Journal of Information & Optimization Sciences
ISSN 0252-2667 (Print), ISSN 2169-0103 (Online)
Vol. 45 (2024), No. 7, pp. 1993–2005
DOI : 10.47974/JIOS-1751

Customer behaviour in a retail store using IoT & machine learning in a cloud environment

Krishna Kumar Singh *
Mugdha Shailendra Kulkarni †
Symbiosis Centre for Information Technology
Symbiosis International (Deemed University)
Pune 411057
Maharashtra
India

Abstract

This paper investigates the integration of machine learning and Internet of Things (IoT) technologies inside a cloud environment in retail stores. Although IoT, ML, and big data have been studied individually, their combined usefulness and efficiency research is still lacking. To fill this gap, our research suggests a strategy that uses ML-machine learning algorithms and IoT devices to improve consumer experiences and streamline operations in retail settings. This paper collects data through Internet of Things (IoT) devices with sophisticated sensing features, such as fuzzy logic, to track and examine consumer interactions, movements, and behaviours in retail settings. Hence, Product movements can be tracked via RFID-tagged product tracking, providing information about consumer preferences and the popularity of particular products. The cloud securely stores and processes the data produced by these Internet of Things devices, using big data techniques to glean insights from enormous datasets. The most frequently used ML technique is association rule mining, which extracts significant patterns and correlations from the collected data. Retailers can use this to give personalised suggestions and enhance product placement and assortment by gaining better insights into customer preferences, purchasing patterns, and product affinities. Using case studies and real-world examples, we illustrate the effectiveness and advantages of combining IoT and machine learning technology in retail environments. This research advances the field of retail analytics by offering a thorough framework for utilising emerging technologies to promote corporate growth, improve operational efficiency, and provide better consumer experiences.

Subject Classification: *94-XX.*

Keywords: *Cloud computing, Internet of Things (IoT), Sensors, Cybersecurity, Associate rule mining.*

* *E-mail:* **krishnakumar@scit.edu (Corresponding Author)**
† *E-mail:* **mugdha@scit.edu**

Introduction

The Retail market is not only about buying goods and products directly from the stores but also personalising the customer's experience and products, which will enhance the sales and gross profit of the stores. The objective is to enhance customer experience using big data analytics, the Internet of Things (IoT) and Cloud Computing. Most research has been done on data and IoT in the retail industry, which has focused on obtaining customer insights to give up heavy discounts, maintaining loyal customers and keeping up with the trend. With the help of data analysis, it will create more marketing strategies to generate more demand but need not convert it into more sales. The paper aims to show the implementation of a fuzzy logic model using the Internet of Things (IoT) to detect the behavioural intentions of customers visiting the shop. This will enable the managers to know the customers' intention to visit the stores and help them keep the items that have a higher tendency to be sold. Also, Associate rule mining, employed in market basket analysis to understand the association between the various items, has been taken into consideration, which can help the retailers understand the purchasing probability of a customer or the customer's behaviour in buying a particular product. The paper has been divided into four major sections. The first section contains reviews from several journals discussing integrating the Internet of Things, Big data, and cloud computing to improve the retail sector. The second part describes the methodology, which is the analytical part. The Third part refers to the implementation of the idea. The last part is the conclusion, which also includes the limitations of the idea and how it can be improved so that the concept will be sustained in the market.

Literature Review

The literature review is divided into Technology used in various services, adoption in the service sector, retail and technology, and how retailers use technology like big data to evaluate real-time data.

Internet of Things (IoT) and big data technologies are both emerging as technological breakthroughs on the rise, and they have been used in various ways for various purposes because of their innumerable benefits. The countries already using these technologies have overcome the problems of fatal car accidents, traffic rise, operational waste, and passenger decision-making using these emerging technologies [1, 2, 3]. Numerous factors affect consumer preferences, determining how a client will behave when using products or services. Customer behaviour has undergone a paradigm shift related to digital technology and is supporting businesses in making quick changes to remain competitive. Adopting future technology will benefit everyone involved in the supply chain, not just customers. Products can self-diagnose problems and be remotely monitored thanks to IoT or connected gadgets. This will make it easier to analyse device data to forecast problems, take proactive corrective action, provide better services, and increase productivity. It has become crucial for the company to consider each customer's preferences [2]. In the literature paper, the concept of Smart Label is introduced and verified using a Web Application module. They have improved the equipment's Overall Equipment Effectiveness (OEE) from 0.52 to 0.76. Intelligent labels with RFID tags and QR codes are placed on the products to track sales. Customers can also get information on the product's expiration date and usage by scanning the intelligent labels. This will decrease the number of salespeople in the store and increase the area available for the customers to shop. RFID is used by web applications with two login modules (one for the store and the other for the manufacturer) developed to display the delivery status, sales statistics, and other information that allows the manufacturers to keep track of the product's progress. This benefits both the retailer and the manufacturer and will provide higher sales. This paper also proposes a recognisable and data-driven smart factory that collects data from retail shops. The products communicate their state via IoT (Internet of Things), analyse the massive data collected, and provide essential facts to the manufacturing industry and the client [3]. The retail environment is rapidly altering worldwide due to the expansion of the Internet and network. Retailers are now creating innovative ways for customers to communicate with them and new customer channels. The following paper discusses how the Internet of Things (IoT) improves innovation in the retail process and how it generates and uses sensitive data. When dealing with customer satisfaction, several ways can provide solutions to the IoT, like Augmented Reality, Smart Customer Relationship Management, Real-time In-store promotions, Automated Checkout and Customer behavioural tracking. Shopping has become un-interactive in the current retail market, and the ordinary store does not provide sufficient tools and applications to enhance the shopping experience. Also, retail is now focused on creating new things and tools to improve services. The Internet of Things (IoT) needs to be tested in a few areas, which include the ability of the IoT devices to predict and meet customer needs by improving the shopping experience, the ability of the IoT as an analytical instrument to analyse the customer market and the unpredictability of the technology's acceptance [4] The authors of this literary work are seasoned experts who have studied data for the retail industry. Based on their market understanding, this article theorised fresh ideas and potential applications of big data technology in retail. This is in addition to e-commerce businesses tracking user browsing behaviours. By identifying the characteristics of products in a category, estimating the demand for attributes from sales of existing products, and then judging the market for a potential new product based on the demand for its constituent characteristics, it is possible to estimate the potential sales of new products. Retailers should assess whether the additional revenue from one effort is sufficient to justify the expense before investing in product distribution and warehousing to reduce the time it takes to fulfil client orders. The authors claim that big data has enormous promise for the retail industry. There are several successful pilot initiatives, but the fundamental difficulty has been scaling them up. [5]

An IoT platform can only access client behaviour and assist with innovative improvements in the fashion retail sector. To give the salesperson a better understanding of the customer's intention, the fuzzy logic concept is used to identify the product qualities and the purchasing intention. Using big data analytics, it is possible to gain a deeper understanding of a customer's intent to buy a certain item and their shopping habits. [6] . RFID tagging is discussed in the paper, along with how it improves inventory management, makes inventory more visible, and enables the sales associate to assist customers, leading to higher customer satisfaction and additional sales chances. Passive RFID tags provide real-time item location information. RFID tags close the generated supply-demand gap. RFID chips are incorporated into items with anti-cloning characteristics that stop someone from copying the RFID chip or its data. Networks of cameras are used to monitor employee and customer behaviour. Based on sentiment analysis, the acquired data helps to customise the consumer experience. The history of a customer's purchases can also be tracked when they have a smart card. Another factor that IoT can handle is the cost of the goods. To embrace dynamic pricing and predict future demand for new products more accurately, merchants can employ machine learning algorithms and demand forecasting based on customer shopping data and real-time inventory visibility. The restrictions include the size of the retail stores and branches, affecting the total ownership cost for distributing and consuming data. IoT device security and privacy issues include cryptographic attacks, spoofing, and authentication.[7] [8] Additionally, the Cloud can assist in lightening the financial strain of infrastructure on businesses. Retailers are no longer required to put up physical stores before waiting for customers to arrive. They only need to provide the product after responding to the customer's wants online. Using this online retail application, customers can use any portable device to shop [9]. The shopkeeper can continuously monitor the data while all of the shop's data can be deployed on the cloud. However, the fundamental issues—which have not been addressed in the research paper—are security, privacy, and a lack of standards that could discourage buyers from completing online purchasing. As a result, the topic of integrity verification of huge records in cloud computing is a thriving area that is quickly drawing increasing research interest. Despite this, many studies are now being conducted in this area. The topics of cloud and huge records are expanding quickly. Therefore, even if current research has already achieved several excellent aims, we are confident that integrity verification procedures will continue to develop in tandem with the advancement of cloud and big data programs to meet new requirements and address fresh security concerns. [10] In this study, the authors presented an "intelligent shopping platform architecture" advanced in the European project "Social Local Mobile iNdoor shopping experience", or, in short, "SoLoMoN". Here, the authors discussed a location tracking service that will track the stores' locations and delimitation areas from a map based on the available picture of a shopping centre. They also presented an IoT implementation for an intelligent parking solution. The journal authors have discussed technologies relevant to indoor localisation systems before describing the architecture of "SoLoMoN" and how the proposed enhancement fits into the existing setup. After that, they presented prototype UI designs for location tracking and intelligent parking services and showed us the results of preliminary validation tests of the services. We have learned about new advancements in Bluetooth technology, such as energy-efficient BLE (Bluetooth Low Energy) with simpler receiver architecture and ongoing research on ultrasonic, infrared, and ultra-wideband (UWB) technologies. Application of these technologies include museum guides, paramedics, personal task reminders, asset tracking, search and rescue, advertising, and location-enabled social networks. "Solomon" has a three-tier architecture with Android and web applications as the front-end, a Java TCP server connected with MySQL through JDBC as the back-end, and a series of IoT devices for sensing and tracking information. In conclusion, the authors suggested possibly bypassing the BLE technology problems using image segmentation and multiple noise filtering techniques. [11] The research compares and contrasts two IOT-based smart basket systems that use various technologies, namely cloud and edge computing. Both of the research's suggested solutions use Wi-Fi for IoT, a webcam, and a Raspberry Pi. The objects in the smart basket are detected using software like Python, TFLite, OpenCV, and Google Cloud Vision API. The RSSI value and time performance of two separate systems are used to compare their performance. Instead of employing RFID and barcode scanning to identify the object, systems are educated using artificial intelligence. Whereas System B is based on cloud computing and processes identified objects in the cloud server, System A is based on edge computing and processes detected objects in the shopping basket device. A webcam and a power source are connected to the Raspberry Pie 4 device. For System B, TFLite is installed, and for System A, Google Cloud Vision API is deployed. Systems' performance is evaluated in LOS (Line of Sight) and NLOS (Non-Line of Sight). System A's time performance under LOS conditions exhibits consistency in time rate, averaging 1.74 seconds. System B's time performance in LOS conditions exhibits significant time rate fluctuations, with an average delay of 10.46 seconds. The average time rate for System A is 1.75 seconds, while System B's is 5.36 seconds when it comes to time performance under the NLOS scenario. According to the analysis results, edge computing is more reliable than cloud computing. [12] More people are increasingly exchanging information over the Internet, utilising devices like PCs, laptops, tablets, mobile phones, and comparable information transmitting and receiving devices thanks to the development of computing and the Internet. This evolved into a conversational tactic centred on the host. The Internet of Things (IoT), which connects physical items to the community and enables record-changing, is the next technological development in computation, networking, and verbal communication. The things exchange information verbally using a variety of protocols securely and reliably, as well as information learned from their environment. Given that the devices, their cardinality and alignment, the data type, the data rate, and the specifics of the area of the IoT programs vary, it may be necessary to define architectures that include the devices, verbal exchange mediums, storage, and evaluation capabilities, as well as the consumers of the derived output. The Internet of Things (IoT) is a collection of different structures with networking, processing, sensing, and communication capabilities. Because of improvements in sensor devices, IoT is becoming more widespread with software across industries and domain names. [13]

Methodology

This study uses a mixed-method approach to examine customer behaviour in retail outlets using IoT and machine learning in a cloud environment. It combines case study analysis, Associate Rule Mining (ARM), and literature review. Using IoT devices with sensors and RFID technology, a variety of retail outlets are chosen for the case study phase to collect qualitative and quantitative data. A thorough literature search in parallel finds pertinent papers on retail analytics, machine learning, and IoT that shed light on customer behaviour analysis and technical applications. The gathered data is then analysed using ARM, which finds significant correlations between consumer behaviour and product preferences. Retailers can more easily adopt this integrated methodology and make strategic decisions by thoroughly understanding customer behaviour in retail settings.

Association Rule Mining

Association rule is one of the techniques in data mining that can be used to carry out the relationship between the items that the customer has touched and felt. The concept has been determined from the terminology of market basket analysis, which is about searching for the relationships of various products during purchase production.

Although many algorithms have been created concerning association rules, the Apriori algorithm is a well-known and frequently used one. This algorithm's fundamental idea is to create a frequent item set. With a single item, they create iteratively frequent item sets of two, three, and so forth, up to frequent item sets of all sizes.

Apriori Algorithm

The Apriori algorithm operates on databases containing transactions and constructs association rules using frequent items. It establishes how strongly or weakly two items are associated using these association rules. This algorithm quickly calculates the itemset associations using a breadth-first search and a hash tree. Finding the frequent item sets from a vast dataset involves an iterative procedure.

Results and Discussion

When the customer enters the retail shop, they will go through a metal detector sensor, which will be attached to the doors, and all the customer's details will be tracked and stored in the cloud platform. The customer then searches for the item in the kiosk machine, which can be seen when one enters the store. The sort of clothes, such as whether they are for the summer or the winter, the fabric, such as cotton or silk, the price range they are looking for, etc., can all be included in the list of questions. The algorithm will be developed in such a way that it can analyse the customer's responses, guide the customer to a specific section where it will be simpler for the customer to identify the products, and propose related products that are also available. The IoT Embedded Vision Solution devices will be attached to the areas or across the shelves and be able to track customers in the stores, determine how confidently customers reach out for the products, how quickly customers can find the products, and also determine how customers are drawn to the discounted products. The raw data coming from the sensors henceforth can be collected and uploaded to the cloud for further processing. Another method that can be used to detect the behaviour of the customer behaviour on the items that are present in the retail store is Associate Rule Mining, which the algorithm that can be used is the Apriori algorithm that can be used to understand and analyse the behaviour of the customers while shopping at the retail store. The following data that can occur in the central retail is shown in Table 1, considering minimum support as two and the minimum confidence as 50%.

Table 1

Minimum Support and Confidence

Transaction ID	T1	T2	T3	T4	T5	T6	T7	T8	T9
Itemset	M, N	N, P	N, O	M, N, P	M, O	N, O	M, O	M, N, O, Q	M, N, O

In the first step, a table is created. That will have the support count, i.e., the frequency of each item in the dataset, called the candidate set. The transaction IDs and associated itemsets offer valuable information about the purchasing patterns seen in various transactions. For instance, the items M and N in transaction T1 show that items M and N were bought together. In the same way, it appears from the items M, N, and P in transaction T4 that they were purchased jointly. Retailers can evaluate inventory management strategies, product bundling, and cross-selling in retail environments by using this data to identify common purchasing patterns and relationships between different products.

Table 2

Frequency Count

Itemset	M	N	O	P	Q
Support_Count	6	7	5	2	1

Subsequently, the item with the highest number of supports will be eliminated, resulting in the frequency itemset Table 2. Hence, only the item with a minimum support count of two will be considered throughout the analysis—the others will be disregarded. Subsets will be created from the pairs of items in the following phase. Thus, the table 3 will be made as it is displayed below.

Table 3

Subset

Itemset	{M, N}	{M, O}	{M, P}	{N, O}	{N, P}	{O, Q}
Support_Count	4	4	1	4	2	0

Interpreting the support counts for each item set provides a meaningful approach to determining how often specific item combinations occur in the dataset. For instance, itemsets {M, N} and {M, O} have a support count of 4, indicating that these combinations are present in 4 transactions across the dataset. However, itemset {O, Q} has no support count, meaning this combination is absent in any transactions. Retailers can use these support counts to identify popular item combinations and learn more about their customers' preferences. Using this information, they can make decisions about product placement, inventory management, and physical marketing strategies.

As the support has been considered as a minimum of two, the support count with a value of less than two will be ignored, and the rest will be taken ahead. A new dataset will be formed to take the following analysis for associate rule mining. Next, a new table will be created for implementing Associate Rule mining with all the possible rules from the combinations. Confidence is calculated using some formulas like sup(A^B)/A for all the rules. After the confidence value has been calculated, the rules with values less than the threshold values will be rejected, and the rest having values more than the threshold values will be considered. (Ref Table 4)

Table 4

Rule-support and Confidence

Rules	Support	Confidence
M ^N → O	2	Sup {(M ^N) ^O}/sup (M ^N) = 2/4=0.5=50%
N^O → M	2	Sup{(M^N) ^M}/sup (N ^O) = 2/4=0.5=50%
M^O → N	2	Sup {(M ^O) ^N}/sup (M ^O)= 2/4=0.5=50%
O→ M ^N	2	Sup {(O^ (M ^N)}/sup(O)= 2/5=0.4=40%
M→ N^O	2	Sup {(M^ (N ^O)}/sup(M)= 2/6=0.33=33.33%
N→ N^O	2	Sup {(N^ (N ^O)}/sup(N)= 2/7=0.28=28%

The first three rules will be considered since the minimum confidence has been decided to be 50%. So, based on the implementation of the associate rule mining, the seller can take various strategies to increase the sales of the product, like increasing the stock of the item or product when there is a tendency for the customer to buy that product, making sure that the item is not deficit. We are giving discounts on the items most unlikely to be taken along with the products most likely to be purchased by the customer. When the Association Rule Mining (ARM) results are interpreted, meaningful information about consumer behaviour and product preferences in retail contexts is revealed. The developed rules indicate strong correlations between specific client behaviours and product selections. For instance, the “M ^ N → O” rule indicates that customers are likely to engage with item O if they interact with both M and N. Similarly, the “N^O → M” rule implies that customers who interact with items N and O are likely to interact with item M. Using this data about customer preferences and purchasing patterns, retailers may enhance the overall shopping experience for customers in their stores, maximise product placement, and target their marketing campaigns.

A detailed analysis of the provided tables containing transaction IDs and support counts provides valuable new insights into consumer behaviour and product correlations in retail environments. Patterns indicating shared customer preferences and purchase behaviours include association rules and frequent items. By utilising these insights, retailers may enhance customer satisfaction and promote business growth by strategically placing products, implementing marketing strategies, and managing inventories. By using these findings better to tailor their products to the needs of their customers, retailers can boost sales and foster customer loyalty in highly competitive retail situations.

Implications

The study's theoretical implications stem from its contribution to the changing field of retail analytics. It broadens our theoretical understanding of how IoT, machine learning, and cloud computing converge to study consumer behaviour in retail environments. Moreover, the validation of Association Rule Mining (ARM) contributes to the theoretical underpinnings of data mining approaches in retail analytics by providing theoretical justification for its efficacy in detecting significant connections between consumer actions and product preferences.

Practically speaking, this study gives shops real advantages for enhancing company results. Personalised consumer experiences, improved processes, and data-driven decision-making are made possible by the insights gained from IoT and machine learning. Retailers can improve operational effectiveness, maximise marketing initiatives, and obtain a competitive edge by implementing the suggested framework.

Conclusion and Future Aspects

Applying the IoT framework models that may assist them in anticipating the behavioural intention of the clients who are approaching their shop to buy things, the retail business can utilise the fuzzy logical model by using the models from the study above, as shown. Additionally, the model used above attempted to describe how customers moved around the store by taking pictures with sensors or IoT Embedded Visual solutions, which are affordable and open up new opportunities for company planning. Along with the sensors, once the data is collected, unsupervised machine learning like the associate rule mining and apriori algorithm has been implemented to understand the product purchase and the necessary steps the retailers can take to reduce the product deficit, keeping some extra stocks in store. Additionally, data comparisons on how clients are drawn by the marketing information provided by the merchants have been produced with the aid of these IoT sensors. The primary challenge in using IoT devices is maintaining security and privacy. Therefore, while the kiosk machine is collecting customer-related data, it could open the door for cyberattacks in this situation. High-level machine learning and deep learning algorithms can be used to improve security since they quickly adapt to models with changing parameters and let IoT devices change scenarios in real-time.

References

[1] W. N. Hussein, L. M. Kamarudin, H. N. Hussain, A. Zakaria, R. B. Ahmed, and N. A. Zahri, "The prospect of Internet of Things and big data analytics in transportation system," *Journal of Physics: Conference Series*, vol. 1018, no. 1 (2018).

[2] A. K. Singh and P. Thirumoorthi, "The impact of digital disruption technologies on customer preferences: The case of retail commerce," *International Journal of Recent Technology and Engineering*, vol. 8, no. 3, pp. 1255-1261 (2019).

[3] S. Kayalvizhi, D. A. Sughi, and D. R, "Automation in manufacturing and retail industry using smart labels," *International Journal of Advanced Science and Technology*, vol. 28, no. 3, pp. 450-459 (2019).

[4] B. t. Bok, "Innovating the retail industry; an IoT approach," *International Journal of IoT Approaches*, no. 4 (2016).

[5] M. Fisher and A. Raman, "Using data and big data in retailing," *Production and Operations Management*, vol. 27, no. 5, pp. 1665-1669 (2018).

[6] C. C. On, H. C. W. Lau, and Y. Fan, "IoT data acquisition in fashion retail application: Fuzzy logic approach," in Proceedings of the International Conference on Artificial Intelligence and Big Data, pp. 52-56 (2018).

[7] F. Caro and R. Sadr, "The Internet of Things (IoT) in retail: Bridging supply and demand," *Business Horizons*, vol. 62, no. 1, pp. 47-54 (2019).

[8] E. T. Bradlow, M. Gangwar, P. Kopalle, and S. Voleti, "The role of big data and predictive analytics in retailing," *Journal of Retailing*, vol. 93, no. 1, pp. 79-95 (2017).

[9] A. N. Kulkarni and S. G. Lohiya, "Cloud computing expose for online retail management system," *International Journal of Emerging Technologies in Engineering Research*, vol. 5, no. 4 (2017).

[10] C. Liu, C. Yang, X. Zhang, and J. Chen, "External integrity verification for outsourced big data in cloud and IoT: A big picture," *Future Generation Computer Systems*, vol. 49, no. 10, pp. 58-67 (2015).

[11] G. Suciu, C. Balanean, A. Pasat, C. Istrate, H. Ijaz, and R. Matei, "A new concept of smart shopping platform based on IoT solutions," in Proceedings of the International Conference on Electronics, Computers and Artificial Intelligence, pp. 1-4 (2020).

[12] N. Waranugraha and M. Suryanegara, "The development of IoT smart basket: Performance comparison between edge computing and cloud computing system," in Proceedings of the 3rd International Conference on Computer and Informatics Engineering, pp. 410-414 (2020).

[13] G. Mohindru, K. Mondal, and H. Banka, "Internet of Things and data analytics: A current review," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 10, no. 3 (2020).

Received April, 2024