路径中的重要节点。因此, 加购次数越多, 往往反映用户的购买意愿越明确。“访问商品数量”的重要性排名较前, 说明用户接触商品的种类越丰富, 其潜在购买可能性越高。访问广度较大的用户往往表现出更高的探索意图, 对不同商品类别的持续关注可能带来更多购买机会。此外, “活跃天数”也展现出明显的正向作用, 反映用户在平台上的持续活跃程度与平台黏性密切相关。活跃时间越长的用户, 其消费稳定性与平台依赖度通常更高, 从而更容易产生购买行为。结合图 1 的 SHAP 分布图可以进一步观察到, 高特征取值(图中以红色标示)通常对应正向的 SHAP 值, 说明诸如“总浏览次数”“总加购物车次数”以及“活跃天数”等行为在频繁出现时, 会显著提升购买概率。相反, 特征取值较低(蓝色部分)时, 多数 SHAP 值集中在负向区域, 表明较低水平的行为参与度不足以明显促进购买倾向。综上, 这些结果共同揭示: 高频率、持续性强、覆盖面广的用户行为, 是驱动购买决策的关键行为模式, 而行为互动程度较低的用户则更难形成明确的购买意愿。

相较于前文所述的重要特征, “7 天浏览次数”“7 天加购物车次数”“7 天收藏次数”等短期行为特征, 以及“总收藏次数”等代表弱意向的行为指标, 其 SHAP 值整体偏低, 对模型预测结果的正向或负向影响幅度均较有限。该结果表明, 此类特征在判断用户是否最终产生购买行为时贡献相对较弱, 呈现出不稳定性与指向性不足等特点。基于电商消费行为理论, 可以从以下几个方面进行理解。首先, 短期行为具有显著的波动性, 难以稳定反映用户真实的购买意愿。浏览、收藏或加入购物车等行为在短时间范围内常受到节日促销、平台推荐、库存调整、界面曝光等多重因素影响, 使得行为频次出现强随机性。短期行为的高频并不必然意味着购买概率的同步提升。例如, 用户可能在促销节点进行大量浏览, 但最终并未下单。因此, 这类特征在数值上缺乏趋势性与稳定性, 使得模型难以从其中提取明确的预测信号。

其次, 收藏类行为通常反映较弱的兴趣意向, 其转化链路较长。收藏商品多是用户的“兴趣标记”, 需要经过进一步比较、价格判断、预算确认等多环节才能转化为实际购买。在品类丰富的电商环境中, “大量收藏但最终不购买”是普遍现象, 使得收藏行为在预测任务中难以体现显著贡献。因此, 无论是“总收藏次数”还是“7 天收藏次数”, 其 SHAP 值较低反映了收藏行为在本研究数据中的弱转化效应。第三, 短期行为难以刻画用户长期稳定的消费偏好。与之相比, 历史累计行为(如总浏览量、累计加购量、

活跃天数、累计访问商品数等)更能描绘用户的长期兴趣结构与消费倾向。这类长期特征在 SHAP 分析中呈现更高的重要度,体现了其在用户行为模式刻画方面的优势,也符合消费行为理论中“长期偏好决定总体倾向”的核心观点。最后,在实际电商决策路径中,用户的购买行为通常由长期偏好与预算结构驱动,而短期行为仅起辅助作用。一次性浏览或偶发性的加购行为无法独立决定购买结果,更多只是决策链路中的局部信号。SHAP 结果显示短期行为权重较低,表明模型已有效捕捉到这一行为规律:即长期行为轨迹比短期波动更能决定用户的购买倾向。综上,短期行为特征与收藏类特征在本研究的购买预测任务中贡献有限,主要由于其指示性不足、波动性强及转化链路较长。相对而言,长期行为累积更能反映用户稳定的偏好和真实意向,因此在模型中具有更强的解释力与预测价值。该结果不仅符合电商平台的运营逻辑,也验证了模型在学习不同行为类型的重要性时所做出的合理权重分配。

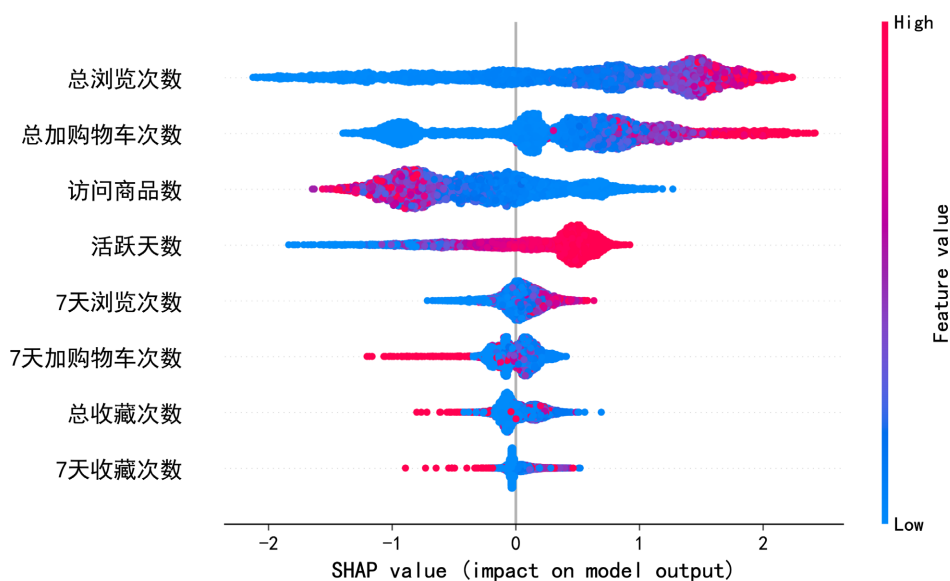


Figure 1. Overview of SHAP

图 1. SHAP 概述图

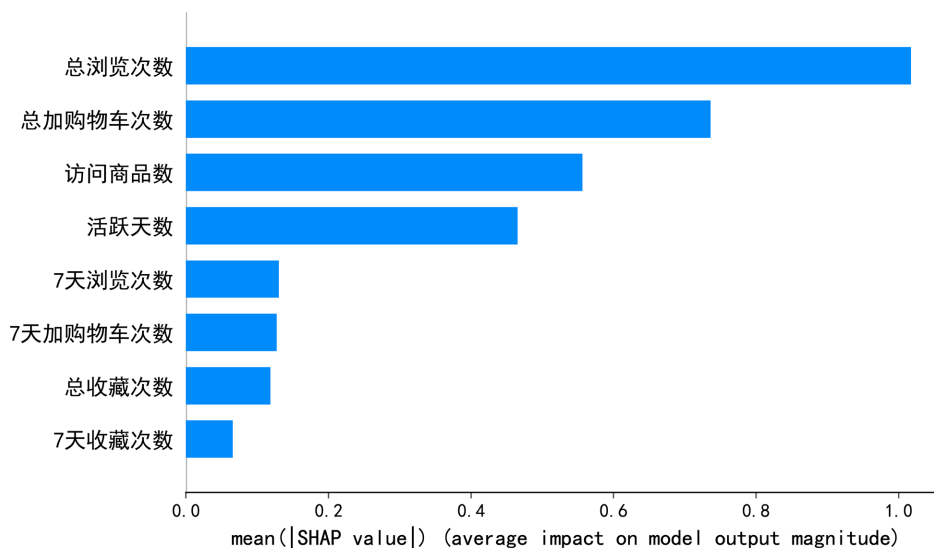


Figure 2. Bar chart of global feature importance

图 2. 全局特征重要性条形图

4.3. 行为特征差异分析

从另一个角度来看,“7 天浏览次数”“7 天加购物车次数”等短期行为特征,以及“总收藏次数”“7 天收藏次数”等涉及收藏的数据,其平均 SHAP 值整体偏低,对模型预测的贡献较为有限。这表明,短期行为和收藏行为在解释用户购买决策时的作用相对较弱。具体而言,收藏行为更多体现用户潜在兴趣,而非即时购买意图;短期行为由于受促销活动、随机浏览或其他临时因素影响波动较大,因此难以稳定反映用户的长期消费倾向。相比之下,长期累计的高频互动行为,例如总浏览次数和总加购物车次数,在揭示用户真实购买意愿方面显示出更显著的作用。这一结果表明,在电商用户购买行为预测中,应更侧重于分析用户的长期行为序列及其在持续平台活动中表现出的行为特征,从而提升模型预测的准确性和实用性。

5. 结论

本研究基于电商平台的多行为序列数据,将 XGBoost 模型与 SHAP 方法结合应用,以系统性地分析影响用户购买行为的关键因素。通过 SHAP 值分析可知,总浏览次数、总加购物车次数、访问商品数量以及活跃天数,是预测用户购买行为的核心特征。在这些特征中,浏览和加购物车行为对用户从兴趣产生到最终购买的转化过程贡献最大,体现了其在购买决策形成过程中的重要作用。访问商品数量能够反映用户在平台上探索商品的广度,活跃天数则体现用户对平台的依赖程度和信任水平。这些长期累计行为和持续互动特征对用户购买可能性均具有显著的正向影响。相比之下,短期行为特征(如“7 天浏览次数”“7 天加购物车次数”)以及收藏行为的重要性相对较低,表明短期行为波动或浅层兴趣难以稳定预测用户的实际购买意图。长期行为轨迹及持续互动特征能够更全面地刻画用户的购买意向,为模型提供更可靠的预测信息。在多行为序列分析中,这类长期累积特征的作用类似于关键拼图,能够更完整地呈现用户的购买倾向和决策模式。尽管本研究在多行为序列数据基础上结合 XGBoost 模型与 SHAP 方法对用户购买行为进行了系统分析,但仍存在若干局限,需要在结果解读与推广时予以谨慎考虑。首先,数据来源单一,本研究采用的是阿里巴巴公开的电商用户行为数据,其用户群体特征和平台运营模式可能存在特定性,因此研究结论在其他电商平台或不同用户群体中的适用性尚需进一步验证。其次,本研究所使用的特征集虽涵盖了浏览、加购、收藏及购买等核心行为,但仍可能遗漏其他影响购买决策的潜在因素,如价格敏感性、促销响应、社交推荐或个性化偏好等,限制了模型解释能力的全面性。此外,XGBoost 虽在处理非线性关系和高维特征方面表现优异,但其预测性能和可解释性仍受参数设置及训练数据分布影响,对于极端样本或稀疏行为序列可能存在一定偏差。最后,本研究基于 SHAP 方法对模型特征贡献进行解释,但 SHAP 分析依赖于所训练的模型结构和特征选择,对于其他类型模型(如深度序列模型)或更复杂的行为交互模式,其可解释性和结论稳健性可能存在一定局限。未来研究可考虑扩展数据来源、引入更多行为与环境特征,并探索多模型集成与跨平台验证,以增强结论的泛化能力和实际应用价值。

参考文献

- [1] Chen, T. and Guestrin, C. (2016) XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, 13-17 August 2016, 785-794. <https://doi.org/10.1145/2939672.2939785>
- [2] Lundberg, S.M. and Lee, S.-I. (2017) A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017*, Long Beach, 4-9 December 2017, 4765-4774.
- [3] Chen, J., Zhang, C. and Xu, Y. (2020) Understanding e-Commerce User Behavior with Multi-Type Sequential Modeling. *Information Systems Research*, **31**, 1234-1252.
- [4] García, S., Luengo, J. and Herrera, F. (2019) *Data Preprocessing in Data Mining*. Springer.

- [5] He, X., Zhang, H., Kan, M.-Y. and Chua, T.-S. (2021) Practical Lessons from Predicting Clicks on Ads at Facebook. *ACM Transactions on Intelligent Systems and Technology*, **12**, 1-18.
- [6] Zhou, Y., Zhao, W. and Tang, J. (2022) Sequential Modeling of Multi-Type e-Commerce Behaviors for Purchase Prediction. *Knowledge-Based Systems*, **250**, Article ID: 109240.
- [7] Wang, H., Chen, L. and Li, Q. (2021) Deep Sequential Models for User Behavior Prediction in E-Commerce. *Expert Systems with Applications*, **175**, Article ID: 114817.
- [8] Li, X., Liu, Y. and Wang, S. (2022) Explainable Machine Learning for e-Commerce Recommendation Systems: A SHAP-Based Approach. *Information & Management*, **59**, Article ID: 103626.