European Economic Letters
ISSN 2323-5233
Vol 14, Issue 2 (2024)
http://eelet.org.uk

Associating Behavioural Patterns with Machine Learning to Understand Consumer Behaviour – A Jewellery Industry Case Study

Dipanwita Dey[1], Dr. Shantanu Chakraborty[2], Amit Chakraborty[3]
Department of Management Studies, Swami Vivekananda University[1,2]
Center for Data Science, JIS IASR[3]
post.for.phd@gmail.com1
shantanuc@svu.ac.in2
amit13.ons@gmail.com 3

Abstract
The current economic landscape presents a dynamic scenario marked by globalization, intensifying competition, and rapid advancements in communication and information technology. In response to these challenges, businesses are compelled to move away from conventional marketing principles and embrace a customer-centric paradigm. This approach prioritizes the effective management of customer relationships as the cornerstone of success. Improving customer relationships at different stages of the customer lifecycle through machine learning is a transformative approach that has revolutionized the way businesses interact with their clientele. In this paper, we delve into the potential of utilizing machine learning techniques to empower businesses in multiple phases of the customer lifecycle. Specifically, during the acquisition phase, machine learning proves invaluable in enabling targeted marketing campaigns increasing Return on Investment (ROI) and crafting personalized offers. This, in turn, yields higher conversion rates and effectively reduces acquisition costs, optimizing the early stages of customer interaction. As customers progress to the retention phase, machine learning models come into play, forecasting customer churn and providing valuable insights for proactive retention strategies. By anticipating potential churn, companies can take swift actions to retain their most valuable customers. Additionally, as businesses seek opportunities for expansion, machine learning unveils upselling and cross-selling prospects, tapping into increased revenue streams from their existing customer base. The research is conducted with retail data from premium jewellery shop.

Keywords - globalization, customer-centric paradigm, machine learning, customer lifecycle, retention strategie

Introduction

For thousands of years, India has cherished jewellery, excelling in pearl production during the ancient Indus Valley period. Over two millennia, India dominated gemstone supply worldwide, valuing these adornments for their symbolism of power, success, and prestige, especially for women. Today, the jewellery industry is a major economic force, contributing 7% to India's GDP and 15.71% to merchandise exports, employing over 4.64 million people. In 2020, India accounted for 3.5% of global jewellery exports, ranking among the top seven exporters. Diamonds led the way, followed by silver-cut diamonds, lab-grown diamonds, and synthetic stones. This industry significantly bolsters India's foreign exchange reserves, with exports surging by 54.13% to $39.14 billion in 2021-22. The Indian government aims to achieve $100 billion in jewellery exports by 2027. A premium jewellery brands wants to implement personalized customer specific approach by using Customer Life Cycle Management. This lifecycle has various phase describing in the Fig.1

- Acquisition - how a company turns a potential customer into a new customer.
- Usage – frequency and using pattern of product by customer
- Cross -Sell - marketing additional products to existing customers
- Customer Experience - consumer reactions throughout every phase of the life cycle
- Retention – Keep the existing customer with business
- Advocacy - convert into a company advocate.

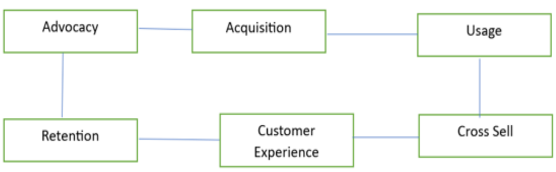

Fig 1. Different Phases of Customer Lifecycle Management

European Economic Letters
ISSN 2323-5233
Vol 14, Issue 2 (2024)
http://eelet.org.uk

Behavioural Economics for Jewellery Industry for Indian Perspective

Behavioural economics, an interdisciplinary field that combines concepts from psychology and economics, tries to understand how people make economic choices. In contrast to classical economics, which presupposes rational decision-making, behavioural economics recognizes that humans frequently behave irrationally due to cognitive biases, emotions, and social pressures. This concept is especially pertinent to India's jewellery business, which is rich in cultural importance, emotional attachments, and significant financial investment. This paper investigates how behavioural economics concepts emerge in the Indian jewellery sector, impacting consumer behaviour, marketing methods, and pricing structures. Jewellery in India is more than just a product; it is an essential component of the socio-cultural fabric. It is essential for religious rituals, weddings, and other important life events. This cultural context substantially influences consumer behavior, frequently. This leads to decisions based on tradition rather than realistic economic factors. For example, gold is typically regarded as a symbol of riches and prosperity, resulting in persistent demand despite price volatility.

Social conventions and cultural expectations have a significant impact on purchase behaviour. The urge to conform to conventional standards, especially at weddings, frequently results in considerable jewellery purchases. This conduct is consistent with the concept of "herd behaviour," in which people replicate the acts of a larger group, assuming that those actions are indicative of proper behaviour. Emotional attachment to jewellery, which is frequently considered as heirlooms passed down through generations, encourages heuristic decision-making. Heuristics are mental shortcuts that make decision-making easier but may lead to biases. The "affect heuristic" plays an important role in jewellery buying.role. Consumers' decisions are influenced by feelings connected with jewellery, such as love, status, and security, rather of a detailed cost-benefit evaluation.

Furthermore, the "endowment effect" is common in the jewellery industry. This cognitive bias causes people to overvalue their possessions merely because they own them. As a result, even during times of financial crisis, consumers may be hesitant to sell their jewellery, preferring to keep it because of its perceived higher sentimental and financial value. Anchoring and framing effects are effectively used by jewellery sellers in India to influence consumer perceptions and decisions. Anchoring happens when people make decisions based mostly on the first piece of information they receive (the "anchor"). Retailers frequently display high-priced items prominently as a reference point, making other products appear more Items deemed to be unusual or unique are given higher value. This method is especially beneficial in the Indian market, as customers frequently seek uniqueness in their purchases for special occasions.

For example, commercial campaigns showcasing limited edition jewellery for festivals such as Diwali or unique collections for wedding seasons instil fear of missing out (FOMO), prompting immediate purchases. Scarcity has a strong psychological influence since it prompts a competitive response and increases the product's perceived value. Richard Thaler proposed the notion of mental accounting, which refers to the cognitive process by which people categorize, assess, and track their financial activities. When purchasing jewellery, people frequently dedicate monies for certain goals, such as weddings or investments. The mental compartmentalization determines how much they are willing to pay for jewellery.

Jewellery stores take advantage of this by providing a variety of payment methods and schemes that are compatible with consumers' mental accounts. Instalment arrangements, gold savings schemes, and exchange offers appeal to various financial mental accounts, making large purchases appear more doable. Retailers can improve affordability and boost sales by organizing payments in accordance with consumers' mental accounting processes. Price perception is another important factor controlled by behavioural economics. Consumers' perceptions of pricing fairness and value for money have a substantial impact on their purchasing decisions. The notion of "reference pricing," in which customers compare the current price to a reference point, is widely utilized in the jewellery industry. Sales and discounts are presented in a way that emphasizes significant savings. Compared to a higher reference price, the discounted price appears more appealing.

Furthermore, the decoy effect, in which the availability of a higher-priced choice persuades customers to select a mid-range option, is widely used. By presenting a high-priced product, shops can direct people to a more profitable mid-priced product, which appears to be a more reasonable and valuable option.

Literature Review

Relationship Marketing/Management

During the 1990s, there was a burgeoning interest in Relationship Marketing (RM). Pioneering research, exemplified by the work of Reichheld and Sasser in 1990, illuminated the significant profitability gains attainable through even marginal improvements in customer retention rates. This revelation heightened awareness within the marketing community about the imperative of cultivating enduring customer relationships. Presently, it is an unequivocal consensus that customers form the nucleus of any business, and as such, a company's triumph is intricately tied to its adept management of these relationships. In 1994, Gronroos, one of the pioneers in the field of Relationship Marketing (RM), offered a definition that characterizes RM as a marketing approach focused on the identification, establishment, maintenance, and enhancement of relationships with customers and stakeholders, all while ensuring profitability and the achievement of objectives for all parties involved. This involves a reciprocal exchange of fulfilment and promises. This definition underscores the necessity of treating each customer as an individual, necessitating the company's ability to discern the particular stage of the customer relationship and provide tailored interactions. This framework subsequently led to the development of the concept of the customer life cycle within RM thinking. In a defining manner, RM has been characterized by Swift (2001) as an encompassing strategy aimed at comprehending and shaping customer actions through significant interactions, with the ultimate goal of enhancing customer acquisition, retention, and loyalty. Andersen and colleagues (2006) assert that the trust and commitment cultivated between banks and their clientele represent the preeminent source of competitive advantage for financial institutions. This insight elucidates the shift in the banking sector, transitioning from transaction-centric models to relationship-oriented paradigms. Their study revealed that customers in the banking sector are increasingly discontented with product-centered approaches, as they desire more consultative and attentive interactions. Consequently, adopting a customer-centric approach in the fiercely competitive banking industry, characterized by growing standardization of products and services, provides a distinctive and differentiating factor.

Machine Learning Model

According to Erevelles, Fukawa & Swayne (2016) Machine learning algorithms have a wealth of research data at their disposal, facilitating their involvement in the investigation of customer lifetime value within the research domain. Scholars harness machine learning techniques to unearth valuable insights from the concealed customer behavioural information within datasets. Traditionally, marketing and consumer behaviour studies have relied on theoretical frameworks and straightforward parameter models rooted in interpretive principles. The examination of customer lifetime value has followed a similar pattern. However, as technology and methodologies advance, data mining technology is poised to revolutionize research methodologies in the marketing and consumer behaviour field, shifting from deductive approaches to inductive ones.

Objective

We tried to end-to-end customer life cycle solution by using machine learning algorithm. Customer Acquisition is optimized by classification algorithms logistic regression. It will find out the most probable lead conversion optimizing the acquisition cost and ROI. In the case of upsell, the customer are segmented into multiple segmentation. Each segment/cluster has unique marketing strategy and offers. Existing customer will be recommended a set of products by using market basket analysis. During the retention phase , random forest algorithm is used to identify the customer to be churned out in near future.

Methodology

A) **Segmentation**: In the current segmentation approach, we utilize clustering methods as the primary means of segmentation. Specifically, the k-means algorithm, an unsupervised learning technique, is employed for this purpose. Determining the appropriate number of clusters involves some guidelines: a) Ensuring that the smallest cluster is not excessively small, typically not less than 2%-3% of the population, b) Ensuring that the largest cluster doesn't become too large, usually within the range of 35%-40%, c) Maximizing the distance between the centroids of clusters to enhance their distinctiveness, and d) Minimizing the radius of clusters. These criteria correspond to inter-cluster heterogeneity (c) and intra-cluster homogeneity (d). To identify the optimal cluster count, a combination of the Scree plot and the Elbow method is applied.

B) **Association Rule Mining**: Association Rule Mining is a data mining technique used in the field of machine learning and data analysis. It focuses on discovering interesting patterns, relationships, and associations within large datasets. This technique is particularly valuable in market basket analysis, where it helps businesses understand customer buying behaviour and make data-driven decisions. The core concept behind Association Rule Mining is the identification of frequently occurring itemset in a dataset. An itemset is a collection of items or attributes that tend to appear together. These itemset are then used to generate association rules, which describe the relationships between various items or attributes.

Support: These measures how frequently an itemset appears in the dataset. It indicates the popularity of the itemset and is usually expressed as a percentage of transactions that contain the itemset.

Confidence: This measures the strength of the association between two items. It is calculated as the percentage of transactions containing one item that also contain another item. High confidence suggests a strong association.

C) Logistic Regression: Logistic Regression is a statistical and machine learning technique used for binary classification problems. It's a widely applied method for modelling the probability of a binary outcome, which means it's used when the dependent variable is dichotomous (having two categories, usually 0 and 1).

Binary Classification: Logistic regression is primarily used for binary classification tasks. For example, it can be employed to predict whether an email is spam (1) or not spam (0), whether a customer will buy a product (1) or not (0), or whether a patient has a disease (1) or does not have the disease (0).

Sigmoid Function: The logistic regression model uses the logistic function (also known as the sigmoid function) to transform a linear combination of input features into a probability value between 0 and 1. The sigmoid function "S" is defined as $S(z) = 1 / (1 + e^{(-z)})$, where "z" is the linear combination of features.

D) Kolmogorov-Smirnov Table: A KS (Kolmogorov-Smirnov) table, often referred to as a KS statistic table or KS chart, is a graphical or tabular representation used in statistics and data analysis to assess the performance of a binary classification model, typically in the context of credit scoring, risk assessment, and other similar applications. The KS statistic is a measure of the discrimination or separation power of a model in distinguishing between two classes (usually positive and negative) in a binary classification problem.

Here's how a KS table is typically constructed and interpreted:

Data Preparation: To create a KS table, you start with a dataset that includes a binary outcome variable (e.g., 0 for non-default and 1 for default) and a predicted probability or a score generated by a classification model for each observation. The dataset is then sorted in descending order based on these scores or probabilities.

Binning or Percentiles: The sorted dataset is divided into several bins or percentiles. Each bin typically represents a range of predicted probabilities or scores. The number of bins can vary, but common choices include deciles (10 bins) or quartiles (4 bins).

Calculating Metrics: For each bin, you calculate several metrics, including:1)The cumulative percentage of positive cases (defaults) up to that bin.2)The cumulative percentage of negative cases (non-defaults) up to that bin.3)The difference between the cumulative percentages of positive and negative cases, which is the KS statistic for that bin.

Plotting or Tabulating: The KS statistics for each bin are plotted on a graph or tabulated in a table. The x-axis typically represents the bins, while the y-axis represents the KS statistic. The KS chart often shows a curve that illustrates how the KS statistic varies across the bins.

Interpretation: The KS statistic measures the maximum difference between the cumulative percentages of positive and negative cases. The higher the KS statistic, the better the model's ability to discriminate between the two classes. In practical terms, a higher KS statistic indicates that the model is better at ranking or sorting individuals based on their likelihood of belonging to the positive class.

Selecting a Threshold: The KS table helps in selecting an optimal threshold for classification. Typically, the threshold is set at the point where the KS statistic reaches its maximum value, which corresponds to the point on the chart or in the table where the largest gap between the cumulative percentages occurs.

KS tables are particularly useful in scenarios where it's crucial to prioritize one class over the other, such as in credit scoring, where identifying individuals at a higher risk of default is a priority. By using the KS table and selecting an appropriate threshold, organizations can make informed decisions about risk management and classification.

European Economic Letters
ISSN 2323-5233
Vol 14, Issue 2 (2024)
http://eelet.org.uk

Data Analysis and Result

A) Acquisition: - During the acquisition phase there are three steps as follows Fig(2).Machine Learning model is used to identify which of the lead will be converted into customer.

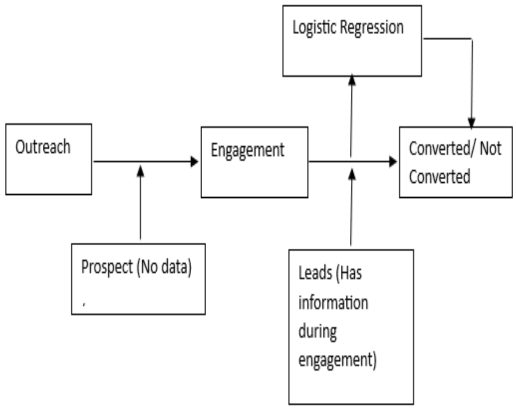

Fig.2: Flow of the acquisition steps and ML model implementation

The steps of identifying the probable lead conversion are as follows

a) Raw data is collected from customer support team of the luxury jewellery brand.
b) Data processing is done on the raw data using feature selection, Null Value treatment
c) Logistic Regression is applied on the processed data and Kolmogorov-Smirnov (KS) table is built by Table .1

Table.1 : Kolmogorov-Smirnov table to identify the class of Non Converted/Converted

	Counts	Events	non-Events	Expected Event Rate	Actual Event Rate	Cumm Events	Cumm non-Events	Cumm Event Rate	% Events	% non-Events	KS Measure
1	15,000	3,591	11,409	19.06%	23.94%	3,591	11,409	23.94%	35.82%	8.15%	27.67%
2	15,000	2,367	12,633	12.81%	15.78%	5,958	24,042	[illegible]	59.43%	17.18%	42.25%
3	15,000	1,361	13,639	10.12%	9.07%	7,319	[illegible]	16.26%	[illegible]	26.92%	46.08%
[illegible]	15,000	816	14,184	7.73%	5.44%	8,135	[illegible]	[illegible]	[illegible]	37.05%	44.09%
5	15,000	575	14,425	5.80%	3.83%	8,710	66,290	11.61%	86.87%	47.36%	39.52%
6	15,000	419	14,581	4.34%	2.79%	9,129	80,871	10.14%	91.05%	57.78%	33.28%
7	15,000	289	14,711	3.14%	1.93%	9,418	95,582	8.97%	93.94%	68.29%	25.65%
8	15,000	240	14,760	2.11%	1.60%	9,658	110,342	8.05%	96.33%	78.83%	17.50%
9	15,000	207	14,793	1.25%	1.38%	9,865	[illegible]	[illegible]	[illegible]	89.40%	9.00%
10	15,000	161	14,839	0.49%	1.07%	10,026	[illegible]	6.68%	[illegible]	100.00%	0.00%
	150,000	10,026	139,974	6.68%	6.68%						

Source: Using Author's Calculation

From Table.1 ,customer residing in first three rows will be considered as "Non Converted". Without using the logistic regression algorithm , out of 30% of the "Leads Not Converted" at random, only 6.68% of the events are identified . With using the random forest algorithm, out of the 30% of "Leads Not Converted" , 16.26% of the events will be identified. So the acquisition cost will be properly implemented and ROI value is much more better. This is called Lift.

B) Cross-Sell

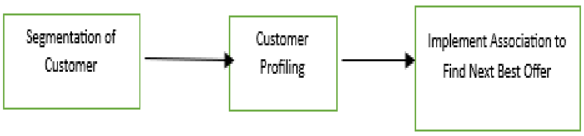

Fig.3 Flow for identifying Upsell Product

The steps of creating the segmentation along with customer profiling and association rule mapping is done in the following way.

a) First of all, number of optimal clusters is determined by Elbow Method, which is three here Fig.4
b) By using k-means algorithm, three clusters are created Cluster-1, Cluster-2, Cluster-3
c) Based on the different population mean, cluster will be marked as Classic, Premium and Platinum respectively as in Table.2
d) Now for each cluster is applied with Association Rule Mining to find the rule for Next Best Offer. In present case confidence threshold =75% and support threshold=20%.
   Table.5 will represent the next best offer for "Platinum" customer.

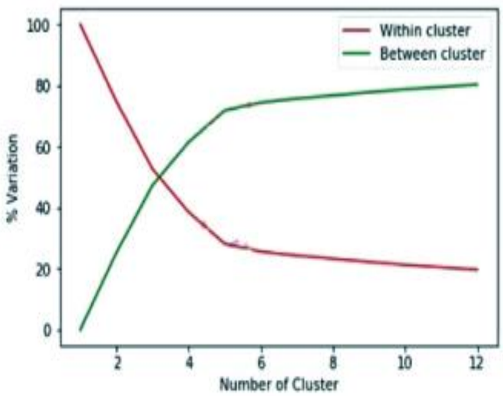

Fig.4 Finding cluster by Elbow Method

Table.2: Customer profiling chart for three cluster

Cluster	Age	income	jame&jewellary_sale	watch_sale	purchase_active_month	frequency
Cluster-1	165	217	81	113	190	201
Cluster-2	87	345	112	178	114	223
Cluster-3	40	567	213	65	167	311

High	>160% of population mean
Medium	85%-115% of population mean
Low	<70% of population mean

Source: Using Author's Calculation

Table.3 Association Rule Mapping for "Premium" segment customers

Association Rule(A->B)	Support(A)	Support(B)	Confidence	Lift
Purchase Gold above 1lakhs--> 20% discount on Diamond Purchase	30%	35%	80%	70%
Total purchase value more than 2 lakhs--> 10% discount in Watch	25%	34%	70%	60%
Diamond purchase of Bridal Collection--> honeymoon Ticket for Couple	22%	25%	82%	75%

Source: Using Author's Calculation

Conclusion:

In today's dynamic landscape of shifting consumer behaviours, the cornerstone of success lies in personalized customer care. Customers now view themselves as beneficiaries of such tailored attention, fostering enduring relationships with businesses. This paper primarily concentrates on elevating customer lifecycle management, offering insights that can be applied both in theory and practice. These findings can prove invaluable to hypermarkets and e-commerce stores looking to refine their customer segmentation strategies. Moreover, the relevance of this study transcends the retail sector, extending its applicability to industries such as tourism and finance, adapting to the unique spending behaviours within each.

Reference

1) Investigating the Impact of CRM Resources on CRM Processes: a Customer Life-Cycle Based Approach in the Case of a Greek Bank
https://www.sciencedirect.com/science/article/pii/S2212567115000313
2) Keeping Track of Customer Life Cycle to Build Customer Relationship https://link.springer.com/chapter/10.1007/11811305_41
3) Customer Life Cycle Management- Time and Beyond…- Expertise recognized by clients, Analysts like. https://www.researchgate.net/publication/272984563_Customer_Life_Cycle_Management-_Time_and_Beyond-_Expertise_recognized_by_clients_Analysts_like
4) Medha Gore J." Customer Life Cycle Management- Time and Beyond…- Expertise recognized by clients, Analysts like." IOSR Journal of Business and Management (IOSR-JBM) e-ISSN: 2278-487X, p-ISSN: 2319-7668. Volume 12, Issue 6 (Sep. - Oct. 2013), PP 78-82
5) Keramati, A., Mehrabi, H. and Mojir, N., 2010. "A process-oriented perspective on customer relationship management and organizational performance: An empirical investigation." Industrial Marketing Management, Vol. 39, No. 7, pp. 1170–1185.
6) A. Martínez, C. Schmuck, S. Pereverzyev, C. Pirker, M. Haltmeier, A machine learning framework for customer purchase prediction in the non-contractualsetting, Eur. J. Oper. Res. 281 (2020) 588–596.
7) A.M. Estrella-Ramon, ´ M. Sanchez-P ´ ´ erez, G. Swinnen, K. VanHoof, A marketing view of the customer value: customer lifetime value and customer equity, South Afr. J. Bus. Manag. 44 (2013) 47–64.
8) Srinivasan, R. and Moorman, C., 2005. Strategic Firm Commitments and Rewards for Customer Relationship Management in Online Retailing. Journal of Marketing, Vol. 69, No. 4, pp.193–200
9) Zineldin, M., 2005. Quality and customer relationship management (CRM) as competitive strategy in the Swedish banking industry. The TQM Magazine, Vol. 17, No. 4, pp.329-344.
10) Durkin, M.G and Howcroft, B. , 2003. Relationship marketing in the banking sector. Marketing Intelligence and Planning, Vol.21, No. 1, pp. 61-71
11) Exploring Novel Techniques to Detect Aberration from Metal Surfaces in Automobile Industries - Mohapatra, D., Chakraborty, A., Shaw, A.K. (2021). Exploring Novel Techniques to Detect Aberration from Metal Surfaces in Automobile Industries. In: Sabut, S.K., Ray, A.K., Pati, B., Acharya, U.R. (eds) Proceedings of International Conference on Communication, Circuits, and Systems. Lecture Notes in Electrical Engineering, vol 728. Springer, Singapore. https://doi.org/10.1007/978-981-33-4866-0_5
12) A, Chakraborty, A.k, Shaw, Scalable IoT Solution using Cloud Services – An Automobile Industry Use Case - https://doi.org/10.1109/I-SMAC49090.2020.9243544
13) Chakraborty, A., Shaw, A.K. and Samanta, S. (2022). On a Reference Architecture to Build Deep-Q Learning-Based Intelligent IoT Edge Solutions. In Convergence of Deep Learning In Cyber-IoT Systems and Security (eds R. Chakraborty, A. Ghosh, J.K. Mandal and S. Balamurugan). https://doi.org/10.1002/9781119857686.ch6