

基于机器学习的商超收益优化与定价策略研究

王哲, 杨渠钊, 卢灏, 陈静琳, 梁兰青, 吴延科*

广东海洋大学数学与计算机学院, 广东 湛江

收稿日期: 2023年11月27日; 录用日期: 2023年12月23日; 发布日期: 2023年12月30日

摘要

在生鲜商超中, 蔬菜类商品的保鲜期都比较短, 且品相随销售时间的增加而变差, 隔日就无法再售。本文对多种单品编码进行分析。基于XGBoost模型, 通过PSO算法构建以最大商超收益为目标的模型。经过计算, 第一种单品编码的最优销售单价为6.299999948531562, 最优的销售数量为41.19558, 最大商超收益为110.80144372618932; 第二种单品编码的最优销售单价为102900011030059, 最优的销售数量为30.746147, 最大商超收益为74.18823167795038。本研究的一个关键创新点是通过XGBoost模型和PSO算法, 实现了对不同蔬菜类商品的个性化最优定价。这意味着商超可以根据每种商品的特性和市场需求, 制定最佳的销售单价, 以最大化商超的收益。通过更合理的定价和销售策略, 商超可以减少蔬菜类商品的滞销和浪费, 有助于减少资源浪费, 提高经济效益, 同时对环保也有积极影响。本文的方法为商超提供了科学的管理决策支持, 使其能够更好地应对市场变化和商品特性, 提高了经营效益。

关键词

机器学习, 商超(商业超市), 定价策略, 收益优化, 数据分析

Research on Revenue Optimisation and Pricing Strategy of Superstores Based on Machine Learning

Zhe Wang, Quchuan Yang, Hao Lu, Jinglin Chen, Lanqing Liang, Yanke Wu*

Department of Mathematics and Computer Science, Guangdong Ocean University, Zhanjiang Guangdong

Received: Nov. 27th, 2023; accepted: Dec. 23rd, 2023; published: Dec. 30th, 2023

Abstract

In fresh food superstores, the freshness period of all vegetable items is relatively short, and the

*通讯作者。

文章引用: 王哲, 杨渠钊, 卢灏, 陈静琳, 梁兰青, 吴延科. 基于机器学习的商超收益优化与定价策略研究[J]. 计算机科学与应用, 2023, 13(12): 2623-2628. DOI: 10.12677/csa.2023.1312261

quality deteriorates with the increase of selling time, and they cannot be re-sold on the next day. This paper analyses a variety of single-item codes. Based on the XGBoost model, the PSO algorithm is used to construct a model with the goal of maximising supermarket revenue. After calculation, the optimal unit price of the first single product code is 6.29999999948531562, the optimal number of sales is 41.19558, and the maximal hyper-merchandising revenue is 110.80144372618932; the optimal unit price of the second single product code is 102900011030059, the optimal number of sales is 30.746147, and the maximal hyper-merchandising revenue is 74.1818932; the optimal unit price of the second single product code is 102900011030059 and the optimal number of sales is 30.746147. The maximum hypermarket gain is 74.18823167795038. A key innovation of this research is to realize the personalized optimal pricing of different vegetable products through XGBoost model and PSO algorithm. This means that the supermarket can set the best selling unit price according to the characteristics and market demand of each commodity to maximize the profit of the supermarket. Through more reasonable pricing and sales strategies, supermarkets can reduce the unsalable and waste of vegetable commodities, help reduce resource waste, improve economic efficiency, and have a positive impact on environmental protection. The method in this paper provides scientific management decision support for the supermarket, so that it can better cope with the market changes and commodity characteristics, and improve the operating efficiency.

Keywords

Machine Learning, Superstore (Commercial Supermarket), Pricing Strategy, Revenue Optimisation, Data Analytics

Copyright © 2023 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

商业超市(商超)作为零售行业的关键部分,扮演着满足广大消费者需求的重要角色。然而,商超行业一直以来都面临来自市场竞争、供应链管理和价格波动等多方面的挑战。在如此竞争激烈的市场背景下,商超企业需要不断寻求新的方法来提高其盈利能力和长期竞争力。

商超行业的核心决策之一是定价策略,它在不仅需要根据市场需求和竞争环境来制定价格,还需要考虑产品特性、季节性波动、顾客忠诚度等多种因素。然而,传统的定价方法通常基于经验和常规,往往无法全面考虑这些因素,导致定价策略不够精确和优化[1]。

机器学习技术作为一种强大的数据分析工具,已在多个领域取得显著成就。在商超行业,机器学习的应用可以为企业更好地理解市场趋势、消费者行为和供应链管理的机会。它不仅可以提高定价策略的准确性,还可以优化库存管理、产品定位、促销策略等多个方面[2]。

然而,现有的商超收益优化与定价策略存在一些问题,包括精确性不足、未能全面考虑各种因素和无法应对市场变化。本研究旨在探讨如何应用机器学习技术来解决这些问题,优化商超的收益和定价策略。我们将结合数据分析和机器学习方法,构建优化模型,通过商超的历史数据和市场情报,为企业提供更好的决策支持。

具体来说,本研究将解决以下问题:

1) **精确性不足**: 现有的商超定价策略可能存在精确性不足的问题,导致价格制定不够精确。我们将使用机器学习技术来提高定价策略的精确性,从而更好地满足市场需求。

2) **未全面考虑因素**: 传统的定价方法未能全面考虑产品特性、季节性波动和顾客忠诚度等因素。我们将通过数据分析和机器学习来综合考虑这些因素, 制定更全面的定价策略[3]。

3) **无法应对市场变化**: 市场环境经常发生变化, 传统的定价方法可能无法及时应对这些变化。机器学习可以帮助商超企业更灵活地调整定价策略, 以适应市场的变化。

本研究的目标是为商超行业提供更多有关机器学习应用的见解, 并为企业实现更好的经济效益和市场竞争能力提供指导。通过充分利用机器学习技术, 商超可以更好地满足消费者需求, 提高盈利能力, 降低成本, 并在竞争激烈的市场中脱颖而出。

2. 相关工作

在进行机器学习的商超收益优化与定价策略研究之前, 值得回顾一下与该主题相关的既有研究和相关工作。早期的研究主要集中在价格弹性分析上, 这是一种用于确定价格变动对销售量的影响的方法。这些研究通常使用传统的经济学模型, 以及统计工具来估计价格弹性[4]-[9]。机器学习已经被广泛应用于需求预测。这些方法利用历史销售数据、季节性因素、促销策略等因素来预测未来需求, 为定价策略提供基础。研究人员开始利用机器学习算法, 如强化学习和深度学习, 来优化定价策略。这些方法可以在实时或接近实时的基础上调整价格, 以最大程度地提高收益[10]。机器学习也被用来实现个性化定价策略, 根据顾客的历史购买行为、位置和其他特征来制定不同的价格。这可以提高销售量和顾客忠诚度。机器学习可以用于分析竞争对手的价格策略, 以帮助商超企业调整自己的定价策略, 以更好地竞争。机器学习被用于改进供应链管理, 以确保产品的及时供应, 减少库存成本, 以及提高供应链的效率。机器学习可以帮助商超企业更好地理解线上和线下渠道之间的关系, 从而制定跨渠道的定价策略[11]。研究表明, 机器学习可以用于预测和提高顾客忠诚度, 通过个性化推荐和奖励系统, 鼓励顾客保持购买习惯。

这些相关工作展示了机器学习在商超领域中的广泛应用, 从价格策略到需求预测, 再到供应链管理和顾客忠诚度, 机器学习技术正在为商超企业提供更多工具和洞见, 以提高其盈利能力和市场竞争力。在本研究中, 我们将建立在这些工作的基础上, 探索更多机器学习在商超行业的应用, 并提供具体的建议和实施方案。

3. 模型的建立

3.1. 基于 PSO-XGBoost 的收益优化模型

粒子群优化算法(PSO)是模拟鸟群行为的模型, 主要目的是使粒子能够找到向解空间, 并能找到最优解。同遗传算法类似, PSO 也是一种基于群体迭代的, 但并没有遗传算法用的交叉以及变异, 而是粒子在空间追随最优的粒子进行搜索。PSO 的优势在于简单, 容易实现, 无需梯度信息, 参数少, 适合于处理实优化问题。

在粒子群算法中, 每次迭代粒子的速度 v_i ; 与位置 x_i ; 更新公式如下:

$$\begin{cases} v_i^{t+1} = \mu v_i^t + c_1 r_1 (\text{pbest}_i - x_i^t) + c_2 r_2 (\text{gbest}_i - x_i^t) \\ x_i^{t+1} = x_i^t + v_i^{t+1} \end{cases} \quad (1)$$

其中, μ 为惯性系数, 取值在[0, 8-1.2]; t 表示迭代次数, 一般取 0.9~0.4 线性递减值; c_1, c_2 是加速因子, 通常取 $c_1 = c_2 = 2$; r_1, r_2 在(0, 1)之间随机取值; pbest 表示个体极值, 记录单个粒子发现的最好的位置; gbest 表示全体极值, 记录整个种群经历过的最好的位置[12]。

基于上述条件, 设置搜索的上下界:

上界 = $[c_1 + 0.001]$ (设定一个微小的增量确保 $x > c_1$), 下界 = $[(\max_ + 1) * c_1]$, 考虑到销售单价 x 必须

大于批发价，此处在上界加上微小的增量 0.001。

XGBoost 是对 GBDT 的改进，适用于分类和回归问题。原始的 GBDT 基于经验损失函数的负梯度来构造新的决策树，只能在决策树构造完后进行后剪枝[13]。而 XGBoost 直接在构建阶段就加入了正则项。此外，GBDT 在梯度下降时使用了代价函数的一阶导数，而 XGboost 对代价函数进行二阶泰勒展开，可以同时使用一阶导数和二阶倒数；GBDT 在每轮迭代时使用全部数据，而 XGBoost 则采用了与随机森林相似的策略，对数据进行采样。此外 GBDT 等算法难以并行，效率很低[14]。

3.2. 模型求解

两种编码分别为“102900011016701”以及“102900011030059”。

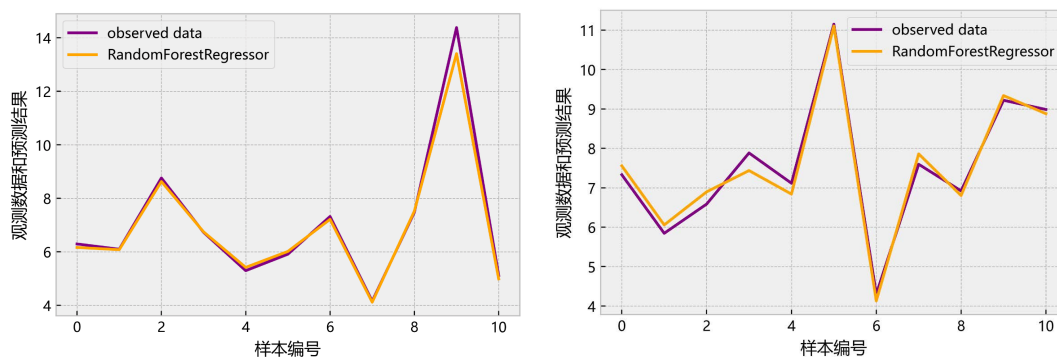


Figure 1. Visualisation of XGBoost model fitting effects

图 1. XGBoost 模型拟合效果可视化

上图 1 为 XGBoost 模型拟合效果可视图，可以看到拟合效果较优。

下图 2 为这两种单品编码经过 PSO 算法变化前后的一个适应函数适应度曲线变化：明显变化后适应曲线更优。

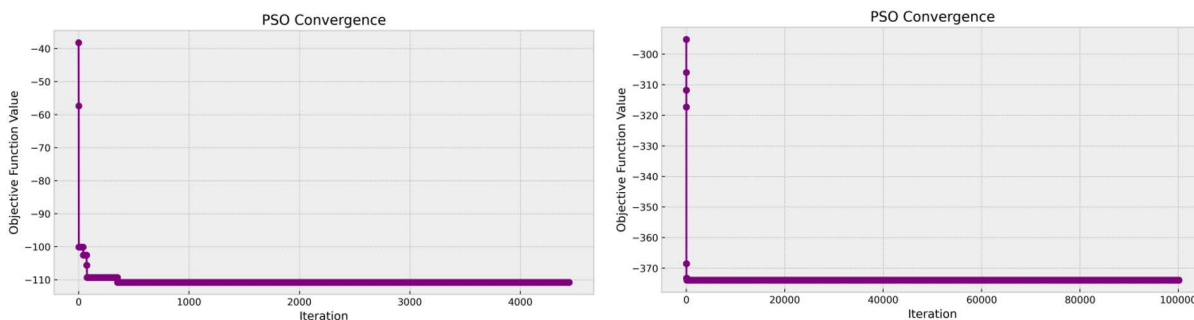


Figure 2. Adaptation curves for the first two single-item coding adaptation functions

图 2. 前两种单品编码适应函数适应度曲线图

基于以上条件，经过算法不断迭代，前两种单品编码最优结果如下表 1 所示：

Table 1. Optimal results for two single-item codes

表 1. 两种单品编码最优结果

最优 x (销售单价)	最优的销售数量	最大的商超收益
6.299999948531562	41.19558	110.80144372618932
4.966432306869793	30.746147	74.18823167795038

上表前两种单品编码“102900011016701”以及“102900011030059”的最优结果。

4. 结论

机器学习在商超行业的应用潜力巨大，可以提高企业的盈利能力和竞争力。通过深入分析数据并应用机器学习模型，商超企业可以在多个领域中获得实质性的好处。机器学习可以通过分析历史销售数据和市场趋势来提高需求预测的准确性。这对于商超企业来说至关重要，因为它们需要满足不断变化的消费者需求。更准确的需求预测可以帮助企业更好地规划库存、减少过剩和缺货，从而提高效益。机器学习可以用于根据实时市场数据来动态调整定价策略。这意味着商超企业可以根据需求、竞争和其他因素来制定最佳的价格，以最大程度地提高收益。个性化定价策略也可以提高销售量和客户忠诚度，因为消费者会感到他们得到了更好的交易。机器学习可以帮助商超企业更好地理解竞争对手的价格策略，并提供市场洞察，以支持决策制定[15]。通过分析竞争者的表现和市场趋势，企业可以更好地定位自己，制定更具竞争力的策略。机器学习在供应链管理方面也具有潜力。它可以用于预测供应链中的瓶颈和风险，以确保产品的及时供应，降低库存风险，提高供应链效率。这对于减少成本并提高服务水平非常重要。通过机器学习，商超企业可以提供个性化的服务和产品推荐，从而提高顾客忠诚度。满足消费者的个性化需求可以增强他们与品牌的联系，并促使他们回头购买。

总的来说，机器学习为商超行业带来了显著的机会，以提高盈利能力、优化定价策略和改进供应链管理。尽管面临一些挑战，但随着技术的不断发展和商超行业的变革，机器学习将继续在该领域发挥重要作用，为企业带来更多创新和竞争优势。未来，我们可以期待更多的研究和实践，以解锁机器学习在商超行业的潜力，为企业的长期成功做出积极贡献。

5. 讨论

在未来的研究中，商超收益优化与定价策略领域仍然有许多有待探索的方向和潜在的研究机会。未来的研究可以深入探讨更复杂的定价策略，特别是在多产品、多渠道和多地区的商超企业中。研究可以考虑不同产品之间的交叉销售、地理差异和渠道特性对定价策略的影响。研究可以集中在实时定价策略的开发和实施上，以更好地应对市场变化和竞争压力。这可能涉及到强化学习、深度学习和实时数据分析等技术的应用。未来的研究可以继续关注供应链管理的改进，以确保产品的及时供应，减少库存成本，并提高供应链的可持续性。这可能包括智能物流和库存管理系统的开发。随着机器学习在商超行业的应用增加，数据质量和隐私问题将愈发重要。未来的研究可以关注如何处理大规模数据并确保数据的质量，同时保护顾客隐私。研究可以集中在更好地理解线上和线下渠道之间的关系，以实现更好的跨渠道整合和互补效应。这可以帮助商超企业更好地满足多样化的消费者需求。未来的研究可以考虑如何通过机器学习来支持商超企业的可持续性和社会责任努力。这可能包括减少食品浪费、提高能源效率和支持社会项目等方面的研究。机器学习在商超行业的应用也会引发政策和法规方面的问题。未来的研究可以关注这些问题，以帮助制定相应的政策框架，平衡商超企业的盈利和消费者权益。未来的研究可以考虑商超行业与其他行业(如物流、制造等)的合作，以实现更高效的供应链和更全面的数据分析。

总之，未来的研究将继续深化商超收益优化与定价策略领域的理解，并为商超企业提供更多创新和竞争优势。机器学习技术将在这一领域中继续发挥关键作用，为企业提供更多工具和见解，以提高其盈利能力和市场竞争力。

基金项目

国家自然科学基金(No. 42275146)，广东海洋大学社科基金(No. C22825)。

参考文献

- [1] Mišić, V.V. and Perakis, G. (2020) Data Analytics in Operations Management: A Review. *Manufacturing & Service Operations Management*, **22**, 158-169. <https://doi.org/10.1287/msom.2019.0805>
- [2] Chen, X., Owen, Z., Pixton, C. and Simchi-Levi, D. (2022) A Statistical Learning Approach to Personalization in Revenue Management. *Management Science*, **68**, 1923-1937. <https://doi.org/10.1287/mnsc.2020.3772>
- [3] Elmachtoub, A.N., Gupta, V. and Hamilton, M.L. (2021) The Value of Personalized Pricing. *Management Science*, **67**, 6055-6070. <https://doi.org/10.1287/mnsc.2020.3821>
- [4] Qu, T., Zhang, J.H., Chan, F.T., Srivastava, R.S., Tiwari, M.K. and Park, W.Y. (2017) Demand Prediction and Price Optimization for Semi-Luxury Supermarket Segment. *Computers & Industrial Engineering*, **113**, 91-102. <https://doi.org/10.1016/j.cie.2017.09.004>
- [5] Guo, S., Chen, C., Wang, J., Liu, Y., Xu, K., Yu, Z., et al. (2019) Rod-Revenue: Seeking Strategies Analysis and Revenue Prediction in Ride-on-Demand Service Using Multi-Source Urban Data. *IEEE Transactions on Mobile Computing*, **19**, 2202-2220. <https://doi.org/10.1109/TMC.2019.2921959>
- [6] Greenstein-Messica, A. and Rokach, L. (2020) Machine Learning and Operation Research Based Method for Promotion Optimization of Products with No Price Elasticity History. *Electronic Commerce Research and Applications*, **40**, Article ID: 100914. <https://doi.org/10.1016/j.elerap.2019.100914>
- [7] Chen, S.S., Choubey, B. and Singh, V. (2021) A Neural Network Based Price Sensitive Recommender Model to Predict Customer Choices Based on Price Effect. *Journal of Retailing and Consumer Services*, **61**, Article ID: 102573. <https://doi.org/10.1016/j.jretconser.2021.102573>
- [8] Pokhariya, J., Mishra, P.K. and Kandpal, J. (2022) Machine Learning for Intelligent Analytics. CRC Press, Boca Raton, 9-234. <https://doi.org/10.1201/9781003269144-13>
- [9] Keskin, N., Li, Y. and Song, J.S. (2022) Data-Driven Dynamic Pricing and Ordering with Perishable Inventory in a Changing Environment. *Management Science*, **68**, 1938-1958. <https://doi.org/10.1287/mnsc.2021.4011>
- [10] Shum, S., Tong, S. and Xiao, T. (2017) On the Impact of Uncertain Cost Reduction When Selling to Strategic Customers. *Management Science*, **63**, 843-860. <https://doi.org/10.1287/mnsc.2015.2355>
- [11] Zhao, J. (2021) Synergy between Customer Segmentation and Personalization. *Journal of Systems Science and Systems Engineering*, **30**, 276-287. <https://doi.org/10.1007/s11518-021-5482-8>
- [12] Wei, W., Feng, L. and Liu, C. (2018) Mixed Causal Structure Discovery with Application to Prescriptive Pricing. *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, Stockholm 13-19 July 2018, 5126-5134. <https://doi.org/10.24963/ijcai.2018/711>
- [13] Zhang, Y., Niyato, D., Wang, P. and Han, Z. (2020) Data Services Sales Design with Mixed Bundling Strategy: A Multidimensional Adverse Selection Approach. *IEEE Internet of Things Journal*, **7**, 8826-8836. <https://doi.org/10.1109/JIOT.2020.2999824>
- [14] Chapparadalli, N.L. (2019) Designing a Dashboard to Support the Decision Process of Dynamic Pricing. Master's Thesis, University of Twente, Enschede.
- [15] Wu, C., Buyya, R. and Ramamohanarao, K. (2019) Cloud Pricing Models: Taxonomy, Survey, and Interdisciplinary Challenges. *ACM Computing Surveys (CSUR)*, **52**, 1-36. <https://doi.org/10.1145/3342103>