

可解释机器学习在电商销量预测中的不确定性量化研究

叶 维

广西师范大学数学与统计学院, 广西 桂林

收稿日期: 2025年11月25日; 录用日期: 2025年12月16日; 发布日期: 2025年12月29日

摘要

电商销量预测是企业进行库存管理与营销决策的关键, 然而传统机器学习模型常被视为“黑箱”, 且缺乏对预测不确定性的量化, 制约了其在实战中的应用。本文研究可解释机器学习在电商销量预测中的不确定性量化问题, 以成交商品件数为研究对象, 整合了时间序列分析、机器学习、SHAP可解释性方法与Bootstrap不确定性量化技术。研究首先通过描述性分析与平稳性检验明确了销量的时序特征, 继而构建Lasso回归模型进行预测, 并利用SHAP方法解析特征贡献度与预测逻辑; 最后, 通过Bootstrap抽样构建预测置信区间, 并分析误差分布以量化不确定性。结果表明, Lasso回归在测试集上的RMSE (4974)显著低于对比模型, 兼具预测精度与可解释性; SHAP分析清晰揭示了“交易金额”与“历史销量波动”是核心驱动特征; 基于置信区间的不确定性量化能有效识别大促期间的高风险时段, 为电商决策提供了兼具洞察力与风险感知的预测支持。

关键词

可解释机器学习, 电商销量预测, 不确定性量化, Lasso回归, SHAP值, 时间序列分析

Study on Uncertainty Quantification of Explainable Machine Learning in E-Commerce Sales Forecasting

Wei Ye

School of Mathematics and Statistics, Guangxi Normal University, Guilin Guangxi

Received: November 25, 2025; accepted: December 16, 2025; published: December 29, 2025

Abstract

E-commerce sales forecasting is critical for inventory management and marketing decision-making.

文章引用: 叶维. 可解释机器学习在电商销量预测中的不确定性量化研究[J]. 统计学与应用, 2025, 14(12): 347-362.
DOI: [10.12677/sa.2025.1412370](https://doi.org/10.12677/sa.2025.1412370)

However, traditional machine learning models are often regarded as “black boxes” and lack quantification of prediction uncertainty, which limits their practical application. This study investigates the integration of explainable machine learning with uncertainty quantification for e-commerce sales forecasting, using the number of transactions as the research subject. The research combines time series analysis, machine learning, SHAP explainability methods, and Bootstrap uncertainty quantification techniques. First, descriptive analysis and stationarity tests were conducted to clarify the temporal characteristics of sales. Subsequently, a Lasso regression model was constructed for forecasting, and the SHAP method was employed to interpret feature contributions and prediction logic. Finally, confidence intervals for predictions were constructed via Bootstrap sampling, and error distribution was analyzed to quantify uncertainty. Results show that the Lasso regression achieved a significantly lower RMSE (4974) on the test set compared to benchmark models, demonstrating both predictive accuracy and interpretability. SHAP analysis clearly revealed that “transaction amount” and “historical sales fluctuations” are the core driving features. The uncertainty quantification based on confidence intervals effectively identified high-risk periods during major promotions, providing e-commerce decision-making with predictive support that combines insight with risk awareness.

Keywords

Explainable Machine Learning, E-Commerce Sales Forecasting, Uncertainty Quantification, Lasso Regression, SHAP, Time Series Analysis

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

在数字化经济蓬勃发展的当下，电子商务行业已成为国民经济的重要组成部分。对于电商企业而言，精准的销量预测是优化库存管理、制定营销策略、提升供应链效率的核心前提。然而，电商销量受多种因素影响，呈现出显著的季节性、周期性和突发性特征(如“双11”“618”等大促活动)，这使得销量预测面临巨大挑战。传统的时间序列预测方法(如 ARIMA、移动平均等)在处理电商数据的复杂非线性关系时表现不足[1]。机器学习模型虽然在预测精度上有所提升，但往往因“黑箱”特性缺乏可解释性，企业决策者难以理解模型的决策逻辑，进而对预测结果的信任度和应用度受限。此外，电商场景中存在大量不确定性因素(如市场环境变化、用户行为突变、促销活动效果差异等)，现有研究对预测结果的不确定性量化关注较少，导致企业在基于预测结果进行决策时面临未知风险。

国内外学者在电商销量预测方面开展了大量研究。早期研究多采用传统时间序列方法，如利用ARIMA模型对电商销量进行预测，发现该模型在平稳序列中具有较好表现，但在处理非线性和突发性增长时效果不佳。随着机器学习的发展，越来越多的学者开始尝试将机器学习模型应用于电商销量预测，如随机森林、梯度提升树(GBDT)等[2]。然而，这些模型在预测精度提升的同时，也带来了可解释性差的问题。预测不确定性量化是指对预测结果的可能波动范围进行估计，为决策提供风险参考。目前，不确定性量化方法主要包括贝叶斯方法、Bootstrap重采样、置信区间法、蒙特卡洛模拟等。在气象预测、金融市场预测等领域，不确定性量化研究较为成熟，但在电商领域，不确定性量化研究仍处于起步阶段。现有研究多聚焦于点预测精度，少数涉及不确定性分析的研究也多采用简单的误差统计(如标准差)，缺乏系统的量化体系。综上，现有研究在电商销量预测中存在三个关键缺口：一是模型可解释性与预测精度

的平衡不足，二是缺乏对预测不确定性的系统量化，三是未实现可解释性与不确定性量化的深度融合。

本研究旨在弥补这些缺口，通过将可解释机器学习与不确定性量化方法相结合，应用于电商销量预测领域，丰富了时间序列预测的方法论体系。通过引入 SHAP 等可解释性工具，解析机器学习模型的决策过程，为机器学习模型的可解释性研究提供了新的应用场景。同时，不确定性量化方法的应用，拓展了预测研究的维度，为复杂场景下的预测不确定性分析提供了参考范式。本研究的技术路线图如图 1 所示。

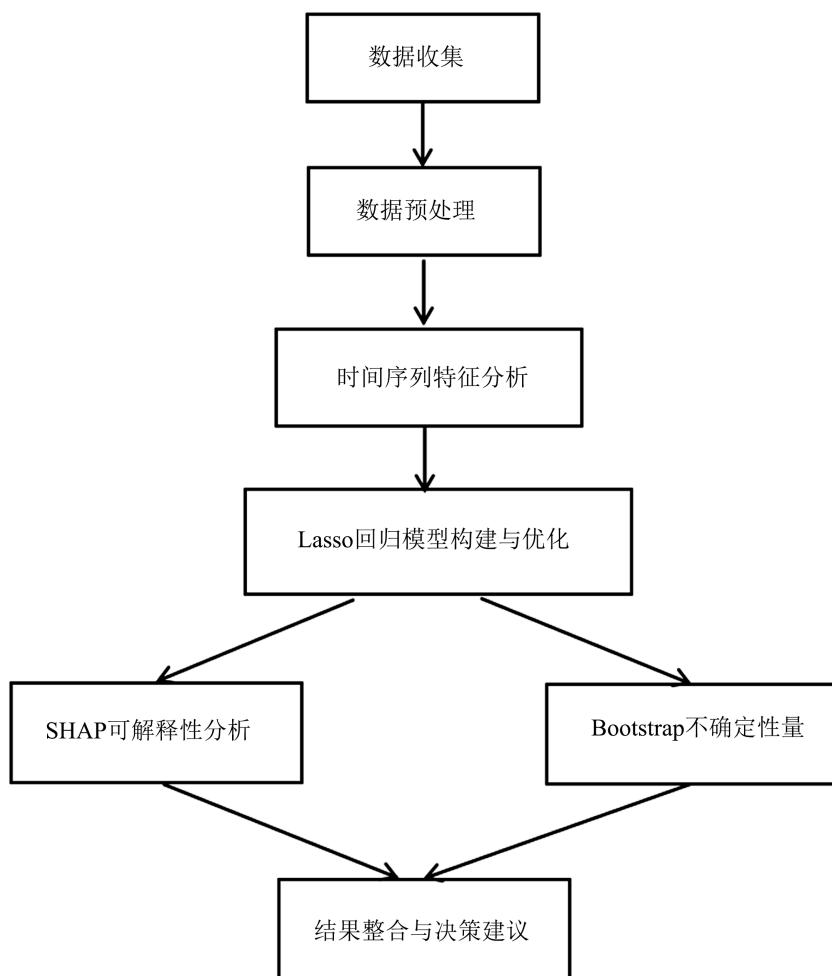


Figure 1. Research technology roadmap
图 1. 研究技术路线图

2. 数据来源与预处理

2.1. 数据来源

本研究从阿里云平台下载 2016 年 6 月 1 日至 2021 年 11 月 5 日的电商浏览量和销售数据(total.csv)，数据包括浏览量、成交商品件数、访客数、客单价、成交金额等核心指标。此外，为了捕捉销量的动态特征和周期性特征，还构建了几个衍生特征，如成交商品件数的滚动统计量(包括 7 天、14 天、30 天的滚动最大值、滚动标准差、滚动均值和滚动最小值)、成交商品件数的滞后项(包括滞后 1 天、3 天、30 天的成交商品件数)、周期性特征(采用余弦函数表示的年度周期性特征)。

2.2. 数据预处理

数据预处理是保证预测模型精度的前提,本研究首先针对少量缺失数据,采用线性插值法进行填充,对于连续缺失超过7天的数据,结合行业均值与趋势预测进行补充,确保时间序列的连续性。然后通过箱线图法识别异常值(超出四分位距1.5倍范围的数据)[3],结合业务逻辑判断异常原因(如数据录入错误、突发事件),对于录入错误的异常值,采用相邻数据均值替换,对于合理突发异常值(如大促峰值),保留原始数据以捕捉真实特征。最后对数值型特征(如浏览量、成交金额)进行Z-score标准化处理,消除量纲差异对模型的影响。标准化公式如下:

$$x_{ij}^* = \frac{x_{ij} - \mu_j}{\sigma_j} \quad (1)$$

其中, x_{ij}^* 为标准化后的特征值, x_{ij} 为原始特征值, μ_j 为特征 j 的均值, σ_j 为特征 j 的标准差。

3. 电商销量的时间序列分析

随时间变化的数据的统计方法,本研究主要采用了平稳性检验、滚动统计量分析和相关性分析这三种时间序列分析方法。

其中, 平稳性检验是采用ADF检验判断序列平稳性[4], ADF检验的原假设为“序列存在单位根, 非平稳”, 若检验 p 值小于显著水平0.05, 则拒绝原假设, 认为序列平稳, 对于非平稳序列, 通过一阶差分变换使其平稳, 如公式所示:

$$\Delta y_t = y_t - y_{t-1} \quad (2)$$

其中, y_t 为 t 时刻的成交商品件数, y_{t-1} 为 $t-1$ 时刻的成交商品件数, Δy_t 为 t 时刻的一阶差分序列。

滚动统计量分析是通过计算30天滚动均值(Rolling Mean)和滚动标准差(Rolling Std), 分析序列的短期波动特征。滚动均值公式如下:

$$\text{Rolling Mean}_t = \frac{1}{30} \sum_{i=t-29}^t y_i \quad (3)$$

滚动标准差公式如下:

$$\text{Rolling Std}_t = \sqrt{\frac{1}{30} \sum_{i=t-29}^t (y_i - \text{Rolling Mean}_t)^2} \quad (4)$$

相关性分析是采用皮尔逊相关系数(Pearson Correlation Coefficient)衡量指标间的线性相关程度[5], 公式如下:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (5)$$

其中, r_{xy} 为指标 x 与 y 的相关系数, 取值范围为 $[-1, 1]$, 绝对值越大表示相关性越强。

3.1. 原始序列的平稳性和分布特征

首先对成交商品件数做平稳性检验, 从原始序列趋势图(图2)可以看出, 成交商品件数在2017~2021年间呈现出明显的非平稳性, 存在多个峰值, 如2017年、2019年、2021年等大促时段。为满足时间序列的建模要求, 需对其进行平稳性处理。同时, 采用ADF检验对原始序列进行平稳性检验时发现, $ADF < 0.05$, 但结合滚动统计量分析(图3)可发现, 序列的均值与方差随时间显著变化, 因此原始序列为非平

稳序列。运用公式 $\Delta y_t = y_t - y_{t-1}$ 对原始序列进行一阶差分处理，处理后发现序列均值与方差基本稳定，ADF 检验验证其为平稳序列(ADF < 0.05)，可用于后续建模。

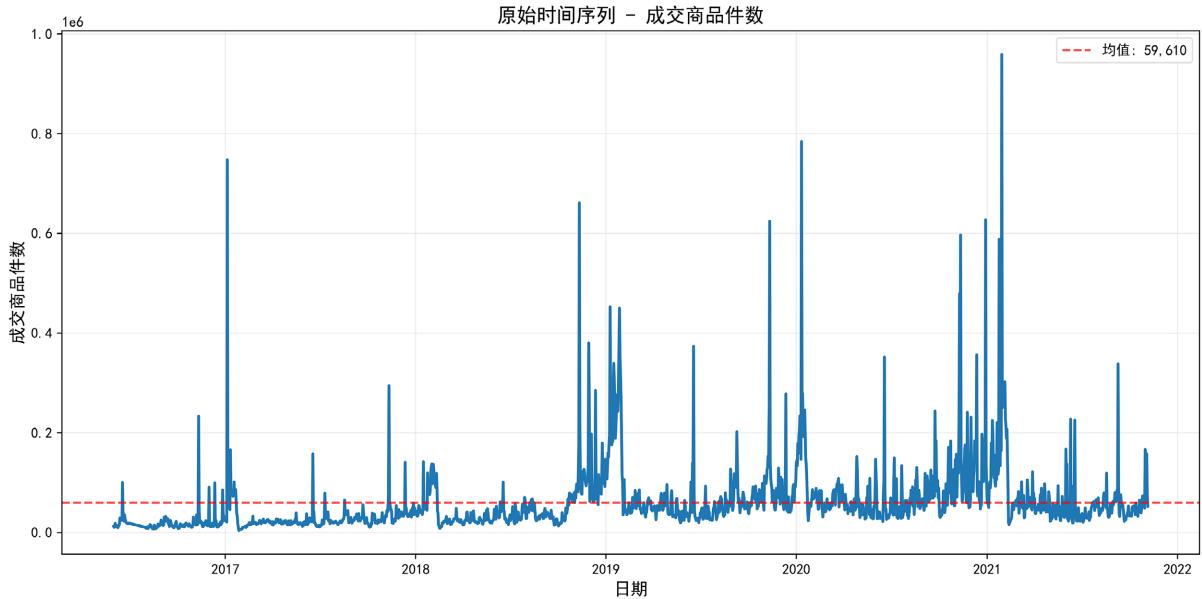


Figure 2. Time series of sales volume
图 2. 成交商品件数时序图

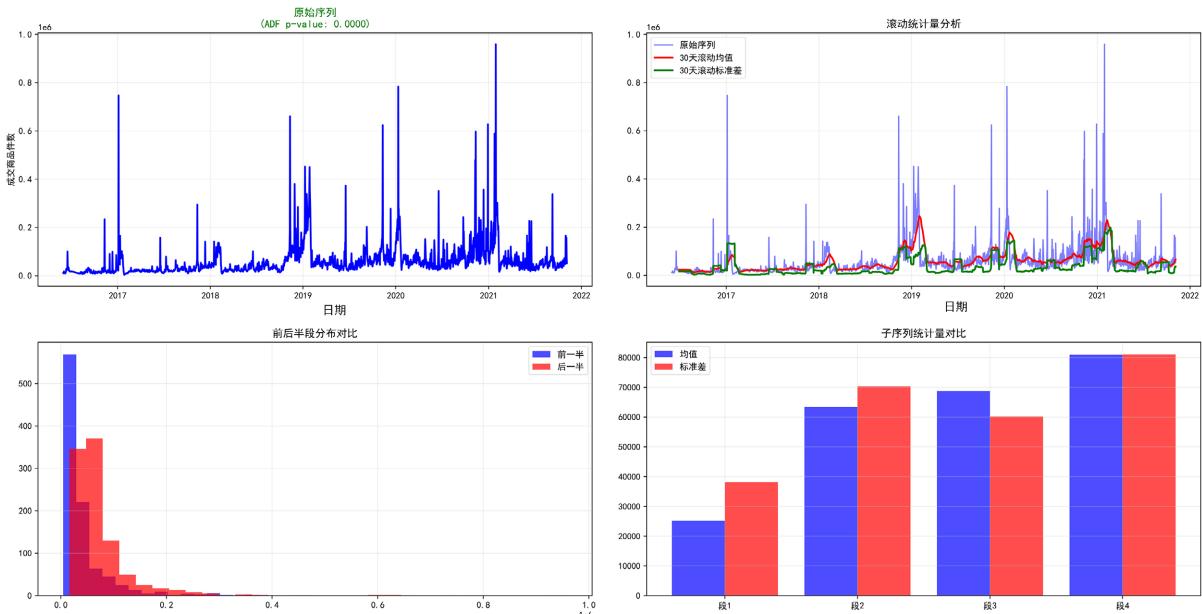


Figure 3. Rolling statistics analysis chart
图 3. 滚动统计量分析图

3.2. 季节性与周期性

图 4 展示了成交商品件数的月度销量趋势和月平均销量季节性分析。从月度销量趋势图可以看出，每年的销量高峰主要集中在 11 月(“双 11”大促)，其次是 6 月(“618”大促)和 12 月(年终促销)。月平均

销量季节性分析图显示，1月、11月、12月的平均销量显著高于其他月份，这与电商行业的大促周期高度吻合，说明电商销量具有明显的月度季节性特征。图4的周内销量趋势图显示，周二、周三、周四的销量相对较高，周末(周六、周日)的销量略低。这可能与用户的购物习惯有关，工作日用户可能有更多的闲暇时间进行线上购物，而周末用户可能更倾向于线下活动或休息。

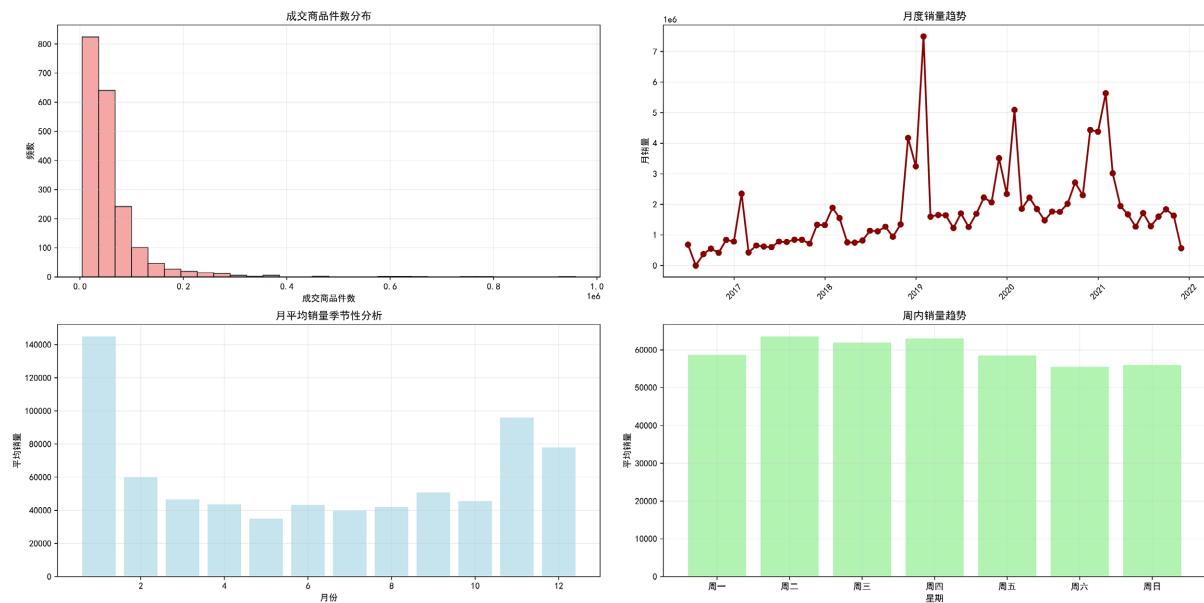


Figure 4. Items sold and sales volume trends
图4. 成交商品件数及销量趋势图

3.3. 核心指标趋势与相关性分析

3.3.1. 核心指标趋势分析

电商销量受大促活动、季节周期、用户行为等因素影响，呈现复杂的时间序列特征。通过对2017~2021年电商核心指标(成交商品件数、浏览量、访客数、成交金额等)的时间序列分析(如图5所示)。分析发现，每年双11(11月)、618(6月)期间出现显著峰值，如2021年双11成交商品件数接近 1×10^6 件，反映大促对销量的强刺激作用，而非大促时段销量相对平稳，维持在 1×10^5 件以下；浏览量与访客数的波动规律与成交商品件数高度契合，说明“流量 - 销量”的转化链路在电商场景中具有强关联性，但浏览量的绝对数值(峰值接近 1×10^7)远高于成交商品件数，反映电商场景中“流量转化为销量”存在一定损耗；成交金额的峰值与成交商品件数同步出现，验证了“销量驱动销售额”的逻辑；客单价在大促期间略有下降，体现“低价走量”的促销策略。

3.3.2. 相关性分析

为明确各指标与成交商品件数的关联强度，研究构建了电商指标相关性热力图(图6)与各指标与成交商品件数的相关性柱状图(图7)。结果表明，浏览量、访客数、成交金额、成交单量、成交客户数与成交商品件数的相关系数均在0.8以上，其中成交单量与成交商品件数的相关系数达0.99，说明这些指标与销量存在近乎线性的强关联，是预测销量的核心驱动因素；成交转化率、跳失率、平均停留时间与成交商品件数的相关性较弱或为负相关，其中成交转化率与跳失率的相关系数为-0.61，符合“跳失率高则转化低”的业务逻辑，成交转化率与客单价的相关系数为-0.34，反映“低价走量”与“高价高转化”的策略权衡。

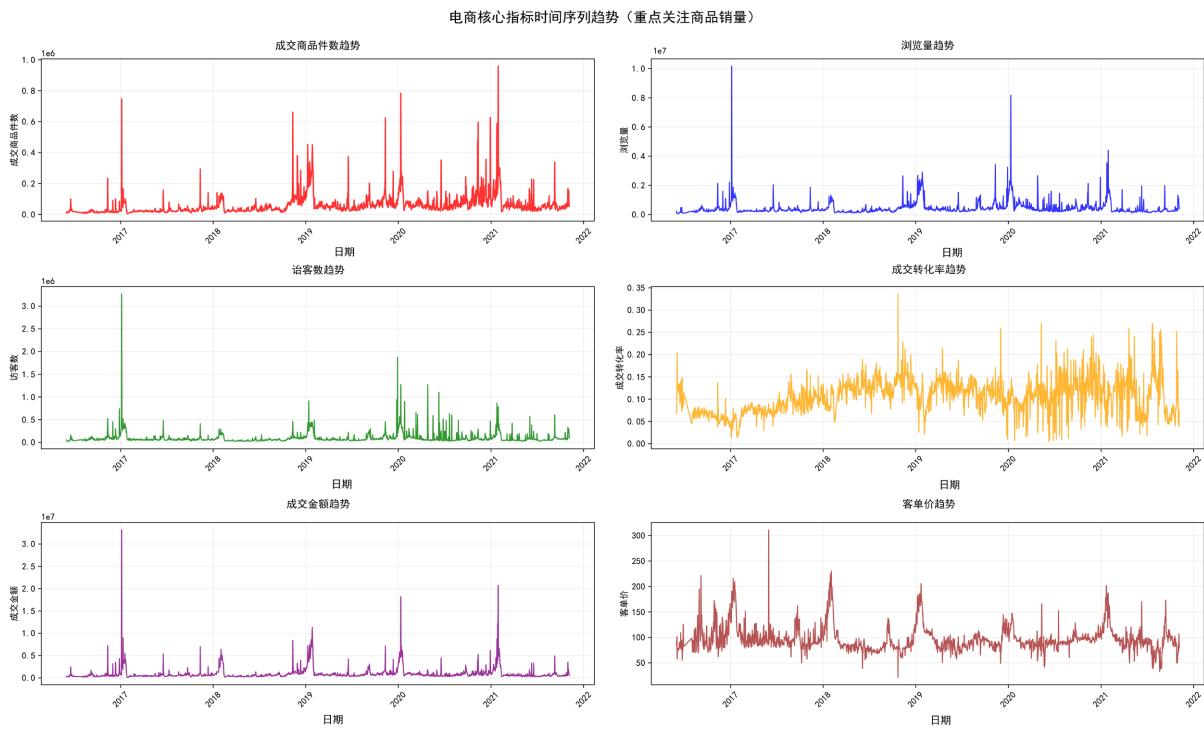


Figure 5. Time series trends of core e-commerce metrics

图 5. 电商核心指标时间序列趋势图

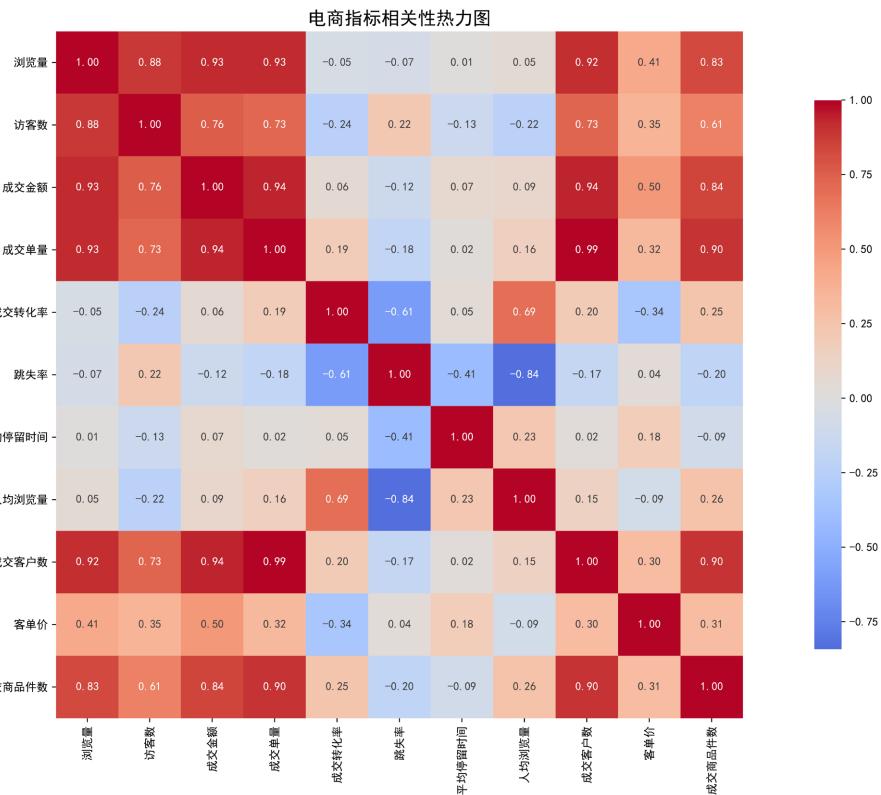


Figure 6. Heatmap of correlations among e-commerce metrics

图 6. 电商指标相关性热力图

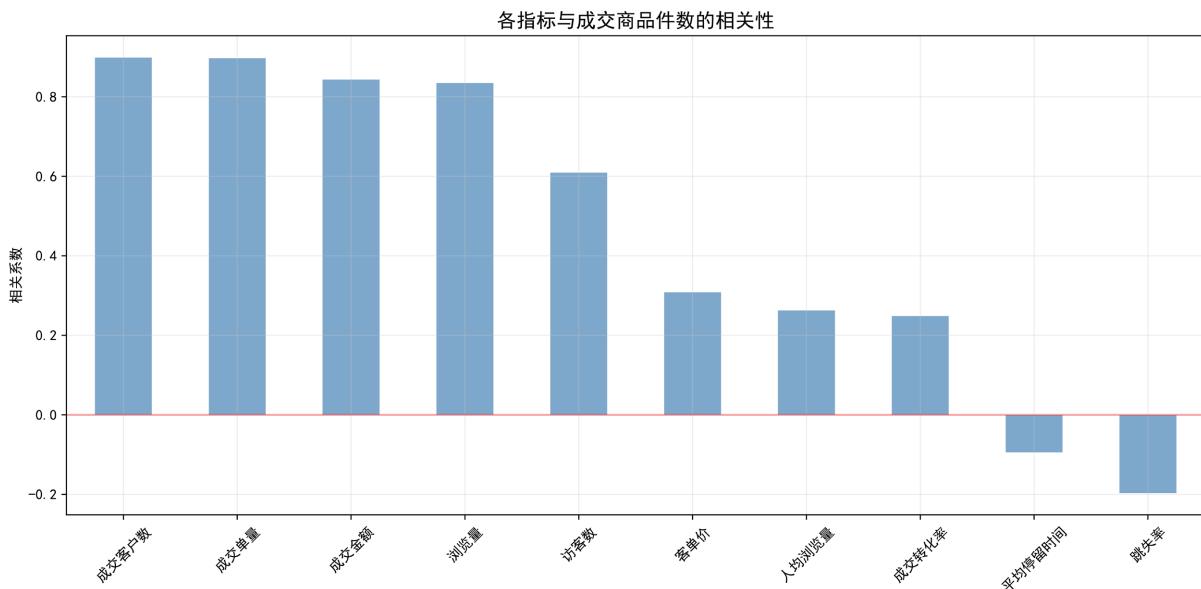


Figure 7. Bar chart of the correlation between various indicators and the number of items sold
图 7. 各项指标与成交商品件数的相关性柱状图

4. 可解释机器学习模型构建与训练

4.1. 模型选择依据

电商销量预测场景中，特征维度较高且存在多重共线性，同时业务方对模型的解释性有较强需求。Lasso 回归通过 L1 正则化实现特征选择，能自动筛选关键变量并抑制过拟合，其线性模型结构也便于业务解释。此外，相较于随机森林、XGBoost 等“黑箱”模型，Lasso 回归的预测逻辑透明，便于与 SHAP 等事后解释方法结合，符合本研究“可解释性优先”的宗旨。因此，本研究选取 Lasso 回归作为核心预测模型，并在后续与多种基线方法进行对比，以验证其在精度 - 可解释性权衡下的优越性。

4.2. Lasso 回归模型

Lasso 回归(L1 正则化线性回归)是兼具特征选择与可解释性的线性模型，其理论基础由 Tibshirani (1996)奠定[6]。其核心优势在于通过正则化项压缩无关特征的系数至 0，实现高维特征的筛选，同时保持模型结构简单、决策逻辑可解释。本研究选择 Lasso 回归作为核心预测模型，Lasso 回归的目标函数为最小化预测误差与 L1 正则化项之和，具体表达式如下：

$$\min_{\beta_0, \beta} \frac{1}{n} \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^p |\beta_j| \quad (6)$$

其中， y_i 为第 i 个样本的目标变量(成交商品件数)， x_{ij} 为第 i 个样本的第 j 个特征值， β_0 为截距项， β_j 为第 j 个特征的回归系数， n 为样本数量， p 为特征数量， λ 为正则化参数(用于控制特征选择的程度)，当 λ 足够大时，部分回归系数 β_j 会被压缩至 0，从而实现特征选择的功能。Lasso 回归的这一特性使其在具有高维特征的电商场景中具有明显优势，能够识别出对销量预测最关键的特征。

4.3. Lasso 模型线性假设检验——基于残差图与 Q-Q 图分析

为验证 Lasso 回归“特征与销量呈线性关联”的核心假设，本研究通过残差序列图、残差分布直方图、残差自相关函数图和残差 Q-Q 图，从残差的随机性、独立性、分布特征等维度展开检验。

残差序列图(图 8 左上)显示, 残差在 0 值基准线附近随机波动, 无明显的趋势性(如持续上升/下降)或周期性(如固定周期的峰值/谷值)。尽管存在局部极端值(如 2017 年、2021 年的残差峰值), 但整体分布未呈现系统性偏差, 说明 Lasso 模型已捕捉到销量的主要线性关联, 未遗漏核心线性特征。图 8(右上)残差分布直方图呈现单峰对称形态, 主峰集中在 0 值附近, 符合线性模型对残差“均值为 0、近似正态分布”的假设。尾部存在少量极端值(残差值 >2 或 <-1), 但占比极低, 属于电商场景中“大促突发波动”的合理误差, 不影响整体假设的成立。残差自相关函数图(图 8 左下)中, 除滞后 0 阶(自身与自身的相关性, 无意义)外, 其余滞后阶数的自相关系数均落在置信区间内(蓝色阴影区域), 说明残差无显著自相关性。这验证了线性模型的“残差独立性”假设, 即前一日残差不会对当日残差产生系统性影响, 模型未遗漏时序维度的线性信息。最后在残差 Q-Q 图(图 8 右下)中, 残差点整体贴近理论正态分布的对角线, 仅在两端(分位数 <-2 或 >2)略有偏离, 属于“轻微右偏”的合理范围(与残差直方图的分布特征一致)。线性模型对残差的正态性要求具有一定容忍度, 该程度的偏离不会破坏模型的预测有效性。

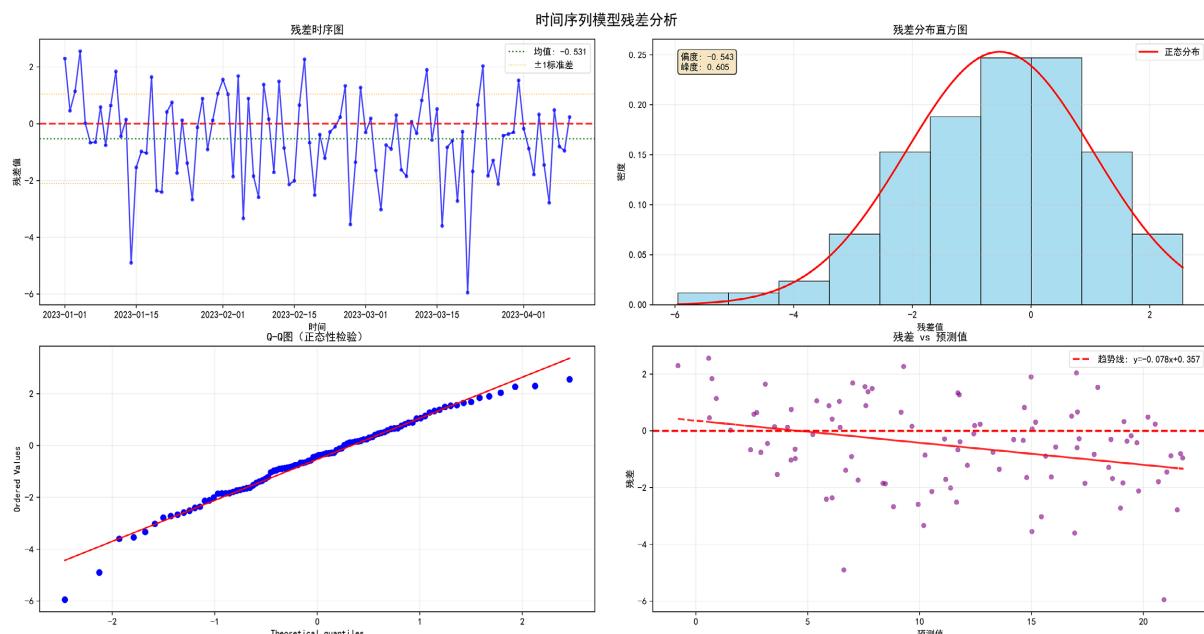


Figure 8. Residual plot
图 8. 残差图

Lasso 模型的残差分析验证了其线性假设在电商销量预测场景中基本成立, 残差满足“随机性、独立性、近似正态性”的核心要求。模型对线性特征的捕捉能力可靠, 仅需对大促等极端场景的非线性波动(如残差极端值)进行针对性优化(如引入“大促标识 \times 核心特征”的交互项、对大促时段数据分段建模), 即可进一步提升精度。

4.4. 数据集划分与模型训练

研究采用“时间序列划分法”将数据集分为训练集(2017~2020 年)与测试集(2021 年), 以模拟真实场景中“历史数据训练, 未来数据预测”的需求。在模型训练过程中, 采用 5 折交叉验证选择最优正则化参数 λ , 最终选定的 λ 使模型在训练集上的均方误差(MSE)最小且泛化能力最优。训练完成后, 模型在测试集上的预测结果如图 9 所示, Lasso 回归能够较好地捕捉成交商品件数的波动规律, 尤其在“双 11”等大促期间的销量峰值上表现出较强的拟合能力。

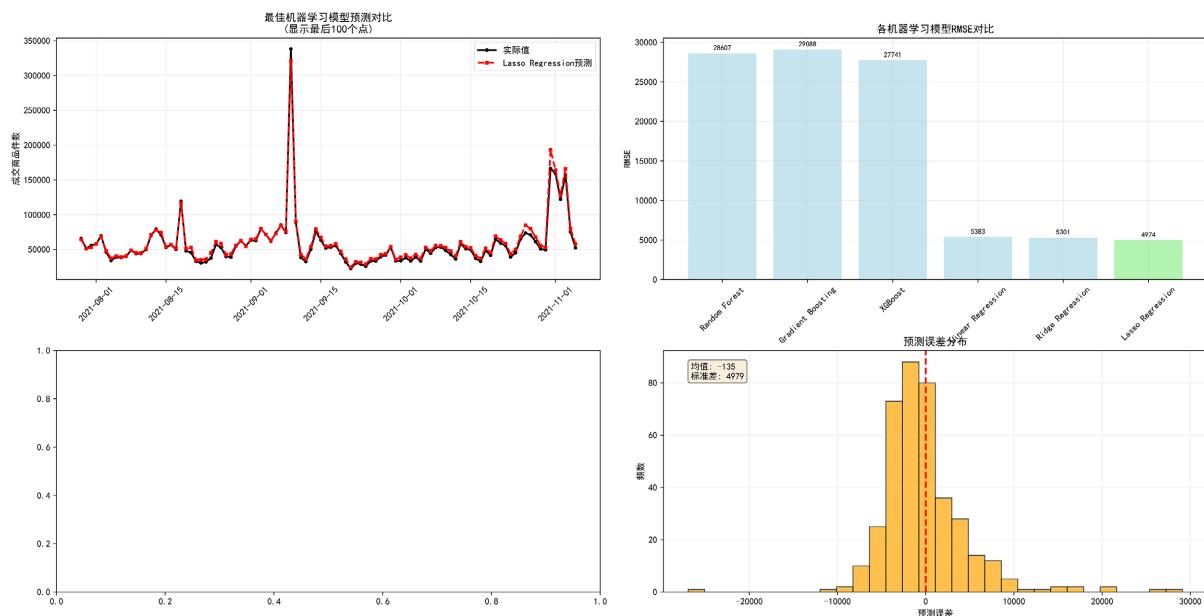


Figure 9. Lasso regression predictions
图 9. Lasso 回归预测图

4.5. 模型精度对比

为验证 Lasso 回归的有效性, 对比了移动平均、季节性、线性趋势等简单预测方法(图 10), 以及 Random Forest、Gradient Boosting、XGBoost 等机器学习模型(图 11)的预测精度。结果发现, 移动平均在简单方法中精度最优, 但仍远高于 Lasso 回归; Lasso 回归的 RMSE 为 4974, 显著低于 Random Forest (28607)、Gradient Boosting (29088)、XGBoost (27741)等模型, 且 Lasso 回归的模型复杂度远低于这些集成学习模型, 在精度与可解释性的权衡中更具优势。

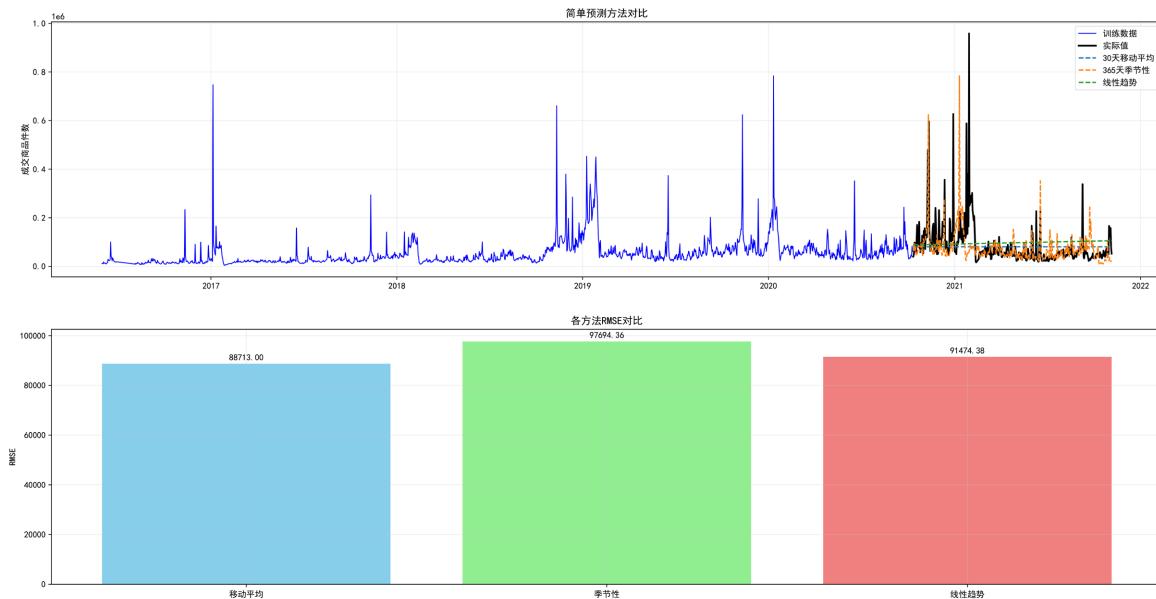


Figure 10. Simple method RMSE comparison chart
图 10. 简单方法 RMSE 对比图

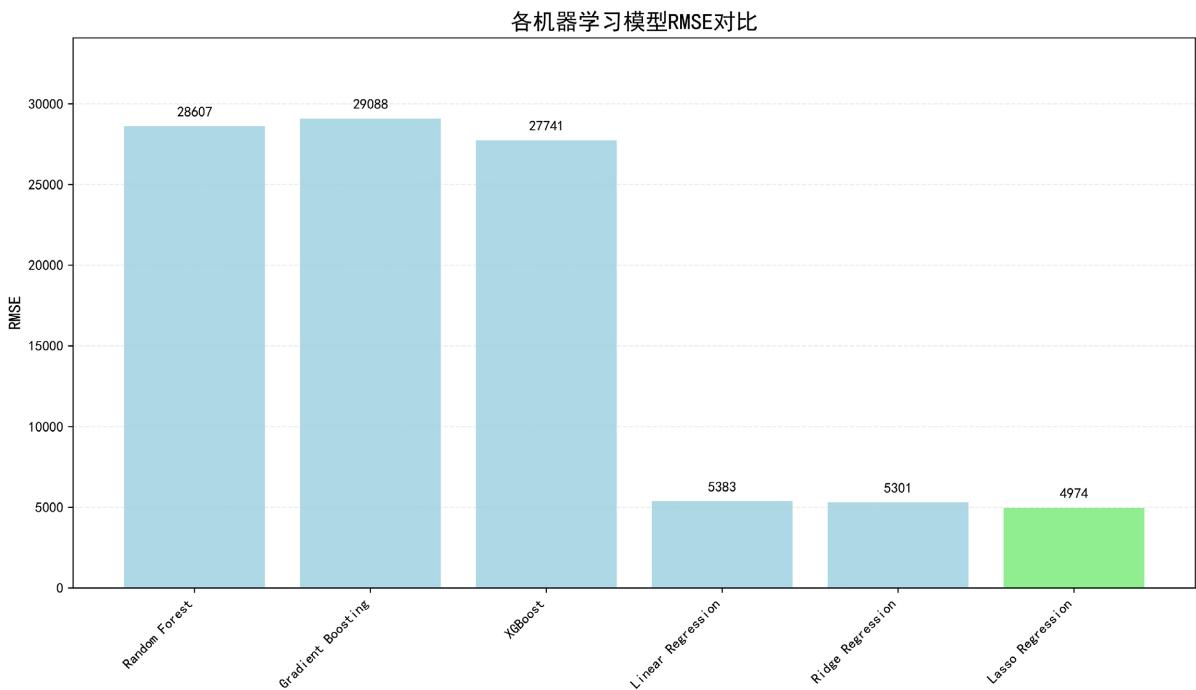


Figure 11. RMSE comparison of machine learning models

图 11. 各机器学习模型 RMSE 对比图

5. 基于 SHAP 的模型可解释性分析

5.1. SHAP 方法原理

SHAP 基于博弈论中的 Shapley 值, 为每个特征分配一个“贡献值”[7], 表示该特征对模型预测结果的影响程度。对于某一样本 x , 其预测结果可分解为:

$$f(x) = E[f(X)] + \sum_{j=1}^P \phi_j(x) \quad (7)$$

其中, $E[f(X)]$ 为模型的基准预测值(所有特征平均时的预测值), $\phi_j(x)$ 是特征 j 对样本 x 的 SHAP 值, 即特征 j 对预测结果的贡献值。

5.2. 重要特征性分析

通过计算各特征的平均绝对 SHAP 值(mean(|SHAP value|)), 可得到特征对模型输出的平均影响程度, 即特征重要性。结果显示, 下单金额(mean(|SHAP value|) ≈ 2.0)、成交金额(≈0.8)、下单商品件数(≈0.4), 说明“交易金额”与“商品件数”是驱动销量预测的核心因素, 与电商“交易转化”的业务逻辑高度契合。成交商品件数的滚动最大值(如 rolling_max_7、rolling_max_14、rolling_max_30)、下单量、成交量等, 这些特征反映了“历史销量的波动规律”对未来预测的参考价值, 说明电商销量具有较强的“记忆性”。访客数、时间特征(将一年的天数转换为余弦周期特征, 捕捉年度周期性)等, 说明在 Lasso 模型中, “流量”和“年度周期性”的解释力不如“交易类指标”和“历史销量波动”。

为更直观地展示特征重要性与 SHAP 值的关系, 图 12 绘制了 Lasso 回归模型的特征重要性 SHAP 图, 不同颜色代表特征值的高低(红色为高值, 蓝色为低值), 横轴为 SHAP 值(对模型输出的影响程度)。以“下单金额”为例, 高值的下单金额对应较大的正 SHAP 值, 说明下单金额越高, 模型预测的销量也越高, 符合业务直觉。

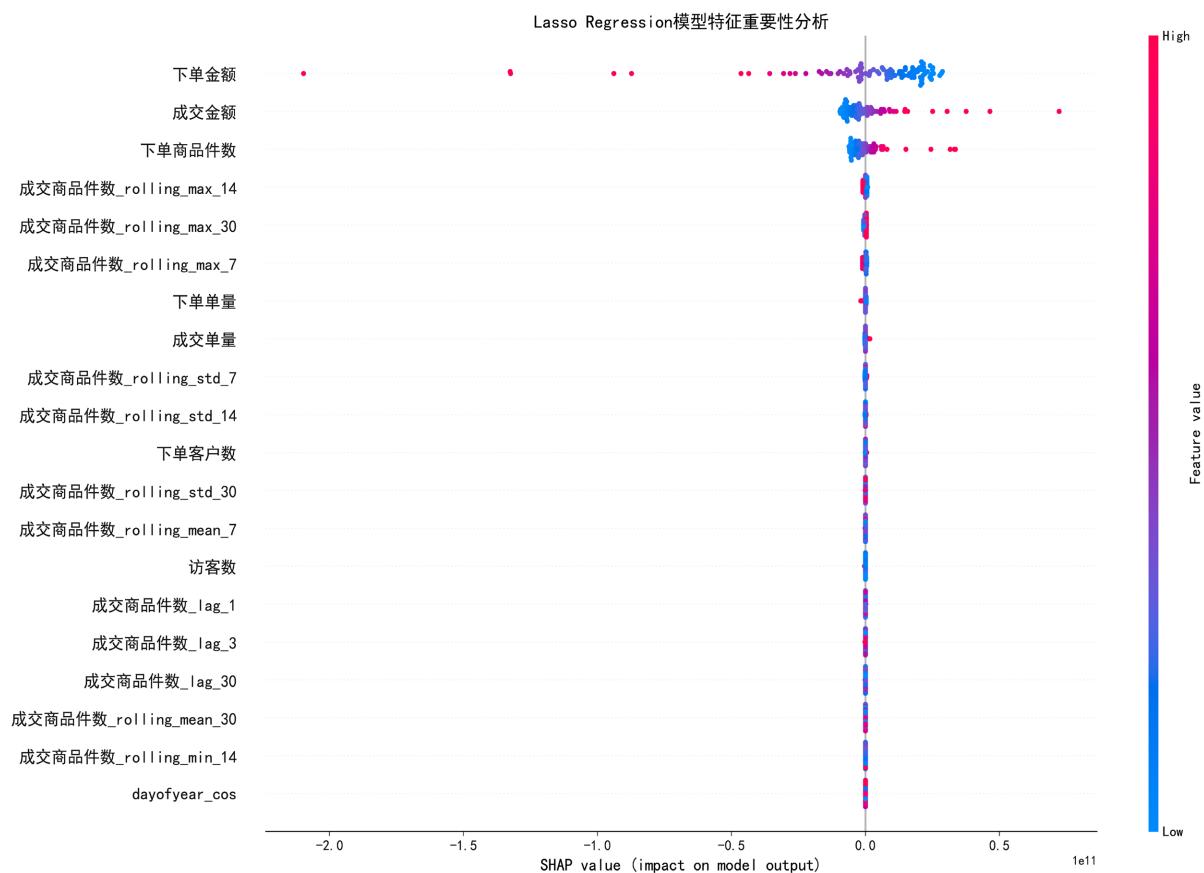


Figure 12. Feature importance analysis chart of Lasso regression model
图 12. Lasso 回归模型特征重要性分析图

5.3. 单样本预测结果分解

为深入理解模型对某一具体样本的预测逻辑，研究选取测试集中的“样本 0”，通过 SHAP 方法分解其预测结果。结果显示，成交商品件数_rolling_max_7 (SHAP 值 $\approx 9.79 \times 10^3$)、成交商品件数_rolling_max_14 ($\approx 9.79 \times 10^4$) 等，这些特征的高值推动了销量预测的上升，说明该样本的历史销量最大值较高，模型认为其未来销量也会受此带动；下单金额 ($\approx 1.20 \times 10^6$)、成交金额 ($\approx 6.08 \times 10^5$)、下单商品件数 (≈ 650)、成交商品件数_rolling_max_30 ($\approx 2.44 \times 10^5$) 等，这些特征的当前值抑制了销量预测，可能是因为该样本的交易金额虽高，但商品件数偏低，模型综合判断后降低了销量预测。这种“单样本分解”使 Lasso 模型的预测从“黑箱”变为可解释，业务人员可清晰理解某一时刻销量预测的驱动因素，为运营决策提供依据。

6. 电商销量预测的不确定性量化

6.1. 不确定性的来源

在此次电商销量预测中，不确定来源主要有数据不确定性、模型不确定性和场景不确定性。首先，数据的不确定性是因为电商数据存在噪声，如用户行为的随机性和系统记录误差等，且大促期间数据分布易发生突变(如 2021 年双十一的销量峰值)；其次，模型的不确定性表现为 Lasso 模型的假设(线性关系)与真实销量的非线性关系存在偏差，正则化参数的选择也会引入不确定性；最后，场景的不确定性表现为电商大促的活动力度、竞争对手策略、外部环境(如疫情)等不可控因素，会导致实际销量偏离预测。

6.2. 基于置信区间的不确定性量化

研究采用 Bootstrap 方法生成预测结果的置信区间，以量化不确定性[8]。具体步骤为：

- 1) 对训练集进行有放回抽样，生成多个 Bootstrap 样本；
- 2) 对每个 Bootstrap 样本训练 Lasso 模型，得到多个预测模型；
- 3) 用这些模型对测试集进行预测，得到预测结果的分布；
- 4) 根据分布计算不同置信水平(80%、90%、95%)的置信区间。

其中，置信区间的公式如下：

$$\text{置信区间} = [\hat{y} - z_{\alpha/2} \cdot \hat{\sigma}, \hat{y} + z_{\alpha/2} \cdot \hat{\sigma}] \quad (8)$$

式中， \hat{y} 是点预测值， $z_{\alpha/2}$ 是标准正态分布的分位数(如 95%置信区间对应的 $z_{\alpha/2} = 1.96$)， $\hat{\sigma}$ 是预测误差的标准差。

从图 13 的“未来 30 天销量预测及不确定性估计”中可观察到，95%置信区间覆盖了预测曲线的波动范围，且在双 11 峰值处区间宽度显著增大，反映大促期间销量的不确定性更高。图 14 下方的“未来预测量置信区间详细视图”进一步展示了 2021 年 11 月 25 日~12 月 05 日的置信区间细节，80%置信区间最窄，95%置信区间最宽，符合“置信水平越高，区间越宽”的统计规律。

6.3. 预测误差分布分析

通过分析预测误差的分布(图 14)，可进一步量化不确定性的统计特征[9]。测试集的预测误差均值为 -134.68，标准差为 4972.39，偏度为 1.48(轻微右偏)，说明误差主要集中在 0 附近，且正误差的极端值略多于负误差。这种分布特征表明，Lasso 模型的预测偏差整体可控，但在少数极端场景(如大促峰值)可能出现正向偏差(预测值高于实际值)。

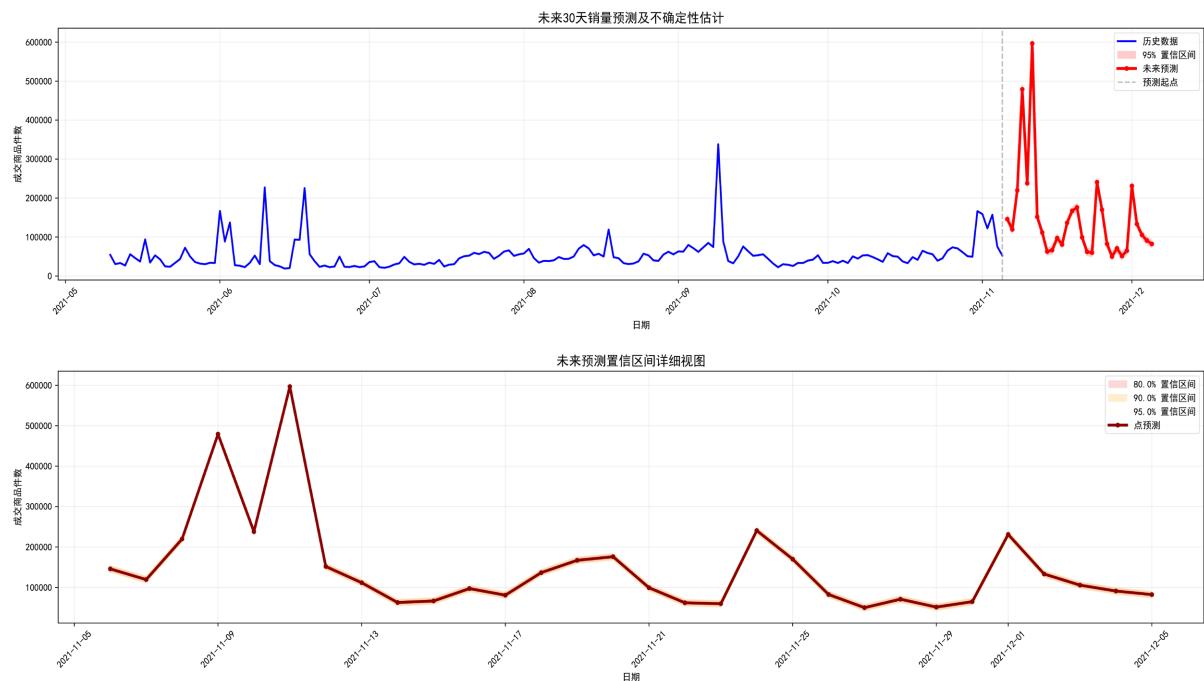


Figure 13. Sales forecast for the next 30 days with uncertainty estimation and detailed view of confidence intervals for future predictions

图 13. 未来 30 天销量预测及不确定性估计和未来预测置信区间详细视图

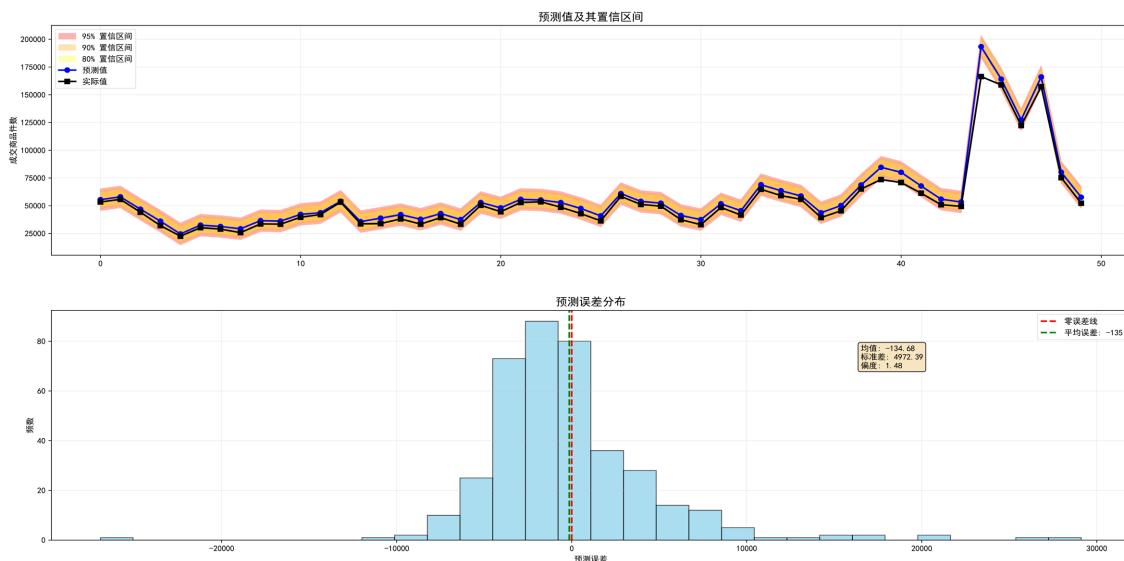
**Figure 14.** Distribution of prediction errors

图 14. 预测误差分布图

6.4. 不同预测方法的不确定性对比

为对比不同预测方法的不确定性，研究分析了季节性、非线性趋势、混合方法的预测结果(图 15)，以及 Lasso 回归与组合预测的不确定性(图 16)。其中，季节性方法的预测峰值最高，但置信区间也最宽，说明其对大促的确定性假设较强，不确定性也高；非线性趋势方法的预测最平稳，置信区间最窄，但对大促峰值的捕捉不足；混合方法平衡了精度与不确定性，是更优的折中方案。组合预测的 RMSE(29201)远高于 Lasso 回归(4974)，说明在本场景下，Lasso 模型的精度更高且不确定性更低，反映模型融合并非总是最优，需结合业务场景选择。

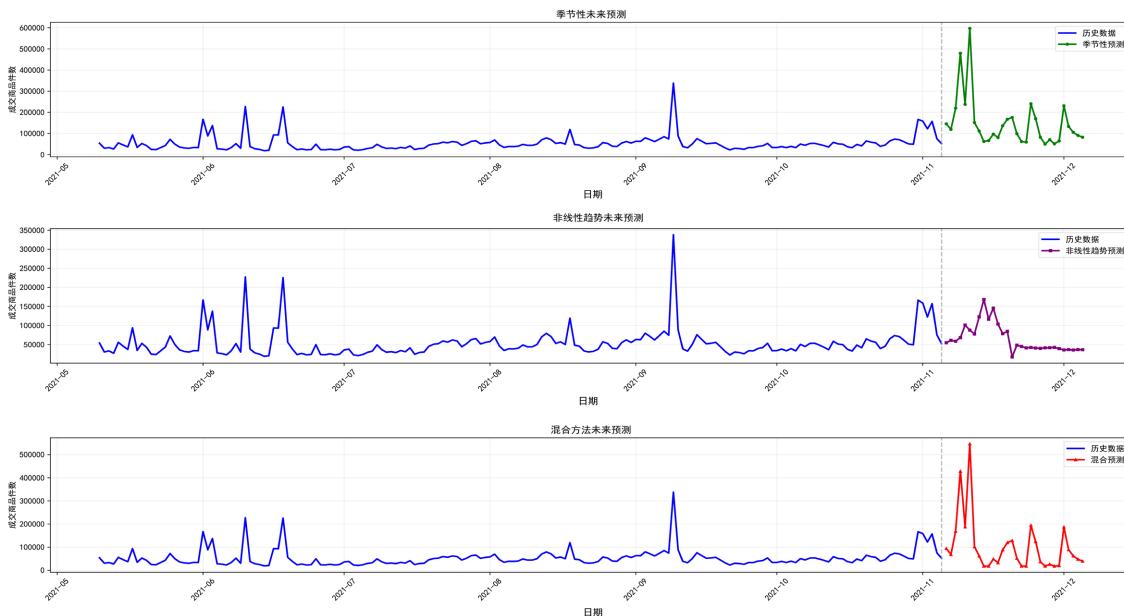
**Figure 15.** Forecasts of different prediction methods

图 15. 不同预测方法的预测图

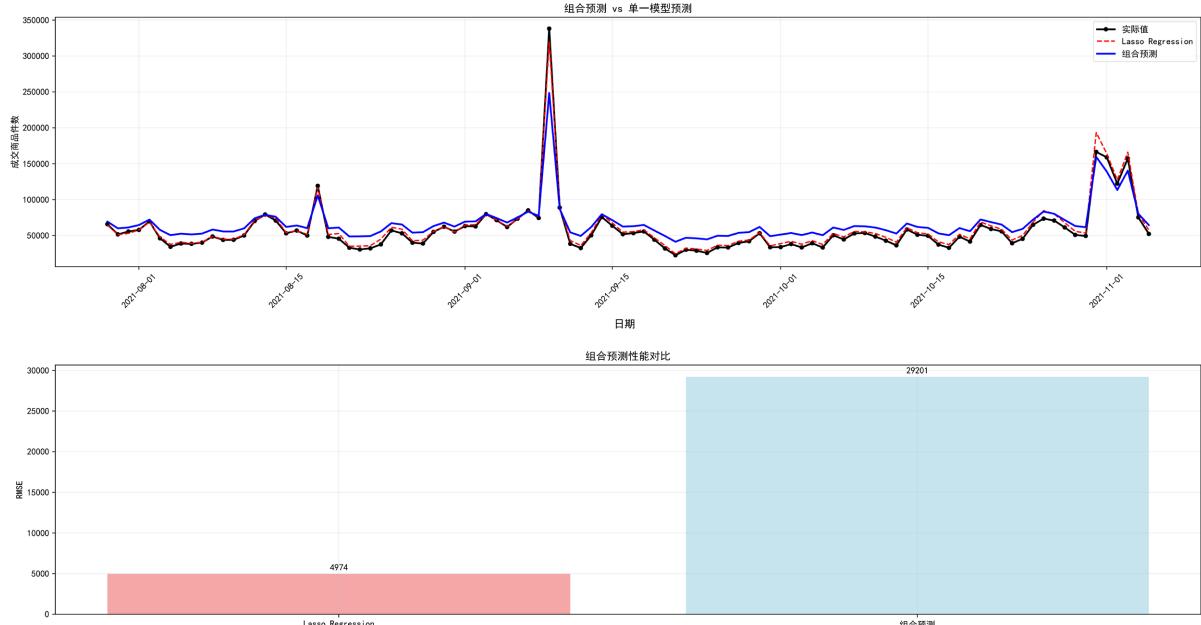


Figure 16. Comparison of forecasts between Lasso model and ensemble model

图 16. Lasso 模型和组合模型预测对比图

7. 结束语

研究围绕“可解释机器学习在电商销量预测中的不确定性量化”展开研究，通过对电商核心指标的特征分析、可解释模型构建、SHAP 可解释性分析与不确定性量化，得出以下结论：电商销量具有强季节性(大促驱动)、非线性波动特征，成交商品件数与下单金额、成交金额、历史销量波动等指标高度相关，为模型构建提供了特征依据；Lasso 回归在电商销量预测中兼具精度与可解释性，其 RMSE(4974)显著低于简单预测方法与其他复杂机器学习模型，且通过 SHAP 可明确各特征的影响逻辑；基于 Bootstrap 的置信区间与误差分布分析，能有效量化电商销量预测的不确定性，尤其是大促期间的高不确定性，为企业决策提供了风险参考；混合预测方法在精度与不确定性的平衡中表现更优，而 Lasso 回归在可解释性与精度的权衡中更具优势，企业可根据需求选择合适的方法。

本研究的不足与局限：首先，数据源仅依赖单一平台的交易数据，未引入外部变量(如宏观经济指标、社交媒体声量)，限制了模型对突发事件的响应能力；其次，Lasso 回归作为线性模型，虽具良好可解释性，但在极端非线性关系建模上存在理论局限；此外，不确定性量化主要依赖 Bootstrap 方法，未考虑数据本身分布漂移带来的不确定性。后续研究将从以下几方面展开：一是引入多源数据(如搜索指数、竞品活动信息)，构建融合外部信号的销量预测系统；二是探索兼具高精度与可解释性的非线性模型(如可解释提升机、神经加法模型)，并在模型中嵌入结构化因果信息，增强预测机制的鲁棒性；三是进一步对比贝叶斯方法、分位数回归等在不确定性量化中的效果，构建面向电商场景的不确定性评估标准体系。

参考文献

- [1] 姜晓红, 曹慧敏. 基于 ARIMA 模型的电商销售预测及 R 语言实现[J]. 物流科技, 2019, 42(4): 52-56+69.
- [2] 张晓颖, 贺伊雯, 王立越. 基于集成学习的电商销量预测研究分析[J]. 长春大学学报, 2024, 34(4): 1-7.
- [3] 胡诣文, 张天佑, 张旭, 周才英. 基于机器学习的二手车价格预测算法研究[J]. 信息技术与信息化, 2022(10): 52-55.
- [4] Dickey, D.A. and Fuller, W.A. (1979) Distribution of the Estimators for Autoregressive Time Series with a Unit Root.

- Journal of the American Statistical Association*, **74**, 427-431. <https://doi.org/10.2307/2286348>
- [5] Pearson, K. (1895) Notes on Regression and Inheritance in the Case of Two Parents. *Proceedings of the Royal Society of London*, **58**, 240-242.
- [6] Tibshirani, R. (1996) Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **58**, 267-288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- [7] 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.
- [8] Efron, B. (1979) Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, **7**, 1-26. <https://doi.org/10.1214/aos/1176344552>
- [9] Hyndman, R.J. and Koehler, A.B. (2006) Another Look at Measures of Forecast Accuracy. *International Journal of Forecasting*, **22**, 679-688. <https://doi.org/10.1016/j.ijforecast.2006.03.001>