Machine Intelligence and Consumer Decision Patterns: A Computational Framework for Retail Behavioural Analytics

Dr. Maria Lenstrand, Associate Professor of Information Technology, Uppsala University, Sweden

1. Introduction

This dissertation offers an extensive treatment of customer behavior analytics (CBA), focused more specifically on the retail field and relying on advanced AI techniques, algorithms, and methods. In recent years, artificial intelligence technologies have been shown to be especially useful when large amounts of data need to be interpreted, resulting in a better understanding of customer needs. This makes it a transformative tool for customer behavior analytics, as it helps gain useful insights from a company's raw customer data. Especially in a competitive environment, modern-day enterprises need to understand the actions and preferences of consumers to find effective strategies to increase their sales. Numerous methods and algorithms have been proposed to address the CBA challenge. However, they focus mainly on historical customer data, and the algorithms are not intended for retail end use.

The use of AI techniques leads to a complete revamping of the way to study CBA, striving to make it more effective. Many factors shape the direction of this research, including advances in digital technology revolutions and changing customer expectations that make consumers demand highly personalized services and product offerings. In line with this, understanding the adoption, use, and preferences of novel technologies is now essential. There are equally "new cultures" that are constantly shaping and being shaped by consumer behavior. This forms the genesis of the need for research into customer behavior analytics as a new way to understand customer actions and preferences. This is especially important because, although the issue of insight into human behavior has received increasing attention in recent research, much of it has focused on behavior as a whole.

1.1. Background and Significance

1.1. Background

There is increasing interest in customer behavior analytics in the retail industry. There have been major advances in the field of data collection and processing, which have revolutionized analysis abilities. The amount of data collected every day is increasing rapidly. Retailers would do well to capitalize on this growing ocean of data to have a clear picture of the market. A greater understanding of customer lifestyles or consumer behavior patterns, for example, may help companies selling retail or consumer goods improve their services and build customer loyalty. Additionally, understanding behavioral patterns can help banks and other financial institutions create a product mix that aligns with customers' needs. In summary, understanding customer behavior has the potential to provide a competitive advantage for businesses.

The analytical matrix representing a nexus of what you sell with to whom you sell combined with who is selling it is termed customer behavior analytics. At present, this detailed information is widely available but not widely examined. Retailers who capture and analyze this data using the most recent techniques might enhance their value proposition. Despite all these promises, retailers face some critical hurdles in exploiting these analytics. This study seeks to highlight the significance of integrating AI with data produced through traditional analytical techniques. These are the different concerns worth considering for future research that were raised throughout the ensuing investigation. The usage of AI to increase customer behavior analysis capabilities is not discussed in detail in the study. The focus of retailers on predictive analysis that integrates AI and customer movement data is not clearly known. Ongoing experimentation is being carried out in this area, which will entail further investigation and findings. AI anticipations for how customer behavior data can be utilized are... The genuine prospective use might vary from what has been predicted. The ongoing investigation aims to hear directly from industry participants on how they wish to utilize predictions and on what retail advantages will be realized.

1.2. Research Objectives

Following the research gaps revealed in the section above, our research aims to fulfill the following objectives: 1) Identify the key factors that influence customer behavior and how they interact, using various AI-based data analytics methodologies. These factors or drivers will largely refer to those details that can be influenced, managed, or controlled by decision-makers in retail environments, such as store attributes, promotional activities, marketing strategies, etc. They should represent more practically actionable insights. 2) Investigate the conformance of data-driven methods with manually collected surveys. The characteristics and the ability to provide comparable results are important to increase the understanding of how such analysis matters in practical retail applications. 3) Examine how various AI methodologies based on connectionist and non-connectionist AI methods can be used independently and cooperatively to uncover key customer behavioral factors that directly increase the businesses' per capita revenue or have a relationship with customer return frequency behaviors in retail environments. 4) Evaluate the effect of investing in AI-based customer behavior analyses, customer attraction and retention strategies, and other associated tools on per capita revenue and other economic indicators in real business scenarios. 5) Investigate the performance of the AI methods both in terms of model diagnostics and predictive power. 6) Explore the challenges of real-world retailer practices of AI investments, providing innovative solutions for business adoption of AI technologies in conjunction with actual commercial customer data.

2. Fundamentals of Customer Behavior Analytics

In retail, understanding various aspects of customer behavior and making customer-centric decisions is always important. With the availability of massive data on business transactions and customer interactions that retail companies have accumulated, data analytics on customer behavioral information becomes possible. In principle, customer behavior analysis discovers meaningful patterns and makes strategic decisions. We illustrate the concept of customer analytics throughout this work to represent the knowledge discovery using retail-related data that is based on clear business objectives. Retail analytics is the process of providing shop owners with the ability to make informed decisions using data. It is a set of analytical techniques that are used to support a wide range of retail decisions including assortment selection, price optimization, customer observation, customer feedback, and store operations management. Retailers can gain a deeper insight into the analysis of customer behavior and preferences to enhance their operations and ultimately benefit from the knowledge. Additional data sources such as customer feedback and comments exist. The plethora of analytics research and techniques includes an increasing number of topics that can be used for unstructured data analytics, such as natural language processing and deep learning methods. Tracking and understanding customer behavior has been identified as a prerequisite to managing businesses effectively. In the context of the retail business, the more we understand about customer behavior, perhaps the better we are able to predict the business trend and its growth, which leads to improved profit. Different from traditional data analysis, customer behavior analysis is capable of predicting business outcomes from customer behavioral patterns.

2.1. Key Concepts and Definitions

Customer behavior is represented by actions performed by customers in a specific transactional environment. The retail space assumes the interchange between consumers and retailers. In the context of the entire retail life cycle, the first concept that retailers work with is customer segmentation. The aim is to understand which are the most important customer segments, in terms of volume or profit, and to establish the marketing strategies needed to convert customers in different segments to perform targeted behavior beneficial to the retailer. In the retail context, customers interact with the retailer by making purchases in one or more stores and talking to customer service.

A desired behavior in this context is loyalty, when customers visit a store or a particular retail network. However, to isolate the shopping chain pattern from the individual service pattern, the concept of customer lifetime value is used. The automatic relevance of customer behavior forecasting in retail, based mainly on transactional relationships, becomes apparent. In the context of AI analytics performed in retail environments, the power of big data in areas like customer analytics can be added to behavioral aspects. Also, the knowledge of the behavior pattern of an individual customer plays a key role in automation tasks in the area of retailing, where customers interact with sellers. The establishment of the basic concepts that are fostered when discussing customer behavior inference in a retail environment necessarily requires the precise delimitation and comprehensive approach to each term. This facilitates a coherent extension from the main concepts to their AI-based implementation or adaptation within the frame of behavioral data analytics.

2.2. Data Sources in Retail Analytics

In retail, customer behavior analytics involves clustering and segmenting consumer data to understand preferences, enabling the identification of purchasing behavior. Retail analytics data sources include transactional data that describes sales volumes, products bought, and individual transactions; customer demographic data such as age, gender, income, and location; and customer behavior data that reflects the frequency of store visits, website traffic, time spent browsing, product views, orders, and abandoned shopping carts. Different sources of customer feedback are also collected interactively in retail in the form of surveys such as face-to-face solicitation, phone calls, and email after transactions and browsing. Analytics help determine the optimal product assortment in a concept category through a variety of store formats that cater to the targeted consumer segment, taking into consideration competitive positions, occasions, and volume.

As the customer segment and the purchase channel evolve with changes in the economic climate, technology, competitive landscape, and customer preferences, an evolving concept category and store-related data are critical forces that drive retail. A main data source in retail revolves around shopper-related sales transactions. These transactions capture data on the product purchased, the price paid, the location and date a customer purchased the product, and whether the customer was new or had previously bought a product from the retailer. However, such transaction data do not include naming information regarding the purchaser, such as name, credit card, or ID number. Technology in database processing known as data warehousing structures and stores large volumes of data effectively and with high throughput for analysis and decision-making.

3. Machine Learning Algorithms for Retail Analytics

Machine learning algorithms are vital in retail. They enable the processing of large datasets of structured and unstructured data, which lead to decision-making processes to gain insights that can enhance consumer experiences. Different types of machine learning algorithms are used to consolidate data for analytics. Among the broad types of machine learning that can be applied are supervised machine learning (the algorithm trains on a labeled dataset), unsupervised machine learning (the algorithm trains on an unlabeled dataset), and reinforcement learning (the algorithm is rewarded for selecting the correct action). Depending on the analytical objectives, different aspects of machine learning can be used to develop, train, and test predictive models that can be adopted to make recommendations and classify, identify, and predict outcomes. In retail analytics, one of the objectives includes clustering the metrics of customer experience.

Machine learning in retail has been applied for consumer behavior and characteristics and the prediction of retailers' strategies. Supervised, unsupervised, and reinforcement machine learning algorithms have been applied in the following: predicting future needs: goods, items, and products; recommending different goods and services based on previous behaviors; and detecting patterns of repeated behaviors in consumer habits. There is a need to avoid applying the wrong machine learning algorithms, as they can provide unusual outcomes or misleading results. Machine learning algorithms select a way to model the data based on clearly marked goals and datasets. With the advancement in computer capabilities, including more power and data to train an algorithm, all aspects of machine learning have been enhanced in accuracy and results to a higher extent.

3.1. Supervised Learning

Machine learning, a subset of artificial intelligence, has played an increasingly significant role in retail analytics. Retailers can leverage the power of accurate and reliable sales forecasting and customer insights that machine learning models are generally built upon. Supervised learning is a fundamental component of machine learning that primarily focuses on learning from data labeled with input-output pairs. In the context of retail customer analytics, some examples of input-output situations are given as follows. Based on a unified transaction history of many customers, a supervised learning algorithm is usually employed to group or partition customers. The labels could be clustered patterns such as "loyal customer," "regular customer," "infrequent customer," "likely big spender," etc. Moreover, a supervised learning algorithm could also be used to predict expected outcomes, for example, "whether the customer will shop again anytime soon" based on her historical shopping patterns. The main purpose of learning in retail analytics is to study the historic customer behavioral patterns and predict the possible future customer actions, so better business decisions could be made upon the acquired knowledge.

In a typical scenario of supervised machine learning, a classifier or a model builder supervises predictions based on the historical data labels. In other words, machine learning models are trained to learn from the previously collected data, where the model builders already know what the output is supposed to be when presented with an input. This kind of problem is called a supervised learning problem, where the model learning or model prediction occurs under guidance or supervision from the training dataset. The training dataset contains a known number of input-output samples: each sample is a record of the attributes of a given object along with correct labels of what or whom the given object belongs to. The prediction model subsequently captures an underlying input-to-output mapping pattern which is applicable to previously unseen record inputs. To build a competitive model, information derived from training data is crucial for feature retrieval, feature engineering, model selection, hyperparameter optimization, model evaluation, and validation with different performance scores.

Because the model captures a pattern of the historical input-to-output entries, accurate and complete information derived from training data is critical to establishing a realistic model mapping pattern. In other words, the prediction of novel inputs with input values that are not covered by the training inputs can only be accurate if the model builder is able to cover the entire variability range of an input feature. Furthermore, once a model is selected and trained based on training data, a set of confidence intervals for predictions could be possible. For example, a 90% confidence interval indicates the model is 90% sure that the true value of an input feature will fall within the specified range. Because predicting the outside-of-range values is unreliable, this kind of quantified uncertainty could be useful. However, this kind of "what is inside is what is outside" logic is fully based on the assumption that the input feature values follow the built model's input distribution. In practice, these aforementioned assumptions should be cautiously discussed and determined for each set of input features. The more accurate choices are, the more competitive the modeled outputs are. In a real-world experiment, choosing input features is a part of optimizing machine learning model hyperparameters. A good choice could make the model more generalized and ready for application to different datasets, so that a better prediction result is likely to occur.

3.2. Unsupervised Learning

Unlike supervised learning, unsupervised learning does not require labeled data. This makes it appropriate for use in a wide range of scenarios. Unsupervised learning is helpful in identifying underlying patterns or structures among observations. In the context of customer behavior analytics, unsupervised learning helps identify consumers with similar behaviors, a key part of market segmentation. Successful market segmentation and recommended customer segment selection can benefit customer relationship management strategies and overall business performance. Common applications of unsupervised learning include clustering and association analysis.

There are several benefits to utilizing unsupervised learning in market and customer behavior analytic applications, as no domain knowledge is required to encode rules that define each segment. More hidden segments and knowledge may be discovered; unsupervised learning can improve the possibilities of insight-driven business management approaches, giving companies a competitive edge in the market. There is growing evidence that, with the right practices, unsupervised learning techniques can be effectively used to overcome business challenges and bring commercial opportunities for companies. In retail and business management, various unsupervised learning methods have been employed, such as clustering, hierarchical clustering, benchmarking clustering, time-series clustering, and market basket analysis association rules. However, there are several research challenges and limitations related to unsupervised learning in the retail and business sectors, including difficulty in interpreting clustering results and deciding the best number of clusters, challenges related to researcher and manager subjectivity, and the sensitivity of clustering results to data preprocessing and scaling operations.

The main idea of unsupervised learning is to uncover the structure of data that is not labeled and is less explored than supervised learning methods. We may use unsupervised learning techniques to understand the data set, uncover the insights and patterns buried in the data, and partition the data set according to the attributes of the observations. One of the most common applications of unsupervised learning is clustering in market segmentation. Each consumer cluster is thought to have similar behavioral characteristics and is unlike those from other clusters. This finding can help managers understand how different segments react differently to service quality factors and make predictions.

3.3. Reinforcement Learning

Reinforcement learning is a new area of machine learning incorporating principles from game theory and operations research. Special interest in reinforcement learning is driven by its application in a dynamic environment, where algorithms learn to solve a problem through repeated interactions, each time choosing an action based upon the predicted reward or outcome. The distinguishing characteristic and primary advantage of reinforcement learning is its capability to automatically learn a learning policy while receiving feedback from the environment. The only thing the program requires is to understand the state space and set of admissible actions of the environment to control. Reinforcement learning can be applicable in a retail context where strategies have to adapt and test with the dynamic environment, like dynamic pricing and inventory management.

A few interesting case studies prove the potential of using reinforcement learning in a commercial environment. For example, reinforcement learning is used to automate optimal prices in a high-frequency trading setting. More traditional bricks-and-mortar retail-oriented applications of reinforcement learning have been shown to improve product recommendation efficacy. A key challenge of applying reinforcement learning is finding proper policies from given noisy data. Efficient exploration strategies are crucial for rapid exploration of unknown states and improved generalization of the system. Therefore, future research can be directed towards better techniques in designing exploration policies, evaluating what an improvement should look like, and adapting to changes in the trading environment. However, in retail, trading agents can explore the market because markets do not change due to an agent's action. Another area of research in dynamic programming in inventory optimization can be pursued since NP-hard problems are memory limited. Thus, memory efficiency is one of the simple characteristics of reinforcement learning algorithms. Reinforcement learning can learn over time to catch customer behavior in a complex environment since creating the environment as close to the actual will be the basic idea of learning the strategy.

4. Case Studies and Applications

Owing to the increase in information about customer transactions, the growing popularity of AI in business, and the support of cutting-edge AI technology provided by commercial and open-source systems, the industry is now well placed for AI to become closely integrated with real business. Therefore, there is an immediate need for business solutions that can utilize AI-based customer behavior analysis tools to achieve the same quality of results in 1:1 targeted marketing that companies have been developing for years. We show many case studies that let us understand how AI tools applied to big data can be effectively integrated into business operations to achieve a specific strategic approach of a company. For retail, analytics around customer behavior can lead to a more strategic approach and better decision-making that impacts operational efficiency. Real-world applications and case studies of customer behavior analytics can be found in many areas. These applications range from small companies to large-scale industries. From personalization, we generalize the strategies to be employed according to the application field and analytics that are applied. These perspectives can be extended to every retail application, presenting a shift from theory to practice. Specific applications that interest us most are personalizing the customer experience to: increase customer engagement by applying personalized recommendations, increase customer purchase frequency, decrease customer churn, decrease customer complaints, and provide relevant upsell opportunities at the time of purchase to maximize sales. In recent work, we also found that analytics around customer behavior and social media data can provide fast feedback from customers to improve product quality. In this study, we focus on providing the challenges associated with the implementation of the retail applications that rely on online customer usage attribute (behavior) applications in a big data setting.

4.1. Personalized Recommendations

For years, retail experts and practitioners from various fields have been actively applying artificial intelligence to customer data, knowing that every day, billions of people interact with each other using digital channels. The fast-growing number of published reviews and white papers confirms the fundamental value given the objective of this study.

APPLICATION OF AI IN CUSTOMER BEHAVIOR ANALYTICS

Personalized Recommendations

The most common AI application in directional customer behavior analytics is personalized recommendations that change the concept of the whole spectrum of customer interaction enhancement through customer data knowledge. Personalized recommendations made by recommendation systems, or recommenders, are based on natural language processing, machine learning, deep learning, and other models used in customer data analysis technologies to be aware of customer behavior in a digital shop.

A recommendation system is typically a computer program that presents a buyer with items they are most likely to purchase, use, or enjoy. There are several algorithms that form mechanisms designed in a recommendation system: collaborative filtering, content-based filtering, combined approaches, hybrid mechanisms, and reinforcement learning. In e-commerce, a recommendation system can be useful in resistant collaboration and revealed collaboration in streams. As a result, it gives non-zero penetration into a recommendation system as a prerequisite in e-commerce-based customer support, as passive communication and language apply to the recommendation system. Potential buyers provide access to collateral new insights for engagement in e-commerce through a recommendation system. AI models must be highly adaptive in order to remain effective; they must process a lot of data in real time and continually learn from it using historical data. As a result, algorithms should be adjusted as they use new real-time data. AI is really adept at analyzing data and learning from it, and it will power retail in the future. In addition, retail use cases demonstrate that shopper and customer behavioral analytics can have significant benefits. A significant percentage of tapped businesses realized a positive impact on sales.

4.2. Dynamic Pricing Strategies

Retail products and services are typically priced based on the costs of production, delivery, and profit margin agreed. However, instead of a fixed price strategy, the retailers can also use dynamic pricing strategies. These strategies change the prices based on data-driven responses to customer demand and other factors. In contrast to the traditional retailers, companies use dynamic pricing for real-time pricing, especially in a digital retail environment. To determine the price, AI-based algorithms are used in analyzing the historical data, market trends, and customer behaviors. There are few parameters that might be involved in those algorithms, for example, demand level, stock level, and customer segmentation.

At the most basic level, such strategies maximize profits given a change in demand. When sales are slow, the price is dropped and people buy more. When sales are up and inventories are low, prices generally rise. For a retailer or service provider, having the capability to efficiently sense changes in consumer demand deals with the perennial sales-marketing problem of establishing pricing and promotions. Identifying the optimal initial price value and the adaptive price policies are the core problems in contemporary pricing strategy. From a retailer's perspective, pricing decisions are driven by a combination of quantitative and qualitative sources including data on customers, competitors, products, and market forces. All these factors must be carefully monitored, analyzed, and evaluated in order to assess the current situation of the overall macro- and micro-environments within which they operate.

The AI system was set to adjust ticket prices not just based on competition but also on the time of the booking. It was noted that many customers book their stays just after payday, and the AI has worked out the best times to enhance that demand and encourage it to be even stronger, which not only increases revenues but also profitability. Regarding those who fear that the price adjusts too high – it does not. It was also mentioned that by managing demand, the AI is able to increase that demand and also the length of time the hotel is full, through knowing the behavior of different customers. There are already many cases where various online websites have tried it, and some antitrust cases are still ongoing. It is important to differentiate between pricing and discrimination. Most importantly, it is essential to communicate to the customer the value and the reasons behind it so they can understand the additional price. Thus, being transparent is crucial. Current pricing, customer value, current stock, package value, and booking volume, as well as extra costs the hotel requires for special room types and pricing outcome, are considered on a moment-by-room classification for individual rooms. The fact that the system considers interest in special rooms and relies on personal data could allow the supplier to claim an exception. If personal data is used for that, and it can be proven that people are being selected because of their interest, then under some interpretations of regulations, it could be that there is a legitimate interest, subject to meeting data protection requirements.

5. Challenges and Future Directions

Retail has only started to take the first steps in the application of AI technologies for customer behavior analytics in general, while very little has been done in the area of AI-driven analytics in real-time retailing. This section addresses potential issues in designing and implementing AI applications. These transdisciplinary works in the deep learning community also open many future research directions in retail analytics. As a result, retail analytics is at a very early stage and its related research efforts remain uninvestigated. As with traditional analytics and technology, some inherent non-technical issues need to be considered and overcome in using AI technologies.

General issues might include organizational, technical, and ethical aspects of retail, while those directly associated with the use of AI include skills, organizational capabilities, data privacy, and regulation. AI-based analytical systems in retailing rely on customer data. Data protection regulations require significant effort in the management and storage of personal data. Data collected often lack quality, as they were not generated to enable an in-depth analysis of customer behavior. Solving such issues requires new insight into the phenomenon of sustainability and the waste of resources, decision-support tools able to overcome traditional staples, and innovative methodological approaches to improve efficiency. The above challenges present a unique opportunity for innovation. These challenges need a combination of academic and industry collaboration to be overcome. If successful, the proposed retail analytics areas might play a small part in the journey from 'Doing Digital' to 'Being Digital'.

Digital transformation not only concerns organizations themselves but also the society and the individuals that compose them. It affects our concept of society and business and hence the way the economy will be developed in the future. Our vision on AI requires a careful deployment of AI-based analytics, which could usher in a sustainable evolution that proliferates. To achieve this, the challenges identified that pertain to change resistance, ethical decision tasks, and lack of data access for the correct training of AI modules, among many others not presented, will need to be turned into strategic guidelines and addressed to move from doing AI in retail to being AI from an enterprise integration and management standpoint. Results of these activities will help frame AI as facilitating useful innovation, rather than being seen as handicapping.

6. Conclusion

This paper has highlighted AI-based customer behavior analytics in retail and aims to understand the implications that accompany this. Retailers today are faced with an increasingly complex environment, both from the store and the customer perspective. As already stated, real-time customer behavior data is collected through various channels, allowing retailers to integrate this data into back-end systems. Several AI technologies can help provide valuable insights into the comprehensive and complex data. The following implications emerge from this research. Most importantly, the acceptance of the transformative effect of AI on retailers' understanding of customer behavior is only the beginning. This will, to a certain extent, has already drastically changed the way customer behavior can be impacted. Secondly, with increasing varieties of AI, we can expect to see new and different forms of customer analytics service opportunities. A third and final implication is a need for retailer academia to concentrate on proactive rather than reactive research and service development.

The main findings of the research have shown that most retailers are not applying artificial intelligence and machine learning algorithms for customer analytics, even though they are seen to have great potential. In addition, it is found that data-driven decision-making is part of the retailers' vision but is hampered by several challenges. Organizations that have already started to work on AI in combination with customer analytics have taken an AI-first approach and invest in data capabilities with the long-term goal that AI should drive parts of their processes and form the base of their analytics infrastructure. The research has also shown that not only factors at the organizational level pose a challenge to retailers, but also a lack of readiness, which is partly the result of decision-makers not willing to change the organization, not willing to interpret hidden insights in customer analytics, and not willing to learn.

The individual case studies have shown us that artificial intelligence can be seen as an enabler to understand and analyze large amounts of customer data, and as such, it helps retailers support data-driven decision-making and follow a customer-centric business policy. However, the research has shown that external factors may act as a barrier to retailers when fully adopting AI in this dynamic sector. With customers undergoing various reasons for this research gap, this research has adopted a retailer perspective and has investigated how retailers are currently using and preparing for AI-based customer behavior analytics and what challenges are currently hampering the application of AI in combination with customer analytics. Even though organization-wide application is currently in its infancy, current practices and studies seem to support the growing belief that customer analytics will continue to grow in importance.

Opportunities for further academic research in the area of AI in combination with customer analytics and practical implications for application include a step-by-step transition to an organization that adopts AI for customer behavior analytics, and applications in retail should opt for the use of big data instead of only information from transaction databases to create an edge in the highly competitive market that is retail. The retail sector requires a continuous pull of innovation and self-learning over time, even if the customer base remains the same. As suggested in the case studies, the use of AI in customer analytics in retail is expected to grow to form the basis of the customer analytics infrastructure in the future by piecing together varying levels of data.