International Journal of Computer (IJC)

ISSN 2307-4523 (Print & Online)

https://ijcjournal.org/index.php/InternationalJournalOfComputer/index

Using Machine Learning to Identify Strategic Brand Growth Points

Daria Malykina*

Lead Generation and Marketing Manager, Roofworx Inc,San Francisco, California, USA

Email:antonbelousov17@icloud.com

Abstract

The study presents a systematic arrangement and theoretical decomposition of machine-learning (ML) approaches with the aim of identifying and testing a brand's key growth levers. Its objective is to construct a unified conceptual framework that integrates diverse ML algorithms into a closed analytical loop for processing market data, uncovering latent insights, and forecasting opportunities for brand expansion. The methodological foundation is built on an analysis and synthesis of leading publications in predictive analytics, natural language processing (NLP), and clustering techniques applied to marketing. The outcome is a multi-layer architecture that enables a staged progression from raw-data collection and aggregation to the formulation of growth hypotheses and their virtual validation. Scientific novelty lies in the description of a framework that eliminates fragmented ML usage in brand management by embedding these techniques into a single strategic process. The results are expected to benefit other researchers as well as strategy and marketing directors seeking to adopt data-driven approaches to brand governance.

Keywords: machine learning; brand growth; strategic marketing; predictive analytics; customer analytics; natural language processing; brand positioning; customization; customer lifetime value (CLV); data-driven marketing.

Received: 1/30/2026
Accepted: 3/30/2026
Published: 4/6/2026

----

* *Corresponding author.*

1.Introduction

The contemporary economic paradigm is undergoing rapid transformations driven by the synergy of technological progress and the evolution of consumer behavior. Under these conditions, a brand's ability to adapt swiftly and to identify new vectors of development in a timely manner becomes a decisive factor in its long-term resilience and competitiveness. Traditional approaches to strategic analysis—relying primarily on retrospective data and expert judgment—are increasingly inadequate in the face of high market volatility and the exponential growth of available information (Big Data). According to industry reports, the global market for artificial intelligence-based marketing solutions was valued at USD 20.44 billion in 2024 and is projected to reach USD 82.23 billion by 2030, expanding at an average annual rate of 25.0 percent between 2025 and 2030. Critical drivers underpinning this acceleration encompass the pervasive integration of machine learning and artificial intelligence to amplify social-media interaction metrics, deliver highly individualized consumer journeys, and propel the rapid growth of e-commerce transactions [1]. Such developments mark a profound epistemological shift—from reliance on experiential intuition to a rigorously data-driven managerial paradigm. Yet, the contemporary application of machine-learning techniques within strategic brand management remains disjointed and largely constrained to tactical use cases, highlighting a pressing absence of an overarching, systemic methodology that can harmonize varied ML approaches to extract strategic (as opposed to merely operational) growth imperatives.

Accordingly, the present study seeks to formulate and validate a unified conceptual architecture that embeds a heterogeneous suite of machine-learning models into a closed-loop analytical ecosystem. This framework is designed to methodically ingest and preprocess market data, surface latent patterns and insights through advanced algorithmic analyses, and generate robust forecasts of strategic expansion opportunities for brands.

The scientific contribution lies in describing a framework that eliminates the fragmented use of machine-learning tools in brand management and consolidates them into a coherent strategic procedure.

The author's hypothesis posits that the systematic application of machine learning in strategic planning not only improves forecast accuracy and decision validity but also identifies latent, counter-intuitive growth points inaccessible to traditional analytical instruments. This approach enables a transition from reactive to proactive brand management, creates a durable competitive advantage, and bridges the gap between the theoretical potential of machine learning and its practical implementation in high-level brand-management tasks.

2.Materials and methods

n contemporary literature on the application of machine learning techniques to identify strategic brand growth opportunities, sources may be grouped thematically into several categories: market and trend review studies; customer segmentation and churn prediction methods; customer lifetime value estimation and recommendation models; textual and visual data analysis for brand image construction; dynamic pricing; and agent-based modeling of business processes.

The first category comprises market review reports and studies that establish the overall trajectory of artificial intelligence in marketing. For example, the Grand View Research report provides an assessment of the AI solutions market in marketing and projects its growth over the coming years [1].

Customer segmentation and churn prediction methods encompass both classical algorithms and deep neural network approaches. Saha L. and his colleagues [2] developed a deep model for churn prediction in the telecommunications sector, achieving high accuracy by incorporating temporal dependencies in subscriber behavior. Tabianan K., Velu S., Ravi V. [3] applied the K-means algorithm for intelligent customer segmentation based on purchase behavior data, which enabled the identification of groups with distinct consumption patterns and the targeted adaptation of marketing strategies. The review by De S., Prabu P. [8] covers a wide range of churn prediction algorithms and highlights key factors influencing model performance, including feature selection, class balancing, and evaluation techniques.

A separate category includes studies focusing on customer lifetime value (CLV) estimation and recommendation systems. Sun Y., Liu H., Gao Y. [9] proposed an ML-based CLV analysis model combined with CRM analytics, enhancing the accuracy of long-term revenue predictions and optimizing retention investment. El-Shaer E. S., McKee G. T., Hamdy A. [6] introduced a hybrid LSTM-EMPG model for next-basket recommendation in e-commerce, demonstrating high relevance of suggested products by accounting for both purchase sequence patterns and product description embeddings. Textual and visual information processing methods play a critical role in brand image analysis and positioning, utilizing social media content and customer reviews. Alzate M., Arce-Urriza M., Cebollada J. [4] performed text mining on online reviews, employing topic modeling and sentiment analysis to identify key brand attributes and assess their impact on consumer perception. Davoodi L. [10] applied BERT-based aspect extraction methods to reviews, enabling a more detailed understanding of user preferences and improving the targeting of marketing messages. Ma R., Furuya K. [5], in a systematic review, described the application of computer vision techniques to social media images for landscape analysis, suggesting the potential transfer of these approaches to the visual content analysis of brands and their environments.Dynamic pricing in competitive environments was investigated by Kastius A., Schlosser R. [7] , who utilized reinforcement learning methods to adapt pricing strategies in real time, considering competitor reactions and demand elasticity.

Finally, the review by Onggo B. S., Foramitti J. [11] addresses agent-based modeling and simulation of business processes, including marketing campaigns and consumer interactions, which facilitates the exploration of complex interaction systems at a macro level and the forecasting of the effects of strategic decisions.

Despite the abundance of research, the literature exhibits certain contradictions. Cluster-based segmentation methods (e.g., K-means) show limited adaptability to evolving customer behavior, whereas deep models (LSTM, neural networks) demand substantial computational resources and large datasets, hindering implementation in small and medium-sized enterprises. Research on review and visual content analysis tends to treat text and images separately, with multimodal approaches receiving scant attention. Moreover, although agent-based modeling provides a broad view of systemic effects, it often lacks detailed integration with specific customer data and actual brand economic metrics. Therefore, there remains a need for studies that integrate multimodal data (text, images, transactions) and for adaptive machine learning models capable of operating effectively under resource constraints.

3.Results and Discussion

Based on the gap revealed in the review studies—namely, the absence of a coherent methodology for integrating machine learning into strategic brand management—a four-level conceptual model has been developed to identify brand growth drivers (see Figure 1). This model (hereafter, the Model) unifies diverse machine-learning algorithms within a single step-by-step framework designed to discover promising directions, verify them, and rank strategic initiatives according to their significance.

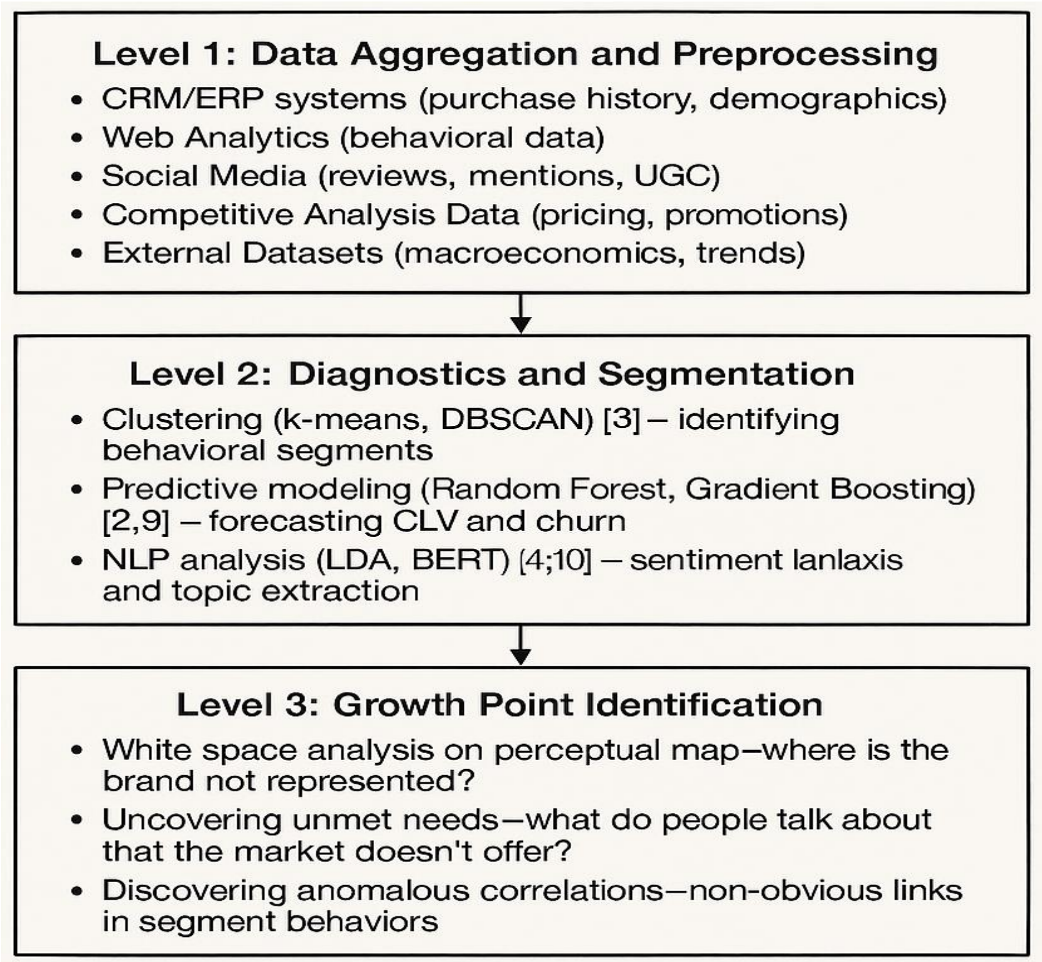

Figure 1: Conceptual model for identifying strategic points of brand growth based on machine learning (compiled by the author based on [2, 3, 4, 9, 10])

As illustrated in Figure 1, the first level entails data collection and preparation. During this stage, information is integrated from heterogeneous sources—internal (CRM systems, ERP platforms, transactional logs) and external (social networks, marketplace reviews, news feeds, and market-trend reports). The objective is not merely aggregation but the assurance of high-quality data by removing noise, imputing missing values, and standardizing all records to a uniform format [11].

Second level—Diagnostics and Segmentation. At this stage, key machine-learning methods come into play:

1. Dynamic clustering. Instead of the classical division by socio-demographic characteristics, clustering algorithms (for example, k-means or the more flexible DBSCAN, which can detect clusters of arbitrary shape) group customers according to their actual behavior: purchase frequency and monetary value (RFM analysis), viewed items, and activity over time. This approach makes it possible to isolate segments such as “loyal customers with a high average transaction value,” “inactive customers,” or “discount seekers” Reference [3].
2. Predictive modeling. On the basis of the generated clusters, predictive models are constructed for key business metrics. Classification algorithms such as Gradient Boosting Machines (GBM) and Random Forests accurately estimate the likelihood of customer churn in the near term or customer lifetime value (CLV) [5, 9]. As shown in Table 1, modern ensemble methods substantially outperform traditional logistic regression in churn-prediction tasks.
3. Unstructured data analysis (NLP). Transformer-based models such as BERT (Bidirectional Encoder Representations from Transformers) or topic modeling (LDA) are applied to textual streams (reviews, user messages, articles). These techniques automatically identify the main topics of discussion concerning the brand and its competitors, determine sentiment (positive, negative, or neutral), and, crucially, extract specific aspects that provoke a given reaction (for example, “long delivery time,” “inconvenient packaging,” or “overpriced”) [8, 10].

Table 1: Comparative accuracy of machine learning models in customer churn prediction (compiled by the author based on analysis [2, 6, 9])

Model	**Advantages**	**Disadvantages**
Logistic regression	Interpretability; speed	Low accuracy on nonlinear data
Random Forest	High accuracy; robustness to outliers	Less interpretable (“black box”)
Gradient Boosting Machine (GBM)	Highest accuracy; flexibility	Requires hyperparameter tuning; sensitive to noise
Multi-layer Perceptron (MLP)	Ability to capture complex relationships	Requires large datasets; computationally intensive

Third level – identification of promising development areas. At this stage, which underpins the scientific novelty of the proposed model, strategic hypotheses are formulated. Drawing on the results obtained at the second level, the search focuses on “white spaces” and anomalies that can serve as departure points for further research:

• Construction of a multidimensional perceptual map: by integrating data from NLP analyses of reviews concerning the focal brand and its competitors, a consumer perceptual map is created. Promising zones are delineated where high demand—frequently mentioned desired attributes—coincides with minimal competitive activity. For instance, if customers systematically request eco-friendly packaging while key market players ignore this factor, the finding unequivocally indicates a growth opportunity.

• Detection of unmet needs: an in-depth examination of negative reviews, not only of the focal products but also across the entire category, reveals recurring market issues. When a substantial share of users complain about the same deficiency in all existing offerings, a direct impetus arises to develop a solution capable of eliminating the identified pain point.

• Identification of atypical correlations: machine-learning techniques make it possible to capture statistically significant yet intuitively non-obvious associations. It may be discovered, for example, that the segment of young mothers purchasing baby food is highly likely to respond to subscriptions for meditation services. Conventional analytical approaches seldom uncover such relationships, whereas for an ML model this constitutes a regularity that can be exploited in cross-promotion or diversification.

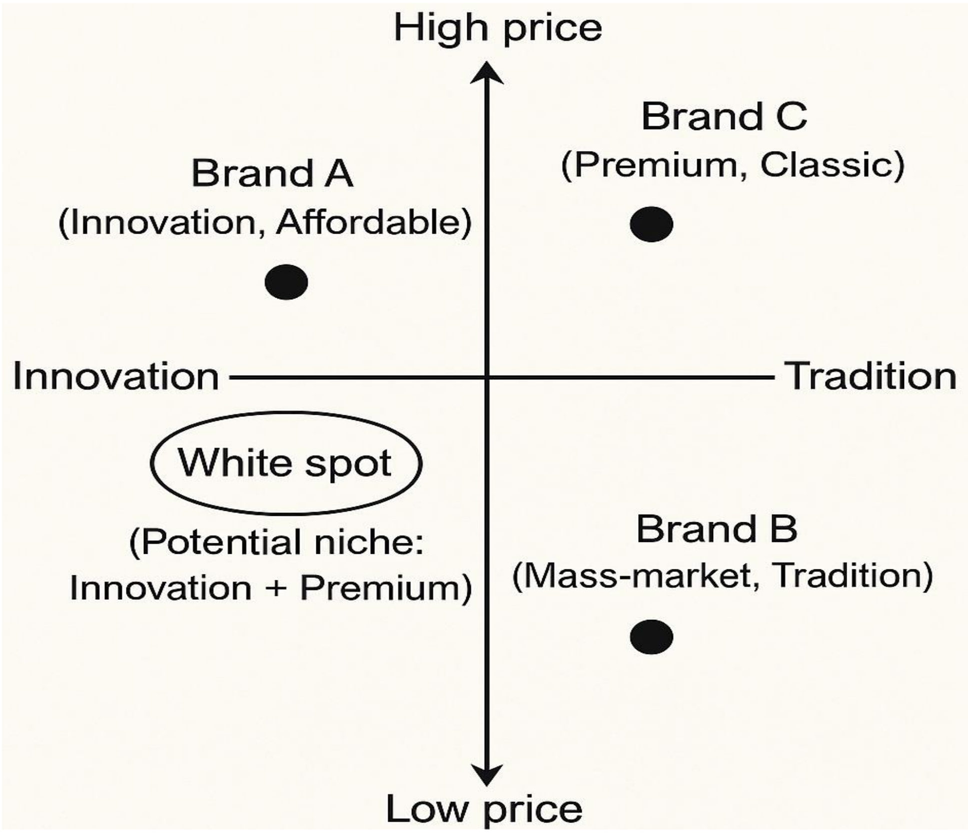

Figure 2: An example of a visual brand positioning map based on NLP analysis (compiled by the author based on the analysis of [3, 4, 7])

Fourth Level — Validation and Prioritization. Before the hypotheses generated at the third level can be scaled up, they require comprehensive verification.

- Micro-testing. Assumptions are examined on small yet statistically representative samples drawn from the segments identified at the second level. For example, demand for a new product flavor can be assessed through targeted advertising that directs the “experimenter” segment to a landing page offering pre-orders. This approach yields actual conversion rates while requiring minimal investment.
- Agent-based simulation. To evaluate complex strategic changes (e.g., adjustments to pricing policy), agent-based models are employed. Within a virtual market environment, the “agents” — digital

counterparts of the real segments — interact with one another and with the brand, enabling prediction of how modifying a single parameter (price) affects the entire ecosystem, including market share, profitability, and related indicators [11].

- Ranking. Initiatives that have passed validation are ordered according to composite metrics: anticipated economic impact (projected ROI, incremental CLV), implementation cost, and alignment with the brand's long-term strategy. This provides management with a transparent foundation for selecting the highest-priority growth opportunities.

When evaluating the proposed model, several key advantages emerge in comparison with traditional approaches. First, objectivity and scalability are achieved: decisions rest upon millions of data rows rather than intuitive judgments or focus groups comprising a handful of participants. Second, speed is dramatically increased; processes that once required months of manual analysis are now completed within hours. Third, the model can detect nonlinear relationships, opening a path to genuine innovation instead of merely local optimization. Nevertheless, implementation entails a series of challenges: stringent requirements for data volume and quality, the need for highly paid analytic specialists and substantial computational resources, and the interpretability issues posed by certain "black-box" algorithms. Overcoming these barriers is particularly important for a successful data-driven transformation of brand management.

The study's findings therefore indicate that machine learning is no longer confined to tactical tools but can become the core of strategic planning. The proposed four-level model provides a practical framework for this integration, turning algorithm use from a fragmented activity into a cyclical, continuous process in which each stage builds on the results of the previous one and enables the ongoing discovery and validation of brand growth opportunities.

4.Conclusion

The research successfully achieved its stated aim by systematising methodological approaches to the application of machine learning and formulating a conceptual model capable of identifying strategically significant areas of brand growth. Analysis of recent academic literature confirmed the urgency of the problem and revealed a gap between the situational use of individual ML algorithms and the absence of their holistic integration into long-term strategic planning.The proposed framework unfolds as a four-layered schema, in which each stratum systematically advances from raw data consolidation through diagnostic scrutiny, onto the formulation of hypotheses about potential vectors for expansion, and culminates in rigorous empirical validation. Rather than treating techniques like clustering algorithms, predictive analytics and natural-language processing as isolated instruments, the architecture integrates them within a cohesive strategic-management scaffolding, orienting every analytical step toward the unearthing of hidden market niches and unaddressed consumer needs. The empirical results affirm the working proposition that a holistic fusion of machine-learning methodologies can illuminate non-obvious growth trajectories, thereby enabling a brand to adopt an anticipatory stance and secure a durable competitive edge.From a practical standpoint, this blueprint offers organizations a methodical pathway for embedding data-centric practices into their brand governance processes. Its chief merits—unbiased evaluation, expedited insight generation and substantive analytical depth—nevertheless coexist with constraints such as the dependence on high-fidelity input data and the requisite availability of specialized technical expertise. Future research should therefore aim at enhancing the explainability of complex ML constructs for decision-makers, as well as at adapting the model's resource demands to the operational capacities of small and medium enterprises.

References

[1]. Artificial Intelligence In Marketing Market Size. [Electronic resource]. – Access mode: https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-marketing-market-report (date of access: 15.05.2025).

[2]. Saha L. et al. Deep churn prediction method for telecommunication industry //Sustainability. – 2023. – Vol. 15 (5). https://doi.org/10.3390/su15054543.

[3]. Tabianan K., Velu S., Ravi V. K-means clustering approach for intelligent customer segmentation using customer purchase behavior data //Sustainability. – 2022. – Vol. 14 (12). https://doi.org/10.3390/su14127243.

[4]. Alzate M., Arce-Urriza M., Cebollada J. Mining the text of online consumer reviews to analyze brand image and brand positioning //Journal of Retailing and Consumer Services. – 2022. – Vol. 67. https://doi.org/10.1016/j.jretconser.2022.102989.

[5]. Ma R., Furuya K. Social Media Image and Computer Vision Method Application in Landscape Studies: A Systematic Literature Review //Land. – 2024. – Vol. 13 (2). https://doi.org/10.3390/land13020181.

[6]. El-Shaer E. S., McKee G. T., Hamdy A. The LSTM-EMPG Model for Next Basket Recommendation in E-commerce //International Journal of Information. – 2024. – Vol. 9 (1). – pp. 9-23.

[7]. Kastius A., Schlosser R. Dynamic pricing under competition using reinforcement learning //Journal of Revenue and Pricing Management. – 2022. – Vol. 21 (1). – pp. 50-63.

[8]. De S., Prabu P. Predicting customer churn: A systematic literature review //Journal of Discrete Mathematical Sciences and Cryptography. – 2022. – Vol. 25 (7). – pp. 1965-1985. https://doi.org/10.1080/09720529.2022.2133238.

[9]. Sun Y., Liu H., Gao Y. Research on customer lifetime value based on machine learning algorithms and customer relationship management analysis model //Heliyon. – 2023. – Vol. 9 (2). - pp.1-16.

[10]. Davoodi L. Enhancing the understanding of e-commerce reviews through aspect extraction techniques: A BERT-based approach //36th Bled eConference digital economy and society: The balancing act for digital innovation in times of instability. – 2023. - pp.1-17.

[11]. Onggo B. S., Foramitti J. Agent-based modeling and simulation for business and management: a review and tutorial //2021 Winter Simulation Conference (WSC). – IEEE, 2021. – pp. 1-15.