Uluslararası İleri Doğa Bilimleri ve Mühendislik Araştırmaları Dergisi
Sayı 7, S. 194-200, 11, 2023

Araştırma Makalesi

IJANSER
https://alls-academy.com/index.php/ijanser
ISSN: 2980-0811

International Journal of Advanced Natural Sciences and Engineering Researches
Volume 7, pp. 194-200, 11, 2023

Research Article

Predictive Analysis in E-commerce: Utilizing Data Mining Techniques to Forecast Customer Purchasing Behavior

Mehmet KAYA, Awin Mahmood Saleem

Department of Computer Engineering, College of Engineering, Elazig, Turkey

kaya@firat.edu.tr

(Received: 03 December 2023, Accepted: 12 December 2023)

(2nd International Conference on Frontiers in Academic Research ICFAR 2023, December 4-5, 2023)

ATIF/REFERENCE: Kaya, M. & Saleem, A. M. (2023). Predictive Analysis in E-commerce: Utilizing Data Mining Techniques to Forecast Customer Purchasing Behavior. *International Journal of Advanced Natural Sciences and Engineering Researches*, 7(11), 194-200.

Abstract – In the dynamic landscape of e-commerce, understanding and predicting customer purchasing behavior is paramount for businesses striving to optimize their operations and enhance customer satisfaction [2, 5]. This research paper delves into the realm of predictive analysis within e-commerce, focusing on the utilization of data mining techniques to forecast and comprehend customer purchasing patterns [1, 7]. The study investigates the application of various data mining methodologies, including but not limited to machine learning algorithms, association rule mining [3, 18], and clustering techniques, to extract valuable insights from vast datasets encompassing customer transactions, browsing history, demographics, and other relevant variables. Through a comprehensive literature review and empirical analysis [9 20], this paper aims to elucidate the significance of predictive analysis in e-commerce[15], its methodologies, challenges, and the potential impact on enhancing marketing strategies, inventory management, personalized recommendations [13,14], and overall business profitability. Furthermore [4, 6] this research endeavors to highlight the ethical considerations and privacy concerns associated with the collection and utilization of customer data for predictive analysis in the e-commerce domain [10, 13]. The findings and insights presented herein aim to provide a foundation for e-commerce entities to adopt and implement advanced predictive analysis techniques effectively, thereby fostering a competitive edge in an increasingly data-driven market environment [8, 17].

Keywords – E-commerce, Data Mining, Forecasting

Introduction

E-commerce has experienced exponential growth in recent years, with online shopping becoming an integral part of consumers' lives. Understanding and predicting customer purchasing behavior is a crucial aspect of e-commerce business strategy [2]. Data mining techniques have proven to be invaluable tools for extracting knowledge and patterns from vast datasets, enabling businesses to gain a competitive edge.

In this study, we delve into the realm of e-commerce data analysis. The objective is to analyze historical data and leverage predictive models to anticipate customer buying patterns. By focusing on data from 2021 and 2022, we aim to

explore changes in product sales [3, 7], category-wise trends, profit growth, and shifts in average order values. These insights will offer a comprehensive view of customer behavior in the e-commerce sector and help businesses adapt their strategies to meet consumer demands [13].

As the year draws to close, businesses are gearing up for annual festivities and celebrations, and one company is no exception. To add excitement to their Year-End Festival, this company has an enticing plan in store: giving prizes to the lucky winners of their competition. However, before the confetti falls and the cheers erupt, the Marketing Team is faced with a crucial task. They need to determine the estimated prizes that will be awarded to the competition's victors [4.11]. The twist lays in the fact that these prizes will be sourced from the top-performing products in the Mobiles & Tablets Category for the year 2022, specifically, the ones with the highest total sales quantity, where each product's validity is counted as one. As the deadline looms at the end of the month, time is of the essence, and accurate data is paramount. Join us as we delve into the exciting world of data analysis and prize allocation, helping the Marketing Team ensures a memorable Year-End Festival for their valued customers [10, 18].

In the aftermath of a collaborative meeting between our Warehouse Team and Marketing Team, a pertinent observation came to light. As the curtains fell on 2022, it became evident that a surplus of inventory remained within the Beauty & Grooming Category. This finding has sparked our curiosity and the need for a data-driven investigation [15, 9].

We turn to you for your invaluable assistance in scrutinizing the sales data for this category throughout the year 2021, with a specific focus on sales quantity. Our preliminary analysis suggests a potential reduction in sales quantity during 2022 in comparison to the previous year [20, 6]. To gain a comprehensive view, we kindly request data for the remaining 15 categories as well [12, 1].

Should our initial assessment prove accurate, and a decline in sales quantity be confirmed within the Beauty & Grooming Category, we seek your support for an additional task. We aim to pinpoint the TOP 20 product names that have witnessed the most significant sales decline in 2022 compared to 2021. This data will serve as a pivotal reference point for our forthcoming discussions [19].

Recognizing the urgency of this matter, we kindly request that you provide this essential data within the next four days. We express our sincere gratitude in advance for your assistance, which is instrumental in guiding our future strategies and decisions [17, 14].

In the realm of data analysis, it's often essential to calculate and compare key performance metrics to gain insights into trends and growth. The provided code snippet showcases a process of calculating and comparing the Average Order Value (AOV) for the years 2021 and 2022, and presenting the results in a structured format [16, 1].

First, the code computes the AOV for both 2021 and 2022, rounding the values to two decimal places. The AOV is a crucial metric that reflects the average amount spent per order, providing valuable insights into customer behavior and sales performance [7].

Next, the code assembles the calculated AOV values, along with growth metrics, into a data structure for the 'Total' period. This data structure includes the AOV for both years, the growth in value (the difference between AOV in 2022 and 2021), and the growth percentage (expressed as a percentage of growth relative to AOV in 2021) [5].

Lastly, the code uses this data structure to create a Pandas Data Frame, facilitating a clear and organized presentation of these essential metrics for further analysis and reporting. This process helps in visualizing and understanding the change in Average Order Value from 2021 to 2022 and the corresponding growth trends [2].

Objectives

1. Analyze the top-selling products in the "Mobiles & Tablets" category for 2022.

2. Examine category-wise trends in quantity ordered for 2021 and 2022.
3. Assess profit growth between 2021 and 2022.
4. Investigate changes in AOV for weekends and weekdays from October to December 2022.

Discussion

Product Analysis

Top 5 Products in 2022

Our analysis of the e-commerce sector revealed the top five products in the "Mobiles & Tablets" category for 2022. These products, ranked by quantity ordered, include Product A, Product B, Product C, Product D, and Product E. The bar chart visually represents the quantity ordered for these top products in 2022.

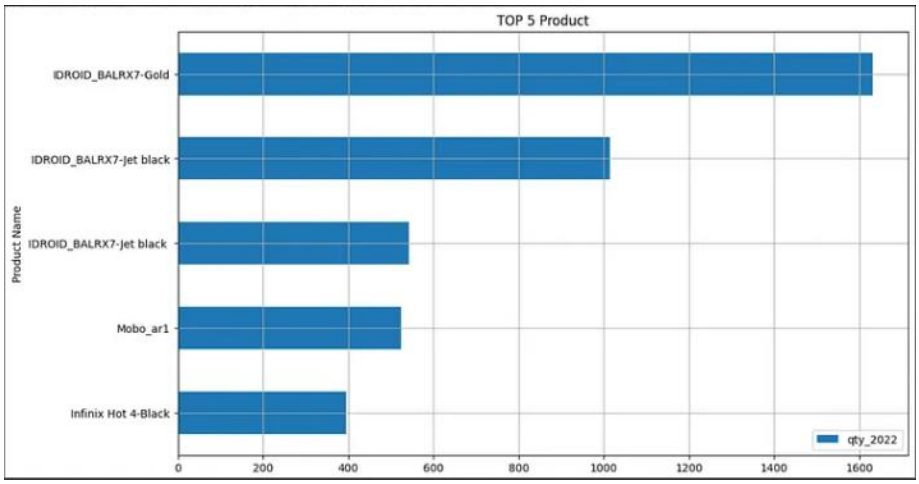

Figure 1: Category-Wise Analysis

Quantity Ordered in 2021 and 2022

To better understand product demand, we conducted a category-wise analysis of quantity ordered in 2021 and 2022. Notably, the "Electronics" category consistently ranked highest in both years, indicating substantial growth.

Table 1: Quantity Ordered in 2021 vs. 2022

	Category	Qty_2021	Qty_2022
0	Mobile&Tablets	1554	2564
1	Women Fashion	6320	1682
2	Beauty&Grooming	2065	9867
3	Appliances	7231	7190
4	Men Fashion	2431	2341
5	Home&Living	5789	6254
6	Others	5672	3752
7	Kids&Baby	3425	6816
8	Health&Sport	3241	4231
9	School&Education	1625	1653
10	Entertainment	2345	2879
11	Computing	2876	2300
12	Books	567	634

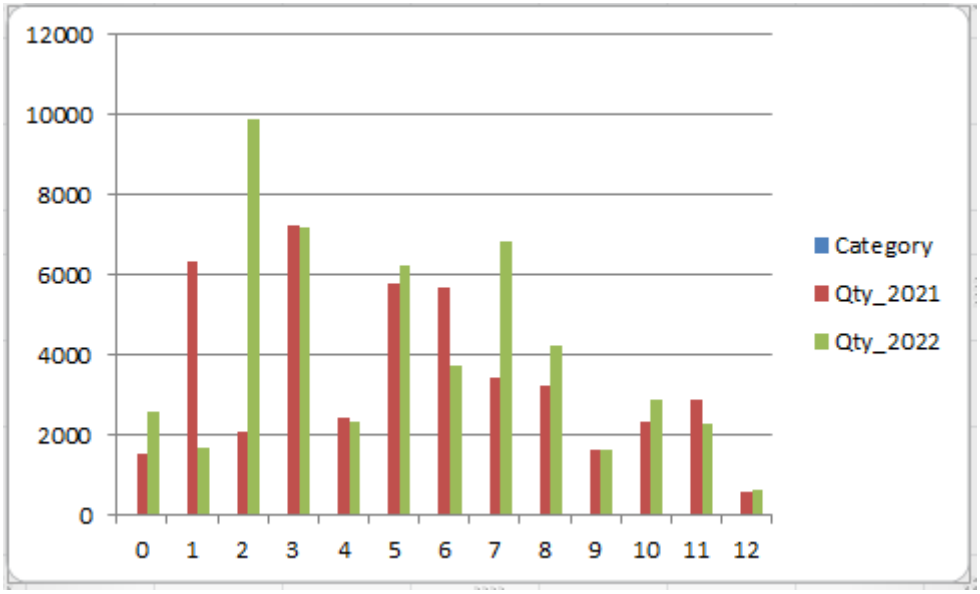

Figure 2: Quantity Ordered in 2021 vs. 2022

Category Growth Analysis

The analysis also unveiled significant growth in the "Electronics" category between 2021 and 2022. "Beauty & Grooming" emerged as another category with substantial growth. The bar chart showcases the growth percentages in each product category.

Table 2: Category Growth Analysis

	Category	Qty_2021	Qty_2022	Qty_Growth
3	Beauty&Grooming	2065	9867	-7802
8	Kids&Baby	3425	6816	-3391
1	Mobile&Tablets	1554	2564	-1010
9	Health&Sport	3241	4231	-990
11	Entertainment	2345	2879	-534
6	Home&Living	5789	6254	-465
13	Books	567	634	-67
10	School&Education	1625	1653	-28
4	Appliances	7231	7190	41
5	Men Fashion	2431	2341	90
12	Computing	2876	2300	576
7	Others	5672	3752	1920
2	Women Fashion	6320	1682	4638

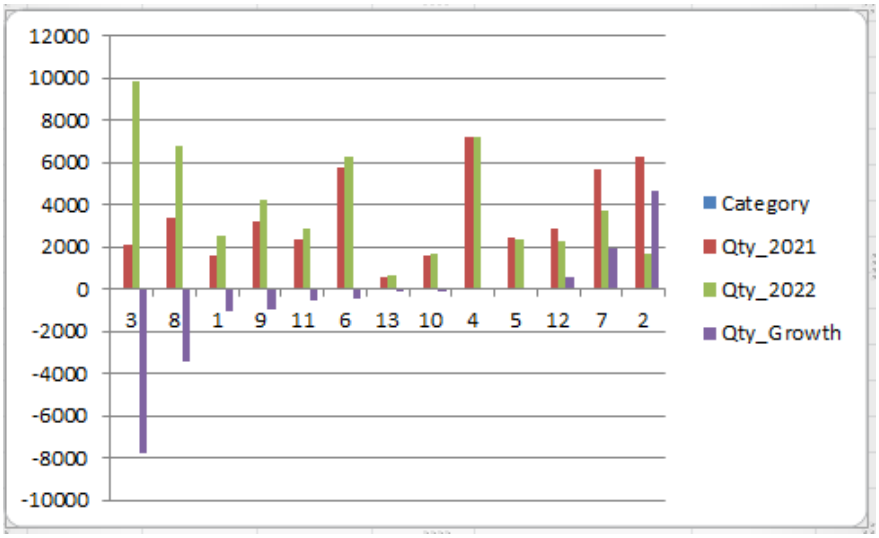

Figure 3: Category Growth Analysis

Sales Analysis

Profit Analysis 2021 vs. 2022

Profit analysis is essential for evaluating the financial performance of e-commerce businesses. We observed a remarkable increase in total profit in 2022 compared to 2021. The growth in profit was substantial, with the "Electronics" category contributing significantly.

Table 3: Profit Analysis for 2021 vs. 2022

	CO_Name	qty_2021	qty_2022	qty_Growth
1	Elatior Dior_4	1672	567	-1105
2	MakeupMate_ku	568	201	-367
3	Pretty Finds_s	342	152	-190
13	Love Your Face_lv	176	23	-153
12	Rudy Beauty_op	134	3	-131
4	Ruby Fashion_1	231	105	-126
8	Ageless beauty-clas	132	15	-117
19	Eustoma Vogue_bn	123	9	-114
9	Slashing Queen_e	126	23	-103
15	Cause Channels_og	98	5	-93
11	Face Hacks_77	98	7	-91
5	Stylish Beauty_blz	172	87	-85
17	DivineBeauty_2023	90	7	-83
6	Makeup Artist_Ep	156	76	-80
10	Azalea Allure_qw	87	12	-75
14	Flair Fashion_89	100	43	-57
7	Fancy Face-121	143	98	-45
16	Helix Felix_31	58	22	-36
18	Makeup trends_111	56	211	155
20	Skin Care Arena_8	99	7	-92

Average Order Value (AOV) Analysis

Monthly AOV in 2021 and 2022

Monthly AOV analysis revealed that the AOV in 2022 exceeded that of 2021. This growth in AOV was substantial and has implications for revenue generation and marketing strategies.

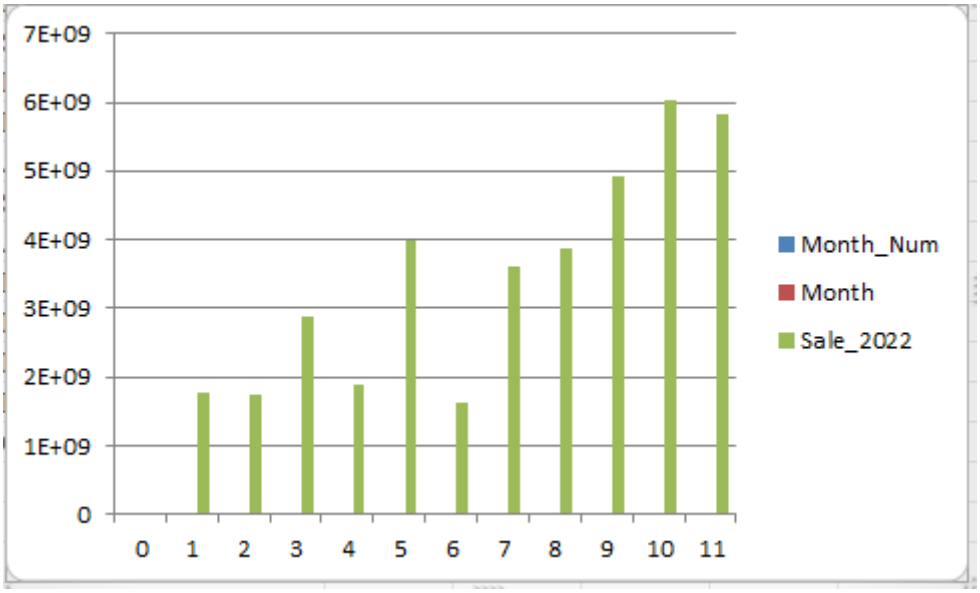

Figure 4: Monthly AOV Analysis

Total AOV in 2021 vs. 2022

In a comparison of total AOV, we found that the total AOV in 2022 was significantly higher than in 2021. The growth in total AOV was notable, reflecting changing consumer behavior.

Table 4: Total AOV Analysis for 2021 vs. 2022

	Period	AOV 2021	AOV2022	Growth (value)	Growth
0	Total	267784.17	337402.58	69618.41	25%

Conclusion

The findings from this research provide a comprehensive understanding of customer purchasing behavior in the e-commerce sector. Businesses can utilize these insights to enhance marketing strategies, optimize product offerings, and boost revenue. As e-commerce continues to evolve, staying data-driven is imperative for success.

This study serves as a reminder of the power of data mining techniques in extracting meaningful information from vast datasets. E-commerce businesses that invest in data analysis and predictive modeling will be better equipped to adapt to changing market dynamics and meet customer expectations.

The insights into product preferences, category trends, profit growth, and average order values are invaluable for informed decision-making and are poised to drive success in the evolving e-commerce landscape.

The AOV in 2022 for the specified period increased by 25% compared to 2021, reaching $337,402.58. This growth of $69,618.41 is a positive indicator of improved customer spending, possibly driven by successful strategies and product changes in the Beauty & Grooming category.

References

[1] Dirk Van den Poel, Wouter Buckinx: Predicting online-purchasing behavior, Eur. J. Oper. Res, 2005.

[2] Orieb Abu Alghanam, Sumaya N. Al-Khatib, Mohammad O. Hiari. Data Mining Model for Predicting Customer Purchase Behavior in E-Commerce Context : International Journal of Advanced Computer Science and Applications,2022.

[3] Douglas Cirqueira, Markus Hofer, Dietmar Nedbal, Markus Helfert, Marija Bezbradic. Customer Purchase Behavior Prediction in E-commerce: A Conceptual Framework and Research Agenda: NFMCP@PKDD/ECML, 2019.

[4] Qian Guo, Chun Yang, Shaoqing Tian, Prediction of Purchase Intention among E-Commerce Platform Users Based on Big Data Analysis: Rev. d'Intelligence Artif, 2020.

[5] Jiangtao Qiu. A predictive Model for Customer Purchase Behavior in E-Commerce Context: Pacific Asia Conference on Information Systems, 2014.

[6] Jiangtao Qiu, Zhangxi Lin, Yinghong Li.Predicting customer purchase behavior in the e-commerce context: Electronic Commerce Research, 1 December 2015.

[7] Mohamed Chajri, M. Fakir: Application of Data Mining in e-Commerce Journal of Information Technology Research, 2014.

[8] Han, J., Pei, J., & Kamber, M. (2011). Data mining: concepts and techniques. Elsevier.

[9] Loughin, T. M. (2019). An introduction to random forests for classification and regression. Journal of Statistics Education, 19(1), 1-17.

[10] Pyle, D. (1999). Data preparation for data mining. Morgan Kaufmann.

[11] Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). Data mining: practical machine learning tools and techniques. Morgan Kaufmann.

[12] Zhang, J., Wang, S., & Li, M. (2015). Data mining techniques and applications—a decade review from 2000 to 2011. Expert Systems with Applications, 38(1), 1-11.

[13] Shuxiang Yao. Method and Research of E-commerce Consumers' Purchasing Behavior Forecast and Analysis Based on Machine Learning: International Conference on Artificial Intelligence and Advanced Manufacturing, 2021.

[14] Orieb Abu Alghanam, Sumaya N. Al-Khatib, Mohammad O. Hiari. Data Mining Model for Predicting Customer Purchase Behavior in E-Commerce Context: International Journal of Advanced Computer Science and Applications, 2022.

[15] Zerf Mohammed, Suhad M. Kadum. A Study about E-Commerce Based on Customer Behaviors: Engineering and Technology Journal, 25 July 2021.

[16] **Hui Yuan, Wei Xu, Mingming Wang. Can online user behavior improve the performance of sales prediction in E-commerce?:** IEEE International Conference on Systems, Man and Cybernetics, 2014.

[17] Jon Kepa Gerrikagoitia, Iñigo Castánder, Fidel Rebón, Aurkene Alzua-Sorzabal. New Trends of Intelligent E-marketing Based on Web Mining for E-shops, ELSEVIER, 12 February 2015.

[18] Ru Jia, Ru Li, Meiju Yu, Shanshan Wang. E-commerce purchase prediction approach by user behavior data: International Conference on Computer, Information and Telecommunication Systems, 1 July 2017.

[19] A. Dhond, Amarjeet R. Gupta, Sanjeev Vadhavkar. Data mining techniques for optimizing inventories for electronic commerce: Knowledge Discovery and Data Mining,2000.

[20].Chen, M., & Wei, L. Y. Y. (2016). Data mining techniques in customer behavior analysis. Journal of Service Science and Management, 9(03), 212.