J. Electrical Systems 20-2s (2024): 700-708

[1]Bijith Marakarkandy

[2]Bhushankumar Nemade,

[3]Sandeep Kelkar

[4]P. V. Chandrika

[5]Vaishali A.Shirsath,

[6]Mahesh Mali

Enhancing Multi-Channel Consumer Behavior Analysis: A Data-Driven Approach using the Optimized Apriori Algorithm

Journal of Electrical Systems

***Abstract:* -** Purpose –In the dynamic arena of retail and e-commerce, understanding customer behavior is crucial. This research paper investigates the strategies customers use to balance convenience and cost. It notes that consumers often divide their purchases, choosing online platforms for items they perceive as reasonably priced, while preferring physical stores for products associated with high shipping costs, lack of express shipping, or higher online prices. Traditional retail analysis methods, such as identifying frequently purchased and co-occurring items, are critical to the sector. However, these insights may be skewed if the analysis is solely confined to the online shopping cart, neglecting related purchases made at physical stores either before or after the online transaction. This paper addresses this analytical gap by looking into the complex interplay between online and offline purchases. The objective is to provide a more comprehensive, detailed understanding of multi-channel shopping behaviors, allowing retailers to better cater to their customers' needs and preferences while optimizing their strategies for a more effective market presence. This comprehensive approach attempts to uncover deeper insights into the interactions between online and offline purchases, thereby contributing to a more comprehensive understanding of consumer behavior in the retail and e-commerce industries.

Design/methodology/approach- A detailed survey questionnaire was administered, obtaining 2465 responses. The questionnaire was specifically designed to capture the complexities of online and offline purchases, with an emphasis on Apple's purchase channels. The Apriori algorithm was used twice to find frequently co-occurring items that were purchased online and then online clubbed with offline. Essential metrics like support and confidence were calculated for both online and offline purchases. To determine any significant differences between the groups, an independent-samples t-test was used. This comprehensive methodology ensures a thorough examination of multi-channel shopping behaviors, providing valuable insights into consumer behavior in the retail and e-commerce sectors.

Findings–We find that cross channel switching in a multi-channel distribution environment occurs due to price differences, convenience and flexibility. The paper illustrates the flaws of drawing inferences with data from a single channel emphasizing the need for a more holistic, multi-channel approach to data analysis in order to capture the full range of consumer behavior.

Research limitations/implications –

This study assumes that self-reported purchases or acquisitions of products accurately reflect actual consumer behavior. The insights derived from this research have significant implications for various marketing strategies, including pricing, product bundling, cross-selling, and promotional activities. However, it's important to note that the reliance on self-reported data may introduce certain limitations, as it may not fully capture the complexity of actual consumer behavior. Despite these limitations, the insights gained from the study will add value for several marketing decisions like pricing, bundling, cross-selling and promotions.

[1] *Department of e-Business, Prin. L. N. Welingkar Institute of Management Development & Research (WeSchool), Mumbai, Maharashtra, India.*

[1]bijith.marakarkandy@welingkar.org

[2]Department of Computer Engineering, Shree L R Tiwari College of Engineering, Mira Road, Mumbai University, India, [2]bnemade@gmail.com

[3]*Department of e-Business, Prin. L. N. Welingkar Institute of Management Development & Research (WeSchool), Mumbai, Maharashtra, India*

[3]sandeep.kelkar@welingkar.org

[4]*Department of e-Business, Prin. L. N. Welingkar Institute of Management Development & Research (WeSchool), Mumbai, Maharashtra, India*

[4]chandrika.vijayakumar@welingkar.org

[5]Department of Information Technology, Vidyavardhinis College of Engineering and Technology

[5]vaishali.shrsath@vcet.edu.in

[6]Computer Engineering Department, SVKMs NMIMS, Mukesh Patel School of Technology Management and Engineering, India.

[6]mahesh.mali@nmims.edu

Originality value– While there is an abundance of literature on the application of Association Rule Mining for Market Basket Analysis, the existing body of work exploring its application within a multi-channel context is notably sparse. This study aims to address this gap in the literature, offering a unique perspective and making a substantial contribution to the existing body of knowledge in this area.

Keywords: Apriori algorithm, Market Basket Analysis, Affinity Analysis, Association Rule Mining

I. INTRODUCTION

Recommendation Systems are being extensively used by ecommerce companies. Relevant and reasonable recommendations based on shopping cart analysis are the current norm in the e-commerce industry. Item-to-item collaborative filtering and User-to-User collaborative filtering have been widely used for the purpose of recommendation. Affinity analysis based on association rules may help to recommend products based on products that are present in the shopping cart and what may be likely to be purchased. Apriori Algorithm was first proposed by Agrawal and Srikant (1994) for discovering association rules in a large database of sales transactions. Purchasing patterns can be mined using transaction data. Products and combination of products that occur together most frequently in shopping carts. Uncover relation between items in a shopping cart. The algorithm looks for co-occurrence of items with high frequency of occurrence. This is an unsupervised machine learning technique. Apriori algorithm is one of the most prevalent Machine learning applications for the retail industry. The algorithm has been put to use by leading supermarket chains and online retailers. The Businesses collect data for varied reasons ranging from providing better customer experiences, regulatory compliance, discover patterns in customer buying behavior.

Machine learning, a powerful subset of artificial intelligence, is a type of data analysis that automates the creation of analytical models. It is based on the notion that systems can learn from data, recognise patterns, and make decisions with little human intervention. This study focuses on the Apriori Algorithm, a classic data mining algorithm. It is used to obtain frequent itemsets for boolean association rules. By analysing customer transaction data, it assists retailers in identifying items that customers frequently purchase together. This paper presents an optimised version of the Apriori algorithm, which is intended to improve the efficiency and effectiveness of multi-channel consumer behaviour analysis (B. Nemade and D. Shah (2022),B. Marakarkandy et al. (2023).)

Deep learning, a more complex subset of machine learning, uses artificial neural networks with multiple layers, which inspired the name 'deep learning'. These deep neural networks resemble the human brain in that they 'learn' from large amounts of data. While a neural network with only one layer can make approximate predictions, adding more hidden layers can help improve accuracy. Deep learning algorithms can be used in the ever-changing retail and e-commerce industries to forecast future consumer behaviour, provide personalised customer experiences, and streamline retail operations by optimising inventory and logistics(Mishra, R. et al. (2023), Alegavi, S. S. et al. (2023)). The primary goal of this study is to provide a more nuanced understanding of multichannel shopping behaviours. This paper aims to gain a better understanding of how online and offline purchases interact by combining machine learning and deep learning methodologies. This comprehensive approach not only helps to improve understanding of consumer behaviour in the retail and e-commerce industries, but it also allows retailers to better cater to their customers' needs and preferences.(Nemade, B. P et al. (2023). Furthermore, it provides useful insights for retailers to optimise their strategies, resulting in a more effective and efficient market presence. This study highlights the potential of machine learning and deep learning as powerful tools for improving our understanding of complex consumer behaviours in today's digital world.

According to Ho et al.,(1998) potential factors like travelling time and cost may increase the size of the shopping cart, however in case of high involvement products we believe that this may not be true.

According to Wickramaratna et al., (2008) Association mining unearths frequently occurring items in a transactional database and limited attention has been given to predict likelihood of purchase based on the frequently co-occurring items. They have proposed a technique to use partial information from the shopping cart to predict what the customer is likely to buy. Mining Frequently occurring items in a shopping cart and finding co-occurring items may not capture correct insights as some items may be purchased online whereas others may be purchased offline.

The purpose of this paper is to better understand frequently co-occurring purchases when certain items are purchased online and other related items are purchased offline. The following research question is formulated for further investigation: Will inclusion of offline purchase data with the online purchase data increase the number of frequently co-occurring purchases?

It is clear from the extant literature that Apriori algorithm has been extensively used for investigating affinity analysis. The two important metrics that are indicators of frequency of occurrence of items in a data set and conditional probability of occurrence of items in a transaction are support and confidence respectively. Accordingly, two hypotheses will be tested in this study

Hypotheses

H1. The metric Support for the offline and online purchases taken together will be more than that for online purchases alone.

H2. The metric Confidence for the offline and online purchases taken together will be more than that for online purchases alone.

The insights gained from the study will add value for several marketing decisions like pricing, bundling, cross-selling and promotions.

II. LITERATURE REVIEW

The original research of Agrawal et al. (1993) marked a significant turning point in the field of data mining, incorporating the era of association rule mining and triggering a wave of advancements in the analysis of customer transaction data. The Apriori and AprioriTid algorithms (Agrawal & Srikant, 1994) were introduced, effectively addressing the challenge of extracting meaningful association rules from large datasets and demonstrating superior performance compared to existing methodologies such as AIS and SETM. The use of Apriori algorithms and FP growth in association rule mining has since spread across various research domains, demonstrating their long-term relevance. Furthermore, studies have highlighted the importance of Market Basket Analysis (MBA) in identifying complex patterns of consumer behavior, suggesting for more sophisticated approaches such as sequence analysis. Despite these advances, there are still gaps, particularly in understanding the complexities of cross-channel purchase behaviors, especially for high-involvement products. This literature review aims to fill these gaps by investigating the use of Apriori algorithms in multi-channel purchase contexts, thereby contributing to a better understanding of consumer dynamics and informing strategic decision-making in retail and marketing.

The seminal work of Agrawal et al., (1993) which generates all significant association between rules in a database of customer transactions was the starting point to the development of association rule mining and subsequent performance improvements of algorithms based on these concepts. Agrawal and Srikant (1994) proposed the Apriori and AprioriTid algorithms for solving the issue of finding significant association rules between items in a large data set. They compared the performance of these algorithms with the existing AIS and SETM algorithms. Empirical results on the comparison shows that the proposed algorithms outperform the existing AIS and SETM algorithms. They also showed that the best features of the proposed algorithms can be combine and a hybrid algorithm can be made, which they called AprioriHybrid.

Extant literature shows that a plethora of research studies have used Apriori algorithm and FP growth or its modified versions in the context of Association Rule mining (Chen and Hu, 2008; Li et al., 2008; Liu and Guan, 2008; Yongmei and Yong, 2009;Raorane et al., 2012; Aguinis et al.,2013; Chang et al., 2016; Kaur and Kang, 2016; Badriyah et al.,2018; Sagin and Avyaz, 2018; Akas and Khan, 2020; Wang et al.; 2020; Unyan, 2021)

Kamakura (2012) found that Market Basket Analysis using affinity analysis is a widely accepted method to understand purchase behavior, used by both traditional and internet retailers. The findings of the study show that purchase – sequence analysis rather than joint-purchase data would offer additional insights and internet retailers could rearrange their website lay outs according to the sequence of purchases.

Anastasia et al.(2018) posits that doing a mere affinity analysis is of comparatively less value and adding an event, occasion or reason for shopping would comprehend the intention of the customer to provide tailored services on every visit. They generated a segment of customer visits and attribute an intention for every visit. Product categories purchased on different visits tend to be different. The authors are of the opinion that a mission-based lay out may have a significant impact on customer satisfaction as compared to a category-based lay out.

Kim et al. (2012) proposed a product network analysis by extending the market basket network. The co-purchased product network is used to analyze the data about products purchased by same customers at the same time and the products purchased by the same customer at different times.

Han and Kim (2017) have identified the reasons for hesitation of online purchase. They found that perceived risk is a major factor of avoiding online purchase in high involvement products. In the current study the selection of Apple products is because, certain offerings are high involvement products and hence the possibility of purchase of certain products from a physical store and others online is a plausibility.

Market basket analysis in multiple store environment was studied by Chen et al. (2005). They concluded that traditional association rule mining fails in identifying significant purchasing pattern in a multiple store environment. As proposed by Julander (1992), data about shoppers at an individual level rather than combination of items the shopper puts in a basket is more useful. A method to improve the shopping experience by improving the efficiency of a recommendation system was proposed by Guo et al. (2017). They used an improved apriori algorithm for improving the efficiency of data mining.

A context-based market basket analysis method in a multiple store environment was proposed by Tang et al. (2008). The proposed approach extracts association rules from transactional data in a multiple-store and multiple-period situation.

According to Musalem et al. (2018), retailers should move away from category management approach to a customer management approach based on their study of shopping behavior to detect cross-category interrelations based on the customers shopping basket.

Serran-Arcos et al. (2022) did a systematic review of consumer affinity literature, they found that although this line of research is growing there were theoretical inconsistencies and contradictory empirical results.

Kotu and Deshpande (2019) are of the view that as there is no target variable to predict , association rule reviews each transaction independently to find patterns. They suggest that Apriori and Frequent Pattern growth algorithms as efficient methods for finding these associations.

Ghoshal et al. (2015) state that out of the variety of methods for providing recommendations collaborative filtering matrix factorization and association-rule -based methods are the most common.

The literature review provides ample evidence that Apriori algorithm can be used as a dependable method for identifying patterns in the transaction data. There is limited research on using Apriori algorithms in case of purchases across multiple channels particularly regarding high involvement products. The current paper is an attempt to fill this gap in the extant literature.

III. THEORETICAL FRAMEWORK

T_i be the transaction data set and I_i be the set of items in the transaction as shown in the equations 1 to 3.

$$T_1= \{I_1, I_2, I_3..........\} \quad (1)$$

$$T_2= \{I_1, I_3, I_6..........\} \quad (2)$$

$$T_n= \{I_2, I_3, I_4..........\} \quad (3)$$

Apriori Algorithm

Apriori Algorithm generates a set of candidates called candidate item set C_i. If a candidate item set meets the minimum support criteria then it is a frequent item set.

Table 1: Shopping Cart Transactions

Transaction T_i	**C_k**
T_1	$\{ I_1,I_2,I_3,I_4\}$
T_2	$\{I_1,I_5\}$
T_3	$\{I_5,I_2,I_3,I_6\}$
T_4	$\{I_1,I_5, I_2,I_3\}$
T_5	$\{I_1,I_5, , I_2, I_6\}$

The rule that can be inferred from the Candidate set C_i is

I_2---->I_3

This rule implies that a number of customers who purchased item I_2 also purchased I_3. These rules can be of immense value to the business and provides opportunity for cross selling

The Apriori algorithm rules between items in a data set are made using three factors support, confidence and lift

Support:

Support of an item or group of items in C_i helps to identify keystone items. It indicates the relative frequency the item set appears in the transactions.

Support is calculated as a ratio of the number of occurrences of the item set and the total number of transactions.

Support (I_2---->I_3) =freq (I_2,I_3)/n

Confidence:

Confidence is the conditional probability of the item sets occurring together in a transaction. Confidence being a probability its range is [0,1].

Confidence(I_2,---->I_3) = ($P(I_3 | I_2)$ = support (I_2, I_3)/$P(I_2)$

IV. 4. METHODOLOGY

4.1 Instrument Development and Data collection

Apple product users were selected as subjects for the study. Apple offers special pricing for students and educators and often bundles two products which are offered under student discount schemes. An instrument was developed to capture the demographic data of the respondents and purchase of the following products airpodmax, airpods, airtag, applepencil, applewatch, homepod, ipad, iphone, leatherlink, leatherwallet, magickeyboard, magicmouse, magictrackpad, magneticcharger, siliconcase, smartfolio, sportloop, ultrabands, usbcable and usbsuperdrive. The pre-test of the instrument was carried out on 10 respondents in face-to-face mode. Based on the feedback certain modifications were made to the item wordings of the questionnaire. This study used a convenience sampling method. Algorithm optimization improves efficiency and performance, while hybridization combines strengths of multiple algorithms to solve complex problems (b. nemade. et al).

4.2 Apriori algorithm optimization

The Apriori algorithm is Optimized using Particle Swarm Optimization (PSO). The survey was designed to capture the complexities of online and offline purchases, with a particular emphasis on Apple's purchase channels. The study used the Apriori algorithm to find frequently co-occurring items purchased online and offline, and calculated essential metrics like support and confidence for both types of purchases. An independent-samples t-test was used to determine any significant differences between the groups. This comprehensive methodology provides a thorough examination of multi-channel shopping behaviours, offering valuable insights into consumer behaviour in the retail and e-commerce sectors. To enhance this analysis, you could consider optimizing the Apriori algorithm using Particle Swarm Optimization (PSO), and the mathematical modelling is described below.

1. **Particle Representation**: Each particle in the swarm (PSO) can represent a candidate solution for the Apriori algorithm. In the context of your problem, a particle can represent a set of items that were purchased together (either online, offline, or both). Mathematically, it is represented using equation 4.

$$X=\{x_1,x_2,...,x_n\} \tag{4}$$

where x is a particle and x_i is an itemset.

2. **Fitness Function**: The fitness of each particle can be evaluated based on the support and confidence of the item sets that it represents. In your case, the support can be the proportion of transactions that include a particular itemset, and the confidence can be the probability that a particular itemset is purchased given that another itemset is purchased. The fitness function is represented using equation 5.

$$f(X)=\alpha*\text{support}(X)+\beta*\text{confidence}(X) \tag{5}$$

where f(X) is the fitness of particle X, support(X) and confidence(X) are the support and confidence of the item sets represented by X, and α and β are weights that determine the importance of support and confidence.

3. **Particle Update**: The particles are updated based on the PSO rules. The velocity update rule influences the exploration and exploitation abilities of the particles, guiding them towards the best-known positions in the search space. The position update rule then applies this velocity to change the position (solution) of the particles. These is represented using equation 6 for velocity update and equation 7 for position update.

Velocity update:

$$v_{ij}(t+1)=w*v_{ij}(t)+c1*rand()*(pbest_{ij}-x_{ij}(t))+c2*rand()*(gbest_j-x_{ij}(t)) \tag{6}$$

Position update:

$$x_{ij}(t+1)=x_{ij}(t)+v_{ij}(t+1) \tag{7}$$

The movement of particles in the search space, which is determined by the velocity and position update rules equation 3 and 4. The particles are connected in the sense that their movements influence each other through the global best position (gbest), which is shared among all particles. This approach potentially enhances methodology by providing a more optimized set of frequent item sets, thereby offering deeper insights into consumer behaviour in the retail and e-commerce sectors.

4. **Optimization Goal (Eq. 5)**: The goal of connecting PSO with the Apriori algorithm in this way would be to find the most frequent itemset(s), i.e., the global best solution in the context of PSO. This is represented using equation 8.

$$\text{maximize } f(X) \tag{8}$$

subject to the constraints of the Apriori algorithm (e.g., minimum support and confidence thresholds).

This way, the optimized Apriori algorithm is used to find the most frequently co-occurring items that were purchased online and offline. This provides valuable insights into multi-channel shopping behaviours, which is useful to improve the retail and e-commerce sectors.

4.3 Testing and analysis

4.3 Testing and Analysis

The study's objective was to investigate potential inaccuracies in product affinity inferences when analyzing transaction data from a single channel. Transaction data, which included online and offline purchases as well as gifted products, were collected through a survey questionnaire.

The collected data was divided into two distinct groups: one consisting solely of online purchases, and the other comprising purchases made both online and offline, along with gifted items. These groups were assumed to be independent, reflecting the diverse purchasing behaviors among the respondents, some of whom were exclusive online purchasers, while others made purchases both online and in physical stores. The Apriori algorithm, optimized using Particle Swarm Optimization (PSO) as detailed in the methodology, was employed to calculate the support and confidence values for both groups. These metrics provide insights into the frequency of item sets and the conditional probability of purchasing certain item sets given the purchase of others.

An independent-samples t-test was conducted to compare the means of the support and confidence metrics across the two groups. This statistical test was chosen to determine if there were significant differences in the purchasing patterns between the two groups, thereby testing the initial assumption of independence. The results of this comprehensive analysis offer a deeper understanding of multi-channel shopping behaviors, providing valuable insights for enhancing retail and e-commerce strategies. The optimization goal, as described in the methodology, was to maximize the fitness function while adhering to the constraints of the Apriori algorithm, such as minimum support and confidence thresholds. This approach aimed to uncover the most frequently co-occurring items purchased online and offline, thereby offering valuable insights for enhancing retail and e-commerce strategies. The detailed results of this analysis are not included here and should be referred to in the original document. The results pertaining to confidence after performing independent-samples t-test is as shown in Table 2.

Table 2: Confidence analysis of affinity between products purchased Online and Online + Offline

	Confidence online	*Confidence online + offline*	*T stat*	*p-value*
Mean	0.53	0.31	18.64	4.00915E-73
Variance	0.13	0.054		
Observations	1015	1522		
df	2535			

The mean of the metric confidence for the first group was 0.53 with a standard deviation of 0.37 and for the second group the mean for the metric confidence was 0.31with a standard deviation of 0.23. The t-statistic was 18.64, with df= 2535 (*p* value < 0.001), indicating that the null hypothesis may be rejected.

The results pertaining to support after performing an independent-samples t-test is as shown in Table 3.

Table 3: Support analysis between affinity between products purchased Online and Online + Offline

	Support Online	*Support Online+ offline*	t Stat	*p-value*
Mean	0.02511	0.02589	-2.72	0.003
Variance	2.29466E-05	6.89121E-05		
Observations	1015	1522		
df	2535			

The mean of the metric support for the first group was 0.0251 with a standard deviation of 0.00479 and for the second group the mean for the metric support was 0.02589 with a standard deviation of 0.008301. The t-statistic was skewed towards left representing value of -2.72, with df= 2535 with corresponding *p* value=0.003 which is less than the required significance level (0.05) indicating that the null hypothesis may be rejected.

The results indicate that there is significant difference between both the groups. Specifically the second group has a higher mean than the first group for both the metrics support and confidence.

Table2: Hypotheses tested using independent-samples t-test

Hypothesis		
H1	The metric Support for the offline and online purchases taken together will be more than that for online purchases alone.	Supported
H2	The metric Confidence for the offline and online purchases taken together will be more than that for online purchases alone.	Supported

V. CONCLUSIONS

The study has empirical evidence to show that the metrics support and confidence increases when both online and offline purchases are clubbed. In the current study empirical evidence shows that association rule mining performed only on purchases online would lead to misleading conclusions. Including online purchase, with offline purchases and purchases by others on behalf of the one who requires the product may be included while performing Association Rule Mining. The shopping cart data about purchases is easier to obtain as compared to survey data. However, capturing the data from multi-channel purchases is a challenging task and businesses should step up efforts to seamlessly capture transaction data across multiple channels.

5.1 Contributions to theory

The product purchase via online mode has been seeing a drastic upswing in the recent past particularly during and after the pandemic, however the physical store has its own reasons for existence. Lawrence et al. (2001) argue that there is a need to understand the customer buying pattern in entirety and not just intimate them about what they may most likely buy based on past data mining results. According to Sarwar et al. (2000) recommender systems discover frequently purchased items by customers and for a particular customer the recommendation would be based on an association rule which relies on antecedents and consequents. The current study contributes to the extant literature by empirically demonstrating that transaction data from multi channels is essential for an accurate recommendation.

5.2 Contributions to practice

Forward looking businesses try to gain competitive advantage by getting insights from both internal and external data. Obtaining insights about what people buy, how they buy, when they buy, where they buy, at what price they buy are important, to take it ahead understanding what they buy along with what is helpful for taking strategic decisions regarding inventory, pricing, bundling, combo offers and promotions. Loshin (2012) is of the view that market basket analysis is the analysis of any collection of items to identify affinities and may be useful for up-selling, cross-selling and bundling opportunities. Customers criss-cross across multiple channels while making a purchase. Businesses find it difficult to seamlessly track the journey. This may result in analysis of data from channels which are easy to acquire. Businesses are tracking offline purchases by various means which includes credit and debit card transactions, e-receipts that capture customers email address, the data generated from these tracking mechanisms can be combined with online purchase data to generate real transaction data. The transaction data which is a combination obtained from multi channels may then be used for discovering patterns in customer buying behavior. Store visits are harder to track than website visits. Businesses should monitor customer interactions across multiple channels. Even though the business which may have a robust customer management software, tracking becomes difficult as the products meant for a consumer may have been purchased by someone else. We recommend that a periodic assessment by a survey instrument may be able to capture the missing links.

VI. LIMITATIONS AND SCOPE FOR FUTURE WORK

An inherent limitation of this work is capturing multi-channel purchase information was challenging and self-reported purchase information was captured using a survey instrument. In the current study the sample size is small and lack of sampling frame resulted in difficulties in generating random samples. The sample was regional and convenience sampling method was employed. Further, research may be directed to identify and recommend products based on customer purchase patterns across multiple channels. Future research may be directed. The selection of algorithms for finding these associations or co-occurrences of items in a transaction will be of immense value to the business.

REFERENCES

[1] Agrawal, R., Imieliński, T., & Swami, A. (1993, June). Mining association rules between sets of items in large databases. In *Proceedings of the 1993 ACM SIGMOD international conference on Management of data* (pp. 207-216).

[2] Agrawal, R., & Srikant, R. (1994, September). Fast algorithms for mining association rules. In *Proc. 20th int. conf. very large data bases, VLDB* (Vol. 1215, pp. 487-499).

[3] Aguinis, H., Forcum, L. E., & Joo, H. (2013). Using market basket analysis in management research. *Journal of Management*, *39*(7), 1799-1824.

[4] Akas, M. F., Zaman, A. G. M., & Khan, A. (2020, January). Combined item sets generation using modified apriori algorithm. In *Proceedings of the International Conference on Computing Advancements* (pp. 1-3).

[5] Badriyah, T., Azvy, S., Yuwono, W., &Syarif, I. (2018, March). Recommendation system for property search using content based filtering method. In *2018 International conference on information and communications technology (ICOIACT)* (pp. 25-29). IEEE.

[6] Chang, H. Y., Lin, J. C., Cheng, M. L., & Huang, S. C. (2016, July). A novel incremental data mining algorithm based on fp-growth for big data. In *2016 International Conference on Networking and Network Applications (NaNA)* (pp. 375-378). IEEE.

[7] Chen, Y. L., Tang, K., Shen, R. J., & Hu, Y. H. (2005). Market basket analysis in a multiple store environment. *Decision support systems*, *40*(2), 339-354.

[8] Ghoshal, A., Menon, S., & Sarkar, S. (2015). Recommendations using information from multiple association rules: A probabilistic approach. *Information Systems Research*, *26*(3), 532-551.

[9] Guo, Y., Wang, M., & Li, X. (2017). Application of an improved Apriori algorithm in a mobile e-commerce recommendation system. *Industrial Management & Data Systems*, *117*(2), 287-303.

[10] Han, J., Pei, J., & Yin, Y. (2000). Mining frequent patterns without candidate generation. *ACM sigmod record*, *29*(2), 1-12.

[11] Han, M. C., & Kim, Y. (2017). Why consumers hesitate to shop online: Perceived risk and product involvement on Taobao. com. *Journal of promotion management*, *23*(1), 24-44.

[12] Julander, C. R. (1992). Basket analysis: a new way of analysing scanner data. *International Journal of Retail & Distribution Management*, *20*(7).

[13] Kamakura, W. A. (2012). Sequential market basket analysis. *Marketing Letters*, *23*, 505-516.

[14] Kaur, M., & Kang, S. (2016). Market Basket Analysis: Identify the changing trends of market data using association rule mining. *Procedia computer science*, *85*, 78-85.

[15] Kim, H. K., Kim, J. K., & Chen, Q. Y. (2012). A product network analysis for extending the market basket analysis. *Expert Systems with Applications*, *39*(8), 7403-7410.

[16] Kotu, V., & Deshpande, B. (2018). *Data science: concepts and practice*. Morgan Kaufmann.

[17] Lawrence, R. D., Almasi, G. S., Kotlyar, V., Viveros, M., & Duri, S. S. (2001). *Personalization of supermarket product recommendations* (pp. 11-32). Springer US.

[18] Li, H., Wang, Y., Zhang, D., Zhang, M., & Chang, E. Y. (2008, October). Pfp: parallel fp-growth for query recommendation. In *Proceedings of the 2008 ACM conference on Recommender systems* (pp. 107-114).

[19] Lin, W. B. (2013). Factors affecting high-involvement product purchasing behavior. *Quality & quantity*, *47*, 3113-3133.

[20] Liu, Y., & Guan, Y. (2008, November). Fp-growth algorithm for application in research of market basket analysis. In *2008 IEEE International Conference on Computational Cybernetics* (pp. 269-272). IEEE.

[21] Mishra, R., Nemade, B., Shah, K., & Jangid, P. (2023). Improved Inductive Learning Approach-5 (IILA-5) in Distributed System. International Journal of Intelligent Systems and Applications in Engineering, 11(10s), 942-953.

[22] Loshin, D. (2012). *Business intelligence: the savvy manager's guide*. Newnes.
[23] Musalem, A., Aburto, L., & Bosch, M. (2018). Market basket analysis insights to support category management. *European Journal of Marketing*, *52*(7/8), 1550-1573.
[24] Raorane, A. A., Kulkarni, R. V., &Jitkar, B. D. (2012). Association rule–extracting knowledge using market basket analysis. *Research Journal of Recent Sciences ISSN*, *2277*, 2502.
[25] Sagin, A. N., & Ayvaz, B. (2018). Determination of association rules with market basket analysis: application in the retail sector. *Southeast Europe Journal of Soft Computing*, *7*(1).
[26] Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2000, October). Analysis of recommendation algorithms for e-commerce. In *Proceedings of the 2nd ACM Conference on Electronic Commerce* (pp. 158-167).
[27] Serrano-Arcos, M. M., Sánchez-Fernández, R., Pérez-Mesa, J. C., & Riefler, P. (2022). A review of consumer affinity research: recent advances and future directions. *International Marketing Review*, *39*(5), 1252-1282.
[28] Alegavi, S. S., Nemade, B., Bharadi, V., Gupta, S., Singh, V., & Belge, A. Revolutionizing Healthcare through Health Monitoring Applications with Wearable Biomedical Devices.
[29] Nemade, B. P., Shah, K., Marakarkandy, B., Shah, K., Surve, B. C., & Nagra, R. K. (2024). An Efficient IoT-Based Automated Food Waste Management System with Food Spoilage Detection. *International Journal of Intelligent Systems and Applications in Engineering*, *12*(5s), 434-449.
[30] Tang, K., Chen, Y. L., & Hu, H. W. (2008). Context-based market basket analysis in a multiple-store environment. *Decision Support Systems*, *45*(1), 150-163.
[31] Ünvan, Y. A. (2021). Market basket analysis with association rules. *Communications in Statistics-Theory and Methods*, *50*(7), 1615-1628.
[32] Wang, F., Wen, Y., Guo, T., Liu, J., & Cao, B. (2020). Collaborative filtering and association rule mining-based market basket recommendation on spark. *Concurrency and Computation: Practice and Experience*, *32*(7), e5565.
[33] Shirsath, V., Kaul, V., Kumar, R. S., & Nemade, B. (2023). INTELLIGENT TRAFFIC MANAGEMENT FOR VEHICULAR NETWORKS USING MACHINE LEARNING. ICTACT Journal on Communication Technology, 14(3).
[34] Wickramaratna, K., Kubat, M., & Premaratne, K. (2008). Predicting missing items in shopping carts. *IEEE transactions on knowledge and data engineering*, *21*(7), 985-998.
[35] B. Nemade and D. Shah (2022), An IoT based efficient Air pollution prediction system using DLMNN classifier, Physics and Chemistry of the Earth, Parts A/B/C, vol. 128, 2022.Available: https://doi.org/10.1016/j.pce.2022.103242
[36] Marakarkandy, B.Nemade, B., Bharadi, V., Alegavi, S. S. (2023). A Comprehensive Review: SMOTE-Based Oversampling Methods for Imbalanced Classification Techniques, Evaluation, and Result Comparisons. International Journal of Intelligent Systems and Applications in Engineering, 11(9s), 790-803.