Management Dynamics

Volume 23 | Number 2 Article 3

June 2023

Application of Machine Learning algorithm for creating sustainable omni-channel retail ecosystem

Somedip Karmakar
Staff Data Scientist, International Technology,Walmart (United States), somedip@yahoo.co.in

Anuja Shukla
Jaipuria Institute of Management Noida, India

Application of Machine Learning algorithm for creating sustainable omni-channel retail ecosystem

Abstract

The retail ecosystem has evolved from simple Kirana stores to large omni-channel retail systems. The advent of omni-channel retailing has seen an increased count of fraudsters who can abuse the gap between the online and offline channels to perform different types of fraudulent activities. Fraud can be related to payment, account take over, refund and cancellation abuse, collusion of customers with associates. Waste can be due to over-production, sub-optimal pricing or discounts, damaged products, throwaways, availability issues, packaging waste. The paper tries to identify the need for better research on improvement in fraud detection systems for sparse data, improvement of the customer experience and reduction of returns. The paper also discusses the need for adoption of ML models through better interpretability. The data analysis is performed on omni-channel retail transactions over a period of a year across multiple countries. Machine Learning methods of supervised and unsupervised learning can be used to highlight the outlier cases to focus upon, and then identify the root causes of such losses. Then causal discovery models can be used to identify root causes and provide prescriptive recommendations. The paper provides a comprehensive evaluation of the challenges in omni-channel retail ecosystem and the proposed machine learning tools which can help resolve the issue at large. The implementation of ML techniques in retail total loss management can lead to multimillion- dollar savings through better fraud detection and waste management. Performance of supervised models improves to close to 98% over the course of months by using feedback loop.

Keywords

Retail loss, fraud management, waste reduction, Machine Learning, anomaly detection, classification, causal discovery, computer vision

Erratum

Article updated to add co-author's affiliation.

RESEARCH ARTICLE

Application of Machine Learning Algorithm for Creating Sustainable Omni-channel Retail Ecosystem

Somedip Karmakar [a,*], Anuja Shukla [b]

[a] SSBM, Switzerland
[b] Jaipuria Institute of Management Noida, India

Abstract

The retail ecosystem has evolved from simple Kirana stores to large omni-channel retail systems. The advent of omni-channel retailing has seen an increased count of fraudsters who can abuse the gap between the online and offline channels to perform different types of fraudulent activities. Fraud can be related to payment, account take over, refund and cancellation abuse, collusion of customers with associates. Waste can be due to over-production, sub-optimal pricing or discounts, damaged products, throwaways, availability issues, packaging waste.

The paper tries to identify the need for better research on improvement in fraud detection systems for sparse data, improvement of the customer experience and reduction of returns. The paper also discusses the need for adoption of ML models through better interpretability.

The data analysis is performed on omni-channel retail transactions over a period of a year across multiple countries. Machine Learning methods of supervised and unsupervised learning can be used to highlight the outlier cases to focus upon, and then identify the root causes of such losses. Then causal discovery models can be used to identify root causes and provide prescriptive recommendations. The paper provides a comprehensive evaluation of the challenges in omni-channel retail ecosystem and the proposed machine learning tools which can help resolve the issue at large.

The implementation of ML techniques in retail total loss management can lead to multi-million-dollar savings through better fraud detection and waste management. Performance of supervised models improves to close to 98% over the course of months by using feedback loop.

Keywords: Retail loss, Fraud management, Waste reduction, Machine learning, Anomaly detection, Classification, Causal discovery, Computer vision

1. Introduction

Retail business is the sale of goods and services to customers either in-stores or online (Mittal & Prashar, 2011; Shukla & Sharma, 2018a, 2018b). Usually in retail businesses, the profit margins are low, as there is fierce competition on price, leading to increasing cost pressure and growing investment in the infrastructure to support the growth. In this hyper competitive environment, millennials are extremely demanding and are expecting delight and convenience across their journey (Singh, 2013). Hence it is crucial to provide a safe, trustworthy, frictionless, and sustainable omni-channel shopping experience. Since the margins are already quite low, it is also important for retailers to prevent loss as much as feasible. Loss can happen either through shrink, fraud, waste, refunds, cancellations and many others.

According to National Retail Security Survey (2019), 1.38% of total sales is due to shrink, which impacts the retail industry by more than $50.6 billion USD. As e-Commerce retail experiences explosive growth, the amount of fraudulent activity naturally rises in parallel. According to LexisNexis (n.d.) successful retail fraud attempts grows by nearly 30% annually, with e-commerce driving a significant share of the fraudulent transactions. In US, fraudsters cause a loss of more than 5 cents out of 100 dollars profit. In a competitive retail industry operating on razor thin margins, this represents significant loss

Received 18 January 2023; revised 22 May 2023; accepted 2 June 2023.
Available online 19 July 2023

* Corresponding author.
E-mail address: somedip@yahoo.co.in (S. Karmakar).

https://doi.org/10.57198/2583-4932.1324

and essentially increases prices for other non-fraud customers. Preventable returns impact the bottom-line and accounts for 2% of revenue. We lose 8 cents on every revenue dollar to return. Returns loss post recovery is slated at $750 million, out of which $200 million was preventable. Reducing the return rate for a department by 50 bps can improve the overall customer NPS by 1 point. Over 60% of all returns across ecommerce retailers are not a choice of the customer. Retail waste management has always been a very significant issue for the retail industry, which causes huge financial loss as well as huge environmental impact. According to the multiple environmental agencies, US generates around 300 million tons of waste. Even if many of the items can be reused, we see that retailers have disposed of 5 billion pounds returned goods usually in landfills and other dumping grounds. As the e-Commerce industry continues to grow at an exponential rate, returns also proportionally rise which in turn causes huge rise in retail waste volumes (www.actenviro.com, 2022). Optimizing the retail waste management plan can seem daunting but if done correctly, could play a big part in maximizing the bottom line, especially during uncertain economic times.

1.1. Retail fraud

With the advent of omni-channel retail, we also have more customers who are trying to exploit the gap between different channels & indulge in fraudulent activities, and there can be potential collusion with store associates, delivery drivers, marketplace sellers, call center agents, as well as between groups of customers to create organized crime rings. Omni-retail has also seen a rise in returns, especially when the customer expectations don't match. Cancellation is another huge source of loss both in form of abuse as well as process issues. Mismatch between the forecasted demand and actual sales is a major cause of overstocking leading to huge waste. Retailers suffer heavy loss by providing heavy markdowns to sell through the inventory and remaining is being disposed of. Majority of these decisions are driven by business leaders subjectively and not using data-driven advanced data science and machine learning techniques. While there has been some historical focus on payments and credit card fraud detection using supervised learning, there is approximately 10%–15% false positives in such models. Also, there is a huge opportunity of identifying any new types of fraud, since these models only focus on learning from previously identified fraud cases. Another major opportunity area is getting good quality labelled data to train such supervised models. One of the prevalent challenges in any Machine Learning problem is the availability of good quality labelled data, and this problem becomes more pronounced in case of rare event scenarios like fraud, where one of the classes is a minority. Another major part of loss reduction is to reduce returns and cancellations systematically either in pre-purchase or post-purchase scenarios. There is a serious gap of Machine Learning techniques to solve such process opportunities.

1.2. Retail waste management

Sustainability and Waste reduction is a major focus for many companies including retail. However there is no clear guidance or recommendation on the optimal path to follow for such efficient waste management. This is where prescriptive Machine Learning models, especially Causal ML could potentially play a significant role, both in identifying the true root causes of wastage, and to provide guided recommendations on how to reduce further waste. Another major gap with such advanced Machine learning solutions is the lack of interpretability and explain ability which causes difficulty in adoption of such solution by business leaders.

We have highlighted that in retail industry it is crucial to minimise loss through reduction of fraud and waste, and also to become more trustworthy and sustainable. There is a significant gap in the current retail landscape and use of advanced predictive and prescriptive models for such efficient loss reduction.

1.3. Research questions

The objective of the study is to understand how can we use advanced predictive and prescriptive Machine Learning models for retail loss management. Following are the existing gaps in research:

i) Better fraud detection in presence of data and label uncertainty
ii) Return reduction and better experience
iii) Waste management through prescriptive ML models
iv) Omni-channel retail loss management especially in emerging technology
v) ML models adoption in retail loss management through interpretable models

2. Literature review

2.1. Retail total loss review

Detection and prevention of shrink is a major task for all total loss prevention teams in retail.

Shrinkage of inventory shrinkage is defined as the loss due to many attributes like is the associate theft, shoplifting, associate error and partner collusion fraud (Holliger, 2002 National Retail Security Survey).

In the book “The Challenges to Preventing Losses in Retailing” (2018), Martin Gill has mentioned that there is very less research performed on fraud detection mechanisms while keeping track of financial loss of missed sales opportunities.

According to Adrian Beck (2020, Total Retail Loss 2.0), retailers face multifarious challenges for loss prevention like inefficient data management (44%), unavailability of required data storage (23%), challenging organisational plan (14%) and blockers due to management (9%).

Beck (2020) classified retail Total Loss as coming from different sources like Store, eCommerce, Supply Chain and Corporate. These are further divided into known and unknown stock loss. Even within the known loss there are malicious loss (fraud) or non-malicious loss which are process opportunities (refer Fig. 1).

Beck further highlights that, returns due to fraudsters cannot be avoidable and is a serious concern since they cause a huge loss and require a lot of resources to control. Such returns significantly affect the financials of the retailers by eating up 15% or more of profit. Many industry reports suggest that in US, more than 75 billion USD of annual loss is due to returns (National Retail Federation, n.d.). Also, over the years such returns fraud is growing at a rate of more than 17%. Fraudulent activities of renting and wardrobing are cases where apparel items are purchased and used at special events and then returned at full cost. There are also cases where big retailers like Amazon has been scammed by return of dirt in packages in place of actual items (Aguiar & Moynihan, 2019).

Sometimes, associates also collude with family members where the spouse returns items and it is accepted by the associate with knowledge that the items are not properly returned or damaged (Speights & Hilinski, 2005). Pre and post analysis of fraudulent returns related to specific intervention actions have not been performed in the past.

2.2. Retail waste review

The Food and Agriculture Organization of the United Nations (FAO, 2018), has recorded that across the world, more than a billion metric ton food is being wasted annually, which is more than a quarter of food produced and also causes waste of natural resources like water and energy. On top of this some government agencies also report that huge quantities of food is wasted without being eaten, and this could be due to many reasons like over-production of fresh food, improper expiry information and poor quality. The waste which gets generated by physical brick and mortar stores is a serious problem for any omni-channel retailer, since it varies a lot based on the inventory changes and internal promotional events. This unpredictable variation is leading to a system that is not efficient, and causes spike or dip in waste during online

Fig. 1. Adrian Beck's total retail loss 2.0 Model (2020).

pickup orders. Retailers will continue to face challenges of proper disposal and management of waste till there is excess inventory either on shelves or backroom (Admin & Admin, 2022). Such waste cycle starts from the point the items that are ordered are shipped from the manufacturer or sellers to the stores or warehouses. Once these shipments arrive, they are wrapped for display in stores or kept for pickup. The packaging using plastic, paper, fillers and cartons also add on to retail waste, and this is true for large retailers like Amazon and Walmart who do many billions of shipments worldwide.

2.3. Application of machine learning

The role of Machine Learning and Artificial Intelligence to solving retail loss and making it safe and sustainable is critical, although existing research has focused only on solving parts of the problem, mainly focusing on payment fraud prevention. According to Nanduri, Jia, Oka, John Beaver, and Liu (2020), The primary challenge of eCommerce fraud is due to the dynamic nature of fraud patterns, which affect any Machine Learning models that are developed to detect certain types of fraud, because of previously unseen patterns. Further the multiple layers of decision-makers and other partners make a lot of reviews and changes which are not properly captured by the models.

Loss reduction by a fraction of percentage can lead to multi-million dollar additional savings and if the false detection of fraud is reduced by even 1%, then payment authorizations improve manifolds.

Ileberi et al. (2022) have proposed a technique of using Genetic Algorithm for feature selection followed by multiple ML classifiers to predict credit card fraud. The results demonstrated that this approach outperformed existing systems. Deep learning approaches for detecting and attributing cyber-attacks can perform better than some of the traditional machine learning models, as per the research by Soleymanzadeh et al. (2022).

Lokanan and Sharma (2022) has noted that till date no researcher has used machine learning to predict fraud in total loss in omni-channel scenarios, however related fraud like payments risk and financial fraud have used such models in past. Even with a lot of preventive policies set up by regulatory bodies and financial institutions, more advanced technology solutions are required to curb these frauds. One path forward can be to develop testing of hypothesis on different fraud modus operandi to identify the features related to fraud.

Weber and Schütte (2019), mentions that ML can have many applications in different parts of retail and wholesale business. There are many difficulties related to data analysis due to the circumstances of retail business. The huge assortment and fast sales pattern across multiple stores creates huge data quantity, which is not possible for manual analysis. In the past there was low priority of such analysis due to technical challenges of outdated systems, so many retailers don't have active initiatives related to ML (Fig. 2). Although ML based solutions can have a huge opportunity to reduce the cost and optimise the business process, yet such processes are not adopted to full capacity.

Malefors et al. (2022) highlights that in any circumstances, fresh food waste is a complex and persistent issue. Covid-19 pandemic put a lot of pressure on the system, and unearthed underlying systemic issues, especially under fast-changing purchase pattern of customers. During such unforeseen events, it may be nearly impossible to predict beforehand. The alternative strategy can be to take care of surplus supply of fresh food items and forming segmentation strategies during unforeseen acts of nature (Ali et al., 2008; Makkar, 2008). If the demand forecast is not reliable and trustworthy, the store associates producing these fresh items will be reluctant to use and hence we will lose out on the benefits of the ML models (Shukla & Mishra, 2021).

Uganya et al. (2022) proposed a new method for effective and intelligent waste management system using IOT through prediction of the possibility of waste. The wastage capacity and level can be monitored continuously using IoT based bins, which can be placed in the stores.

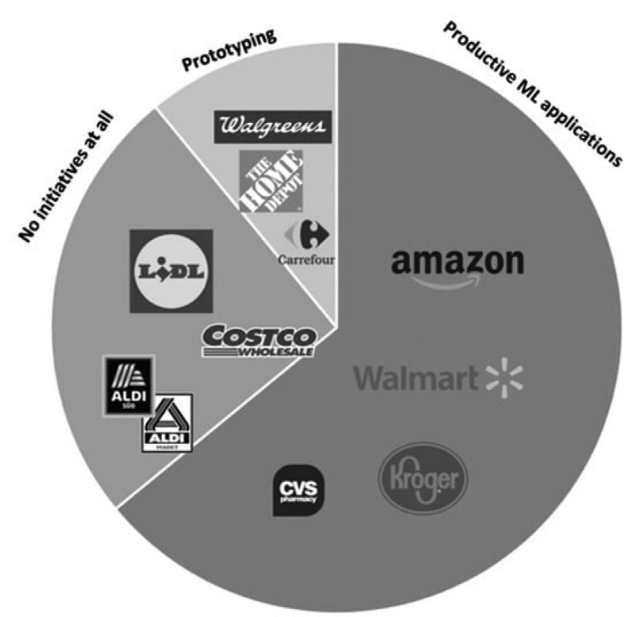

Fig. 2. ML adoption across top retailers in the world.

Glatzel et al. (2016) mentions that a ML based forecasting system can use historical transactions and also additional internal strategies like advertisements and promotional events, operating hours and external attributes like weather and holidays. Additional constraints on delivery timeline, safety stock, pack size, etc can also play an important role.

As per the analysis by Falatouri et al. (2022), among a multitude of traditional and advanced ML models applied for forecasting, LSTM performed better for products with stable demand, while SARIMA showed better results for products with seasonal behaviour. SARIMAX models also performed significantly better for products with promotions. Hybrid approaches by training SARIMA(X) and LSTM on similar, pre-clustered store groups, can improve the forecast quality.

Use of mobile assisted retail shopping (Shukla & Sharma, 2018a, 2018b) and use of smart voice assistants for purchase would also influence retail (Mishra et al., 2022). Retailers need to review their ways of working with different partners like suppliers, negotiate better and also help to improve forecasts of suppliers. Through these steps, they can develop sustainable and long-lasting improvements using the power of Machine Learning.

3. Conclusion

This paper highlights the importance of predictive and prescriptive Machine Learning techniques in retail loss prevention. It identifies the different problem areas in loss prevention like fraud, shrink, waste, availability and provides insights into how such problems can be tackled using Machine Learning techniques like supervised and unsupervised models, anomaly detection, graph analytics, reinforcement learning, computer vision and causal inference. The use of technology in retail can be tested for continuance intention to compare in store as offline retail purchase (Mishra et al., 2023). The future research may focus on identifying the best approaches and designing a ML based system which can solve these problems for omni-channel retail ecosystem and make it safe and sustainable.

Conflict of interest

There is no conflict of interest.

References

Admin, & Admin. (2022). Hazardous Waste Management & Disposal Services - ACTenviro. ACTenviro. https://www.actenviro.com/

Ali, Jabir, Kapoor, Sanjeev, & Nath, Tribhuvan (2008). Consumer's preference on food retail store attributes in India. *Management Dynamics, 8*(No. 2). Article 2. https://doi.org/10.57198/2583-4932.1178.

Aguiar, A. R. , Moynihan, Q., (2019, August 1). A 22-year-old is accused of scamming Amazon out of nearly $370,000 by sending back boxes weighted with dirt instead of the returned product. Business Insider. https://www.businessinsider.com/22-year-old-mud-filled-boxes-scam-amazon-370000-2019-8?r=US&IR=T

Beck, A. (2020). Total Retail Loss 2.0: Moving Beyond the Theory. [online] Loss Prevention Media. Available at: https://losspreventionmedia.com/total-retail-loss-2-0-moving-beyond-the-theory/ [Accessed 20 Nov. 2022].

Glatzel, C., Hopkins, M., Lange, T., & Weiss, U. (2016). *The secret to smarter fresh-food replenishment-machine learning*.

Holliger, R. (n.d.). 2002 National Retail Security Survey: Final Report. [online] Available at: https://www.academia.edu/18038890/2002_National_Retail_Security_Survey_Final_Report [Accessed 20 Nov. 2022].

Ileberi, E., Sun, Y., & Wang, Z. (2022). A machine learning based credit card fraud detection using the GA algorithm for feature selection. *Journal of Big Data, 9*(1), 1–17.

LexisNexis Risk Solutions. (n.d.). The True cost of FraudTM Study. [online] Available at: https://risk.lexisnexis.com/insights-resources/research/us-ca-true-cost-of-fraud-study. [Accessed 20 Nov. 2022].

Lokanan, Mark Eshwar, & Sharma, Kush (2022). Fraud prediction using machine learning: The case of investment advisors in Canada. *Machine Learning with Applications, 8*, Article 100269. ISSN 2666-8270. https://doi.org/10.1016/j.mlwa.2022.100269.

Makkar, Urvashi (2008). Impact of organized retail on the purchase of fresh food items: A comparative analysis of organized and unorganized retail. *Management Dynamics, 8*(No. 2). Article 5. https://doi.org/10.57198/2583-4932.1181.

Malefors, Christopher, Secondi, Luca, Marchetti, Stefano, & Eriksson, Mattias (2022). Food waste reduction and economic savings in times of crisis: The potential of machine learning methods to plan guest attendance in Swedish public catering during the Covid-19 pandemic. *Socio-Economic Planning Sciences, 82*. Part A, 2022, 101041,ISSN 0038-0121. https://doi.org/10.1016/j.seps.2021.101041.

Mishra, A., Shukla, A., Rana, N. P., Currie, W. L., & Dwivedi, Y. K. (2023). Re-Examining post-acceptance model of information systems continuance: A revised theoretical model using MASEM approach. *International Journal of Information Management, 68*, 102571. https://doi.org/10.1016/j.ijinfomgt.2022.102571

Mishra, A., Shukla, A., & Sharma, S. K. (2022). Psychological determinants of users' adoption and word-of-mouth recommendations of smart voice assistants. *International Journal of Information Management, 67*, 102413. https://doi.org/10.1016/j.ijinfomgt.2021.102413

Mittal, K. C., & Prashar, Anupama (2011). A field study on opportunities and challenges faced by organized retailers in tri-city, A field study on opportunities and challenges faced by organized retailers in tri-city. *Management Dynamics, 11*(No. 1). https://doi.org/10.57198/2583-4932.1142. Article 4. Available at: https://managementdynamics.researchcommons.org/journal/vol11/iss1/4.

National Retail Federation. (n.d.). National Retail Security Survey 2021. [online] Available at: https://nrf.com/research/national-retail-security-survey-2021 [Accessed 20 Nov. 2022].

Nanduri, Jay, Jia, Yuting, Oka, Anand, John Beaver, & Liu, Yung-Wen (2020). Microsoft uses machine learning and optimization to reduce E-commerce fraud. *INFORMS Journal on Applied Analytics, 50*(1), 64–79. https://doi.org/10.1287/inte.2019.1017

Shukla, A., & Mishra, A. (2021). Effects of visual information and argument concreteness on purchase intention of consumers towards online hotel booking. *Vision: The Journal of Business Perspective*. , Article 097226292110380. https://doi.org/10.1177/09722629211038069

Shukla, A., & Sharma, S. K. (2018a). A system dynamic approach to analyse the impact of electronic word of mouth on box office revenue. *SAMVAD, 14*, 19–25.

Shukla, A., & Sharma, S. K. (2018b). Evaluating consumers' adoption of mobile technology for grocery shopping: An application of technology acceptance model. *Vision: The Journal of Business Perspective, 22*(2), 185–198. https://doi.org/10.1177/0972262918766136

Singh, Smita (2013). Demystifying the millennials' perception of retail store attributes. *Management Dynamics, 13*(No. 1). Article 3. https://doi.org/10.57198/2583-4932.1116.

Soleymanzadeh, R., Aljasim, M., Qadeer, M. W., & Kashef, R. (2022). Cyberattack and fraud detection using ensemble stacking. *AI, 3*(1), 22–36.

Speights, D., & Hilinski, M. (2005). Return fraud and abuse: How to protect profits. *Retailing Issues Letter, 17*(1), 1–6.

Uganya, G., Rajalakshmi, D., Teekaraman, Y., Kuppusamy, R., & Radhakrishnan, A. (2022). A novel strategy for waste prediction using machine learning algorithm with IoT based intelligent waste management system. *Wireless Communications and Mobile Computing*, 2022.

Weber, F., & Schütte, R. (2019). A domain-oriented analysis of the impact of machine learning—the case of retailing. *Big Data and Cognitive Computing, 3*(1), 11. https://doi.org/10.3390/bdcc3010011

Zhang, D., Frei, R., Senyo, P. K., Bayer, S., Gerding, E., Wills, G., & Beck, A. (2023). Understanding fraudulent returns and mitigation strategies in multichannel retailing. *Journal of Retailing and Consumer Services, 70*, Article 103145.