



The Influence of Brand Community Participation and User-Generated Content on Customer-Based Brand Equity Repurchase Intention and Electronic Word-of-Mouth in Direct-to-Consumer Markets

Samuel Bradley Doe Gadekah ^{1*}, Joy Onma Enyejo ²

¹ Department of Economics and Business, TU Bergakademie Freiberg, Freiberg, Germany

² Department of Business Administration, Nasarawa State University, Keffi, Nasarawa State, Nigeria

* Corresponding Author: Samuel Bradley Doe Gadekah

Article Info

ISSN (online): 2583-6641

Volume: 04

Issue: 03

Received: 24-03-2025

Accepted: 26-04-2025

Published: 28-05-2025

Page No: 204-224

Abstract

The rapid expansion of direct-to-consumer (DTC) markets has transformed brand community participation and user-generated content (UGC) into important mechanisms influencing customer-based brand equity (CBBE), repurchase intention, and electronic word-of-mouth (eWOM). However, conventional analytical approaches frequently examine these relationships through linear structural models, potentially overlooking nonlinear interactions, heterogeneous consumer responses, and complex mediation pathways. This study proposes a novel Brand Community and User-Generated Content Equity Prediction Algorithm (BCUGC-EPA), an explainable hybrid analytical framework designed to quantify and predict the interconnected effects of brand community engagement and UGC characteristics on brand equity formation and post-purchase behavioural intentions. The proposed algorithm integrates Partial Least Squares Structural Equation Modelling (PLS-SEM), Extreme Gradient Boosting (XGBoost), and Shapley Additive Explanations (SHAP) within a unified predictive architecture. PLS-SEM estimates direct, indirect, and mediating relationships among community participation, UGC credibility, perceived authenticity, customer-based brand equity, repurchase intention, and eWOM, while XGBoost captures nonlinear dependencies and higher-order interactions. SHAP provides interpretable estimates of variable contributions to consumer behaviour predictions. The study adopts a quantitative research design using structured questionnaires administered to customers of DTC brands, supplemented, where available and consented to, by anonymised community engagement indicators and UGC interaction records. Model evaluation will compare BCUGC-EPA against conventional PLS-SEM prediction, Multiple Linear Regression, Random Forest, Support Vector Regression, and standalone XGBoost under identical data partitions and validation procedures. Predictive performance will be evaluated using the coefficient of determination (R^2), root mean squared error (RMSE), mean absolute error (MAE), and cross-validated predictive stability. Statistical significance will be examined through bootstrapped path coefficients, mediation effects, and confidence intervals. Comparative results will be presented through predictive performance bar charts, observed-versus-predicted scatterplots, SHAP importance diagrams, mediation path diagrams, and interaction response surfaces. The central hypothesis is that integrating theory-driven structural relationships with nonlinear machine learning will improve predictive accuracy while preserving meaningful interpretation of how community participation and UGC influence consumer outcomes. The study is designed to determine whether customer-based brand equity mediates the relationships between brand community participation, UGC characteristics, repurchase intention, and eWOM, and whether these pathways differ across consumer engagement levels. Its anticipated contribution is an empirically testable framework for explaining and predicting customer retention, brand advocacy, and equity development in DTC markets, offering actionable evidence for optimising community-driven marketing strategies and consumer relationship management.

Keywords: Brand Community Participation, User-Generated Content, Customer-Based Brand Equity, Repurchase Intention, Electronic Word-of-Mouth

1. Introduction

1.1. Background to Brand Community Participation and User-Generated Content in Direct-to-Consumer Markets

Direct-to-consumer (DTC) markets have transformed brand management from a predominantly transaction-oriented activity into a digitally mediated system of continuous consumer interaction, collaborative value creation, and relationship development. Unlike conventional distribution arrangements, DTC businesses maintain direct access to customer interfaces, enabling greater

control over brand communication, community management, and post-purchase engagement. Within this environment, brand community participation encompasses consumer interactions involving discussions, peer assistance, product recommendations, feedback provision, and collaborative content creation. User-generated content (UGC), including product reviews, testimonials, photographs, instructional videos, and social media commentary, provides information that can shape consumers' assessments of product quality and brand credibility. Wu *et al.* (2018) established that engagement within online brand communities is associated with subsequent review-generation behaviour and positive online evaluations. These relationships suggest that participation should be examined as a multidimensional construct rather than measured exclusively through community membership or interaction frequency. Consequently, the present research conceptualises participation through cognitive engagement, social interaction, behavioural contribution, and community identification, while treating UGC credibility, information quality, and perceived authenticity as separate explanatory dimensions.

The strategic relevance of these constructs extends beyond communication effectiveness to the development of customer-based brand equity (CBBE), repurchase intention, and electronic word-of-mouth (eWOM). Zollo *et al.* (2020) demonstrated the importance of consumer benefits and brand experience in explaining relationships between social media marketing and brand equity, indicating that interaction alone does not necessarily generate favourable brand outcomes. Similarly, Ardley *et al.* (2020) identified value co-creation practices within online brand communities, providing support for examining consumer participation as a productive rather than merely receptive activity. Building on these findings, the present study introduces the proposed Brand Community and User-Generated Content Equity Prediction Algorithm (BCUGC-EPA), which combines partial least squares structural equation modelling (PLS-SEM), Extreme Gradient Boosting (XGBoost), and Shapley Additive Explanations (SHAP). This architecture is designed to estimate theoretically meaningful relationships while detecting nonlinear consumer response patterns. Its evaluation will establish whether combined structural and predictive analysis offers stronger out-of-sample performance than conventional statistical and machine learning models.

1.2. Evolution of Customer-Based Brand Equity in Digital and Community-Driven Marketing Environments

Customer-based brand equity has evolved from a construct traditionally associated with brand awareness, perceived quality, associations, and loyalty into a dynamic representation of consumer responses generated across interconnected digital environments. In DTC markets, brand equity develops through repeated encounters with branded content, consumer reviews, social interactions, product experiences, and community-generated evaluations. These interactions can influence consumers' familiarity with a brand, perceptions of product reliability, emotional attachment, and willingness to maintain commercial relationships. Koay *et al.* (2021) established that perceived social media marketing activities are associated with consumer-based brand equity through mechanisms involving

brand experience and co-creation behaviour. Their findings provide a basis for examining brand equity as a multidimensional outcome influenced by both company-controlled communication and consumer-led interactions. Accordingly, the present study operationalises CBBE through brand awareness, perceived quality, brand associations, and attitudinal loyalty, while distinguishing brand equity indicators from behavioural outcomes such as repurchase intention and eWOM. This separation is necessary to reduce construct overlap and preserve interpretability in subsequent mediation analysis.

The growing importance of interactive communities also requires brand equity models to account for consumer identification, engagement intensity, and heterogeneous behavioural responses. Kaur *et al.* (2020) demonstrated that brand community identification and rewards influence consumer brand engagement, which subsequently contributes to brand loyalty. However, the relationship between engagement and brand outcomes may depend on consumer characteristics and specific interaction patterns. Schivinski (2021) applied conditional inference trees to identify distinct consumer engagement patterns, demonstrating the analytical relevance of nonlinear classification and interaction effects in digital brand research. These findings support extending traditional linear brand equity frameworks with computational methods capable of identifying heterogeneous consumer responses. Within BCUGC-EPA, latent brand equity scores obtained through PLS-SEM will serve as theoretically informed inputs or prediction targets in appropriately separated modelling stages. XGBoost will subsequently evaluate nonlinear relationships involving community participation, UGC characteristics, and brand equity dimensions. The resulting framework will facilitate comparisons between structural explanatory performance and predictive accuracy without assuming that algorithmic complexity necessarily produces superior outcomes.

1.3. Brand Community Engagement and Its Influence on Consumer–Brand Relationships

Brand community engagement represents a multidimensional process through which consumers develop psychological, emotional, and behavioural connections with brands through sustained participation in digitally mediated communities. In DTC markets, these relationships are reinforced through customer-to-customer interactions, direct communication with brand representatives, shared product experiences, and collaborative problem-solving. Engagement extends beyond passive content consumption to include active participation, identification with community values, and voluntary advocacy. Kumar and Nayak (2018) demonstrated that customer brand engagement mediates relationships between brand community engagement and loyalty intentions, highlighting the importance of intermediate psychological mechanisms in converting community relationships into brand relationships. This suggests that community participation should not automatically be interpreted as evidence of loyalty, because behavioural engagement may reflect information seeking, social motivations, or temporary incentives. Accordingly, the proposed conceptual framework distinguishes participation frequency, interaction quality, community identification, and perceived community support. These indicators provide a foundation for estimating the relative importance of community-related experiences in

developing customer-based brand equity and shaping consumers' subsequent repurchase and recommendation intentions.

The development of sustainable consumer-brand relationships also depends on social capital, perceived psychological ownership, and the capacity of communities to maintain trust during unfavourable brand experiences. Chi *et al.* (2022) established that social capital contributes to customer brand engagement within online service brand communities, with collective psychological ownership and customer citizenship behaviour functioning as important mediating mechanisms. Yuan *et al.* (2020) further demonstrated that online brand community engagement can influence consumer responses during product-harm crises, particularly through forgiveness-related mechanisms associated with repurchase intention. These findings indicate that community engagement may generate different outcomes under favourable and adverse conditions. The proposed BCUGC-EPA framework therefore incorporates community identification, social interaction, and participation intensity as explanatory variables while evaluating customer-based brand equity as a potential intermediary mechanism. Interaction terms and nonlinear response functions will examine whether high participation strengthens or weakens relationships between UGC characteristics and consumer outcomes. Structural path estimates will be complemented by predictive comparisons to distinguish theoretically supported associations from variables that contribute most strongly to individual-level prediction.

1.4. User-Generated Content Credibility, Authenticity, and Consumer Behavioural Responses

User-generated content constitutes an influential component of digital consumer decision-making because it provides experiential information originating from individuals who have purchased, evaluated, or interacted with branded products. In DTC environments, consumers encounter UGC through customer reviews, social media posts, product demonstrations, community discussions, and visual testimonials. However, its persuasive effectiveness depends on informational credibility, perceived authenticity, argument quality, relevance, and source characteristics rather than content exposure alone. Daowd *et al.* (2021) demonstrated that source presentation, credibility, argument quality, and source similarity influence electronic word-of-mouth credibility and information adoption among younger consumers. These findings indicate that UGC credibility should be assessed through several measurable dimensions rather than represented by a single undifferentiated score. In the present study, credibility refers to the perceived accuracy and trustworthiness of consumer-generated information, while authenticity represents the perceived genuineness and independence of the contributor's experience. Information quality will capture completeness, specificity, diagnostic usefulness, and product relevance. These dimensions will be measured using validated questionnaire indicators and evaluated for discriminant validity before incorporation into the proposed predictive architecture.

Consumer responses to UGC also depend on interactions between message characteristics, perceived source reliability, and the social environment in which information is received. Onofrei *et al.* (2022) identified source and content factors as important mediating mechanisms connecting consumer-to-

consumer social media interactions with purchase intention and behavioural engagement. Similarly, Zhao *et al.* (2020) demonstrated that information quality is associated with consumer trust and purchase intention in social e-commerce settings, suggesting that informational value is an important mechanism through which peer-generated communication influences consumer evaluations. Nevertheless, content perceived as credible may not necessarily generate repurchase intention unless consumers associate the information with favourable brand experiences. The proposed BCUGC-EPA model therefore evaluates both direct relationships between UGC attributes and consumer outcomes and indirect pathways operating through customer-based brand equity. XGBoost will investigate possible nonlinear thresholds and interactions, while SHAP analysis will assess the contribution of individual UGC attributes to predictions. This arrangement allows the study to examine whether credibility, authenticity, or information quality has greater predictive relevance for different consumer outcomes without interpreting predictive importance as evidence of causality.

1.5. Repurchase Intention and Electronic Word-of-Mouth as Outcomes of Customer-Based Brand Equity

Repurchase intention and electronic word-of-mouth represent distinct but potentially interconnected behavioural outcomes of consumer-brand relationships in DTC markets. Repurchase intention reflects consumers' stated likelihood of purchasing products from the same brand again, whereas eWOM intention concerns their willingness to communicate product evaluations, recommendations, and experiences through digital channels. Although both outcomes may reflect favourable brand attitudes, they differ in their underlying motivations, perceived social consequences, and behavioural requirements. Tan (2024) examined relationships among customer participation, online brand community trust, positive eWOM intention, and repurchase intention, demonstrating that participation may operate through intermediary relational mechanisms rather than uniformly producing direct effects. This distinction is important because consumers may recommend a brand without intending an immediate repeat purchase, while habitual purchasers may rarely participate in digital advocacy. Within the present study, repurchase intention and eWOM intention will therefore be modelled as separate endogenous constructs, with customer-based brand equity positioned as a potential mediator between brand community participation, UGC characteristics, and behavioural intentions. Brand loyalty indicators will be carefully distinguished from repurchase intention to avoid artificially inflating structural relationships through overlapping measurement items.

Electronic word-of-mouth additionally influences consumer behaviour through credibility, perceived usefulness, message quality, and information adoption mechanisms. Ismagilova *et al.* (2020), through a meta-analysis of existing research, established that several eWOM characteristics are associated with purchase intention, although their relative importance differs across studies. Cheng *et al.* (2020) further demonstrated the relevance of online community characteristics and satisfaction-related processes in the development of customer loyalty. Together, these findings support examining how consumer experiences contribute to brand equity before translating into retention-oriented and

advocacy-oriented intentions. The BCUGC-EPA framework will estimate the structural relationships linking community participation and UGC attributes to customer-based brand equity, followed by separate prediction models for repurchase intention and eWOM intention. Bootstrapped indirect effects will assess the statistical evidence for mediation, while comparative predictive analysis will determine whether combining latent constructs with nonlinear machine learning improves out-of-sample estimation. The analysis will also investigate whether community participation and UGC credibility interact in shaping consumer outcomes, providing a basis for graphical representations of heterogeneous engagement patterns and their managerial implications.

1.6. Limitations of Conventional Statistical and Machine Learning Approaches to Consumer Behaviour Prediction

Conventional consumer behaviour research frequently employs multiple linear regression, covariance-based structural equation modelling, and partial least squares structural equation modelling to investigate relationships among marketing constructs. These approaches provide established procedures for estimating structural coefficients, assessing latent variables, and testing mediation hypotheses. Nevertheless, models specified exclusively through additive linear relationships may inadequately represent consumer behaviour when predictor effects are nonlinear, conditional, or heterogeneous across individuals. Shmueli *et al.* (2019) emphasised the importance of evaluating PLS-SEM models through out-of-sample predictive assessment rather than relying solely on explanatory indicators such as path significance and coefficients of determination. This distinction is particularly relevant to DTC markets, where consumers differ in community involvement, content preferences, previous experiences, and brand familiarity. Consequently, a statistically significant relationship between brand participation and customer-based brand equity does not establish that a model can accurately predict the responses of previously unobserved consumers. The proposed study addresses this concern by independently examining structural validity and predictive performance, ensuring that theoretical explanation and algorithmic accuracy are evaluated using appropriate assessment procedures.

Marketing analytics also faces difficulties arising from fragmented information systems, heterogeneous measurements, and the limited interpretability of complex predictive algorithms. Aluso (2021) discussed challenges surrounding cross-platform marketing data integration and forecasting, reinforcing the importance of consistent data structures, harmonised indicators, and reliable analytical pipelines. Ma and Sun (2020) further highlighted the predictive flexibility of machine learning while identifying interpretability as an important challenge in marketing applications. Standalone algorithms such as Random Forest, Support Vector Regression, and XGBoost can capture complex response functions, but they do not automatically estimate theoretically specified indirect effects or establish causal mechanisms. Conversely, conventional structural models may be insufficiently flexible when nonlinear interaction patterns are not explicitly specified. BCUGC-EPA is therefore proposed as a hybrid framework combining PLS-SEM measurement and mediation analysis with

XGBoost prediction and SHAP-based interpretation. Its performance will be assessed using identical training and testing partitions, nested hyperparameter optimisation, RMSE, MAE, R^2 , and bootstrap uncertainty intervals. The proposed integration is intended to preserve theoretical interpretability while evaluating whether nonlinear modelling produces measurable improvements over established predictive alternatives.

1.7. Problem Statement, Research Gap and Justification for the Proposed BCUGC-EPA Algorithm

Despite growing investment in brand communities and consumer-generated digital content, a central analytical problem remains: the relationships connecting community participation, UGC characteristics, customer-based brand equity, repurchase intention, and eWOM are frequently examined through separate theoretical or predictive frameworks. Consequently, limited integrated evidence is available concerning how these constructs jointly contribute to consumer outcomes, particularly when nonlinear interactions and heterogeneous engagement patterns are considered. Ngai and Wu (2022) identified the expanding application of machine learning across marketing activities and highlighted the need for frameworks capable of translating diverse consumer information into meaningful managerial insights. However, predictive modelling alone cannot establish whether brand equity serves as an intermediary relationship mechanism or determine the theoretical significance of individual marketing constructs. The problem is particularly relevant to DTC organisations, where community management, consumer-generated content, and repeat purchasing occur within closely connected digital environments. Without an integrated analytical architecture, managers may incorrectly prioritise high engagement volumes over credible content, meaningful community relationships, or brand equity development. This study therefore seeks to connect theory-driven measurement, structural relationship estimation, and nonlinear prediction within a coherent methodological framework.

A further research gap concerns the limited availability of directly comparable evidence evaluating explanatory validity, predictive accuracy, interpretability, and model robustness within a unified brand community and UGC framework. Herhausen *et al.* (2024) identified interpretability, responsible algorithmic deployment, and advanced predictive methods as important directions in contemporary marketing research. Volkmar *et al.* (2022) similarly examined the opportunities and implementation barriers associated with artificial intelligence and machine learning in marketing management. These contributions support methodological integration but do not establish that a specific hybrid algorithm will outperform all existing approaches. The present study consequently proposes BCUGC-EPA as a testable methodological contribution integrating PLS-SEM, XGBoost, and SHAP. Its proposed architecture will estimate direct and indirect structural relationships, identify nonlinear interactions, and generate interpretable predictions of brand equity, repurchase intention, and eWOM. Comparative experiments against PLS-SEM prediction, multiple linear regression, Random Forest, Support Vector Regression, and standalone XGBoost will assess whether the integration improves generalisation.

The resulting statistical tables, structural diagrams, feature-attribution plots, and predictive response surfaces will provide the empirical basis for evaluating the proposed algorithm's contribution rather than presupposing superior performance.

1.8. Research Objectives, Research Questions, and Hypotheses

The principal aim of this study is to develop and empirically evaluate the proposed BCUGC-EPA algorithm for examining how brand community participation and user-generated content influence customer-based brand equity, repurchase intention, and electronic word-of-mouth in DTC markets. The investigation is guided by the following objectives, research questions, and testable hypotheses, which collectively establish the theoretical relationships, statistical evaluation criteria, and comparative predictive requirements of the proposed framework.

Research Objectives

1. To assess the influence of brand community participation on customer-based brand equity among consumers in direct-to-consumer markets.
2. To examine how user-generated content credibility, authenticity, and information quality contribute to customer-based brand equity formation.
3. To determine the effects of customer-based brand equity on repurchase intention and electronic word-of-mouth intention.
4. To evaluate the mediating role of customer-based brand equity in relationships between community participation, user-generated content, and behavioural intentions.
5. To investigate nonlinear interactions between community participation and user-generated content characteristics in predicting consumer behavioural outcomes.
6. To develop the BCUGC-EPA algorithm by integrating PLS-SEM, XGBoost, and SHAP-based explanatory modelling techniques.
7. To compare the predictive performance of BCUGC-EPA against conventional statistical models and established machine learning algorithms.
8. To identify the most influential predictors of brand equity, repurchase intention, and electronic word-of-mouth through explainable machine learning analysis.

Research Questions

1. How does brand community participation influence customer-based brand equity in direct-to-consumer markets?
2. To what extent do user-generated content credibility, authenticity, and information quality influence customer-based brand equity?
3. How does customer-based brand equity relate to repurchase intention and electronic word-of-mouth intention?
4. Does customer-based brand equity mediate the relationships between community participation, user-generated content, and consumer behavioural intentions?
5. Do nonlinear interactions between community participation and user-generated content improve the prediction of brand equity and behavioural outcomes?
6. How effectively can BCUGC-EPA integrate structural

relationship estimation, nonlinear prediction, and explainable artificial intelligence?

7. Does BCUGC-EPA achieve lower prediction errors and greater predictive stability than conventional statistical and machine learning models?
8. Which brand community and user-generated content attributes contribute most strongly to the prediction of customer-based brand equity, repurchase intention, and electronic word-of-mouth?

Research Hypotheses

1. **H1:** Brand community participation is positively associated with customer-based brand equity in direct-to-consumer markets.
2. **H2:** User-generated content credibility, authenticity, and information quality are positively associated with customer-based brand equity.
3. **H3:** Customer-based brand equity is positively associated with repurchase intention and electronic word-of-mouth intention.
4. **H4:** Customer-based brand equity significantly mediates the relationships between community participation, user-generated content, and consumer behavioural intentions.
5. **H5:** Nonlinear interactions between community participation and user-generated content characteristics provide additional predictive information about brand equity and behavioural outcomes.
6. **H6:** The integrated BCUGC-EPA framework provides statistically significant incremental predictive information beyond its individual component models.
7. **H7:** BCUGC-EPA achieves significantly lower out-of-sample prediction errors than the specified conventional statistical and machine learning benchmarks.
8. **H8:** The relative predictive contributions of community participation and user-generated content attributes differ significantly across brand equity, repurchase intention, and electronic word-of-mouth outcomes.

1.9. Contributions and Significance of the Study

This study proposes theoretical, methodological, empirical, and managerial contributions to digital branding and consumer behaviour research by integrating brand community participation, user-generated content characteristics, customer-based brand equity, repurchase intention, and electronic word-of-mouth within a unified analytical framework. Theoretically, it extends existing consumer-brand relationship models by positioning customer-based brand equity as an intermediary construct connecting digital community experiences with subsequent behavioural intentions, while distinguishing the contributions of participation intensity, content credibility, perceived authenticity, and information quality. Methodologically, the proposed BCUGC-EPA algorithm combines PLS-SEM for latent variable measurement and structural mediation analysis, XGBoost for nonlinear predictive modelling, and SHAP for interpreting individual predictor contributions. The intended contribution is not simply the combination of existing computational methods, but an explicitly specified and empirically evaluated workflow that connects theoretical constructs to predictive features while controlling for data leakage and maintaining transparent benchmark comparisons. Empirically, the study will generate evidence concerning direct and indirect relationships, nonlinear interaction effects, variation in consumer responses, and the

predictive stability of alternative modelling approaches. These results will be communicated through structural path diagrams, numerical comparison tables, observed-versus-predicted scatterplots, SHAP importance plots, and nonlinear response surfaces. For DTC organisations, the findings are expected to inform evidence-based decisions concerning community investment, user-generated content management, customer retention strategies, and digital advocacy initiatives. The broader significance lies in developing an interpretable decision-support framework capable of distinguishing statistically supported consumer relationship mechanisms from predictive associations, thereby improving the methodological foundations of data-driven brand equity assessment.

1.10. Scope of the Review and Structure of the Paper

The study focuses on the relationships between brand community participation, user-generated content, customer-based brand equity, repurchase intention, and electronic word-of-mouth among consumers with experience purchasing from DTC brands and interacting with their digital marketing environments. Its empirical scope encompasses structured questionnaire data measuring community participation, UGC credibility, authenticity, information quality, brand awareness, perceived quality, brand associations, attitudinal loyalty, repurchase intention, and eWOM intention. Where appropriately authorised and available, supplementary engagement indicators may be incorporated to strengthen behavioural measurement and predictive analysis. The investigation is restricted to consumer-level relationship modelling and prediction rather than experimental causal identification or firm-level financial performance estimation. Methodologically, it employs PLS-SEM for construct measurement and structural analysis, XGBoost for nonlinear prediction, and SHAP for interpreting predictive contributions, with Multiple Linear Regression, Random Forest, Support Vector Regression, standalone XGBoost, and conventional PLS-SEM prediction serving as comparison approaches. The paper is organised into five principal sections. Section 1 introduces the research background, central problem, theoretical motivation, objectives, research questions, and hypotheses. Section 2 critically reviews the theoretical and empirical literature concerning digital brand communities, user-generated content, customer-based brand equity, consumer behavioural intentions, and computational prediction techniques. Section 3 describes the conceptual framework, research design, mathematical formulations, proposed BCUGC-EPA architecture, and validation procedures. Section 4 presents and discusses the empirical structural relationships, mediation estimates, benchmark performance comparisons, feature-attribution results, and nonlinear interaction analyses. Section 5 provides the conclusions, methodological contributions, strategic recommendations, study limitations,

and directions for future research.

2. Literature Review

2.1. Theoretical Foundations of Brand Communities, Social Identity, Consumer Engagement, and Customer-Based Brand Equity

The theoretical foundation of brand community participation integrates social identity theory, social exchange theory, and customer-based brand equity perspectives to explain how consumers develop enduring relationships with brands through collective interaction. Social identity theory suggests that consumers derive psychological value from identifying with communities whose shared interests, symbols, and values reinforce their self-concepts as represented in figure 1. Peters and Bodkin (2018) demonstrated that community characteristics vary across brand and retail environments, suggesting that membership alone cannot adequately represent the psychological dimensions of community identification. From a social exchange perspective, consumers participate when expected informational, relational, or emotional benefits justify their contributions. Kumar and Kumar (2020) established that experiential benefits, self-esteem benefits, and perceived community investment positively influence community engagement, commitment, and brand loyalty. These findings provide the theoretical basis for operationalising brand community participation through interaction frequency, identification, contribution intensity, and perceived community value. Within the proposed framework, these dimensions represent theoretically distinct antecedents of customer-based brand equity rather than interchangeable indicators of consumer loyalty.

Consumer engagement extends these theoretical relationships by explaining how repeated interactions generate brand knowledge, perceived quality, brand associations, and attitudinal commitment. Ononiwu *et al.* (2023) examined digital engagement through influencer marketing, customer relationship management analytics, and adaptive online technologies, providing a related perspective on how digital infrastructure supports customer interaction. Although this research does not directly establish brand community effects in DTC markets, it demonstrates the relevance of integrated digital engagement systems. The present study consequently conceptualises customer-based brand equity as a higher-order construct comprising brand awareness, perceived quality, brand associations, and attitudinal loyalty. BCUGC-EPA will evaluate whether community identification and participation influence these dimensions directly or through nonlinear engagement patterns. PLS-SEM will estimate structural associations and assess construct validity, while XGBoost and SHAP will examine heterogeneous predictive contributions. This theoretical integration supports empirical differentiation between social belonging, behavioural participation, and brand equity formation.

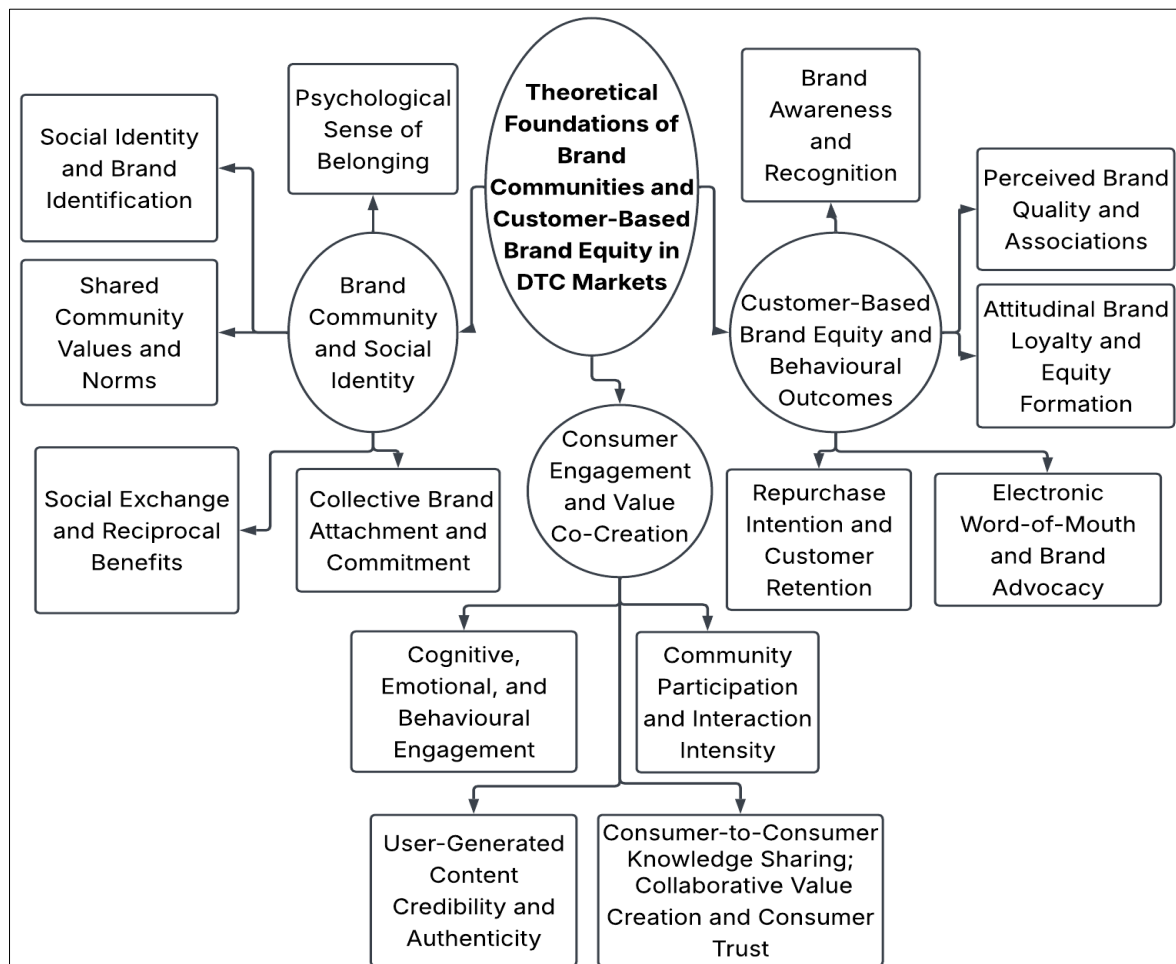


Fig 1: Theoretical framework linking brand community identity, consumer engagement, value co-creation, brand equity, and behavioural intentions.

Figure 1 illustrates the theoretical relationships among brand community participation, social identity, consumer engagement, value co-creation, and customer-based brand equity (CBBE) within direct-to-consumer (DTC) markets. The conceptual framework comprises three interconnected branches representing the psychological foundations, engagement mechanisms, and behavioural outcomes of consumer-brand relationships. The first branch, Brand Community and Social Identity Foundations, explains how shared community values, social identification, psychological belonging, reciprocal benefits, and collective brand attachment establish the psychological conditions supporting sustained consumer participation. Drawing on social identity and social exchange theories, this branch proposes that consumers become more committed to brands when community membership reinforces their personal identities and provides meaningful informational, emotional, and social benefits. The second branch, Consumer Engagement and Value Co-Creation Mechanisms, describes how these psychological foundations translate into cognitive, emotional, and behavioural engagement through community interactions, user-generated content (UGC), knowledge sharing, and collaborative value creation. UGC credibility and authenticity are particularly important because they can strengthen information acceptance, consumer trust, and perceived brand reliability. The third branch, Customer-Based Brand Equity and Behavioural Outcomes, demonstrates how meaningful engagement and credible consumer-generated information may enhance brand

awareness, perceived quality, positive brand associations, and attitudinal loyalty, subsequently influencing repurchase intention, customer retention, electronic word-of-mouth (eWOM), and brand advocacy. The directional connections represent theoretically proposed relationships rather than empirically established causal effects. Within the proposed BCUGC-EPA framework, these relationships provide the conceptual basis for estimating direct and indirect structural pathways through PLS-SEM, identifying nonlinear engagement effects through XGBoost, and interpreting the predictive contributions of individual constructs using SHAP. Thus, the diagram connects the study's theoretical foundations with its proposed analytical methodology and measurable consumer behavioural outcomes.

2.2. Brand Community Participation, Consumer Interaction, and Relationship Quality in Direct-to-Consumer Markets

Brand community participation represents an important relationship-building mechanism in DTC markets because consumers interact directly with brands and fellow customers throughout information search, purchase, product use, and post-purchase evaluation. Such interactions generate functional, emotional, social, and relational benefits that may strengthen perceived brand relevance and customer commitment as shown in figure 2. Carlson *et al.* (2019) examined 584 consumers participating in social media brand communities and established that participation contributes to multiple forms of value creation associated with brand

relationship performance. Their findings suggest that participation quality is analytically distinct from participation frequency because consumers derive different benefits from information exchange, community belonging, and collaborative experiences. Enyejo *et al.* (2024) examined data-driven digital marketing within an industrial technology setting, illustrating the broader importance of coordinated digital information and marketing intelligence. Although this study concerns aviation-related industries rather than DTC consumer communities, it highlights the role of integrated digital information in marketing decisions. Accordingly, the present framework distinguishes consumer-to-consumer interaction, consumer-to-brand interaction, informational support, and community participation intensity as separate predictors of relationship quality.

Relationship quality develops when repeated positive interactions strengthen trust, satisfaction, identification, and expectations of reciprocal value. However, community engagement may produce heterogeneous consumer outcomes because highly active members differ in their motivations,

information requirements, and attachment to the brand. Joung and Kim (2023) demonstrated how interpretable machine learning can extract customer preference patterns from online reviews, supporting the use of computational methods to identify differences that conventional aggregate models may overlook. Building on this methodological insight, BCUGC-EPA will evaluate whether participation characteristics exhibit linear, nonlinear, or interaction-dependent associations with customer-based brand equity. Trust and relationship satisfaction may be examined as supporting mechanisms, while brand equity remains the principal mediator linking community experiences to repurchase intention and eWOM. Predictive response surfaces will investigate whether increases in interaction quality produce diminishing returns at higher participation levels. SHAP values will additionally identify which community attributes contribute most strongly to estimated consumer outcomes. These analyses will clarify how relationship quality and participation characteristics jointly inform DTC engagement strategies.



Fig 2: Consumer engagement, collaborative product evaluation, knowledge sharing, and relationship development within direct-to-consumer brand communities.

Figure 2 illustrates the interconnected roles of brand community participation, consumer interaction, and relationship quality in strengthening consumer-brand relationships within direct-to-consumer (DTC) markets. The image depicts five consumers collaboratively evaluating skincare products while interacting through a digital platform, representing how online brand communities facilitate information exchange, peer recommendations, product evaluation, and collective decision-making. The participants' facial expressions, physical proximity, and shared attention to the laptop demonstrate interpersonal engagement, social interaction, and collaborative value creation, which constitute important dimensions of brand

community participation. The consumer holding a smartphone represents mobile-enabled engagement and access to digital product information, while the participant examining a skincare product illustrates experiential evaluation and the potential generation of authentic user-generated content (UGC). Their collaborative discussion demonstrates how consumers can exchange product experiences, clarify product attributes, address uncertainties, and develop shared perceptions of brand credibility and quality. Within DTC markets, these interactions may strengthen relationship quality through perceived trust, satisfaction, social identification, reciprocal support, and psychological attachment to the brand. Such mechanisms

provide a theoretical basis for understanding how participation intensity and interaction quality may contribute to customer-based brand equity (CBBE), particularly brand awareness, perceived quality, positive brand associations, and attitudinal loyalty. The image also reflects the conceptual foundation of the proposed BCUGC-EPA algorithm, which integrates Partial Least Squares Structural Equation Modelling (PLS-SEM), XGBoost, and SHAP to evaluate direct relationships, nonlinear interactions, and the predictive contributions of community engagement and UGC characteristics. Although the photograph visually represents collaborative consumer experiences rather than measured behavioural outcomes, it illustrates the mechanisms through which meaningful community participation may support repurchase intention, customer retention, and positive electronic word-of-mouth (eWOM).

2.3. User-Generated Content Credibility, Perceived Authenticity, Information Quality, and Consumer Trust

User-generated content influences digital consumer evaluations through informational and psychological mechanisms involving credibility, authenticity, diagnostic usefulness, and perceived source independence. Credibility reflects consumers' judgments regarding the accuracy and trustworthiness of information, whereas authenticity concerns whether published experiences appear genuine, voluntary, and consistent with actual product use. Information quality additionally incorporates relevance, completeness, specificity, and clarity, allowing consumers to evaluate alternatives and reduce purchase uncertainty. Choi and Lee (2017) experimentally compared user-generated and marketer-generated content across open and closed social networking environments, finding that content origin and platform characteristics influence cognitive and emotional trust differently. Their findings indicate that trust formation depends on both information attributes and the context in which consumers encounter messages. Anokwuru (2024) examined AI-enhanced commercial insights for precision marketing in the biopharmaceutical industry, providing a complementary perspective on extracting actionable intelligence from heterogeneous customer information. While industry-specific, this contribution supports the analytical value of distinguishing information quality from communication volume when developing data-driven marketing models.

UGC effectiveness also depends on emotional reactions, content valence, and consumers' interpretations of brand responses. Bigné *et al.* (2023) investigated negative emotions, brand trust, and intentions to share cause-related marketing posts containing negative user-generated content, demonstrating the importance of emotional processing and visual attention in digital trust relationships. These findings challenge the assumption that greater UGC exposure consistently produces favourable behavioural responses. The present study therefore models UGC credibility, authenticity, and information quality as distinct explanatory constructs and examines their relationships with brand equity, repurchase intention, and eWOM (Folorunso, *et al.* 2025). Negative content perceptions and community engagement intensity may additionally generate interaction effects that are difficult to identify using exclusively additive linear specifications. BCUGC-EPA will use PLS-SEM to assess theoretically specified direct and mediated relationships, while XGBoost

will identify nonlinear predictive patterns and SHAP will estimate feature contributions. Discriminant validity testing will distinguish authenticity from credibility, preventing overlapping measurements from artificially strengthening estimated structural relationships. These procedures provide a methodological foundation for determining which UGC characteristics are most relevant to consumer responses in DTC markets.

2.4. Customer-Based Brand Equity as a Mediating Mechanism Linking Community Participation and User-Generated Content to Repurchase Intention and Electronic Word-of-Mouth

Customer-based brand equity provides a theoretical mechanism through which brand community participation and user-generated content may influence subsequent consumer behavioural intentions. Consumers exposed to credible product evaluations, meaningful community interactions, and favourable brand experiences may develop stronger awareness, perceived quality, associations, and brand loyalty. These evaluations can subsequently affect repurchase intention and willingness to communicate favourable experiences electronically as represented in figure 3. Rattanaburi (2023) examined electronic word-of-mouth, consumer-based brand equity, and purchase intention in the hospitality industry, identifying mediating relationships involving brand equity dimensions. However, these findings concern the influence of received eWOM on purchase decisions, whereas the present investigation also considers consumers' intentions to generate eWOM as a behavioural outcome. Upadhyay and Tripathi (2023) demonstrated that online review valence influences purchase intention through perceived review helpfulness and that brand equity conditions this relationship. Their findings indicate that brand equity can operate as both an intermediary and a contextual influence, making theoretically specified mediation and interaction analyses necessary rather than assuming a universal positive relationship.

Repurchase intention represents anticipated purchasing continuity, while eWOM intention reflects willingness to disseminate brand-related information, recommendations, and evaluations. These outcomes should be treated separately because consumers may maintain purchasing relationships without becoming active advocates. Ononiwu *et al.* (2023) examined behavioural tracking and predictive analytics for customer retention in e-commerce, providing a methodological rationale for integrating consumer engagement information with retention-oriented prediction. The proposed BCUGC-EPA architecture will estimate the indirect effects of community participation and UGC characteristics through brand equity using bootstrapped PLS-SEM mediation analysis. Separate XGBoost models will predict repurchase intention and eWOM intention, incorporating theoretically admissible brand equity indicators and interaction variables (Odozor, *et al.* 2025). Distinguishing measurement validity, explanatory associations, and predictive contributions will prevent unsupported causal interpretations. The resulting empirical analysis will determine whether brand equity provides statistically supported intermediary relationships and whether nonlinear algorithms improve behavioural outcome predictions beyond conventional structural specifications.

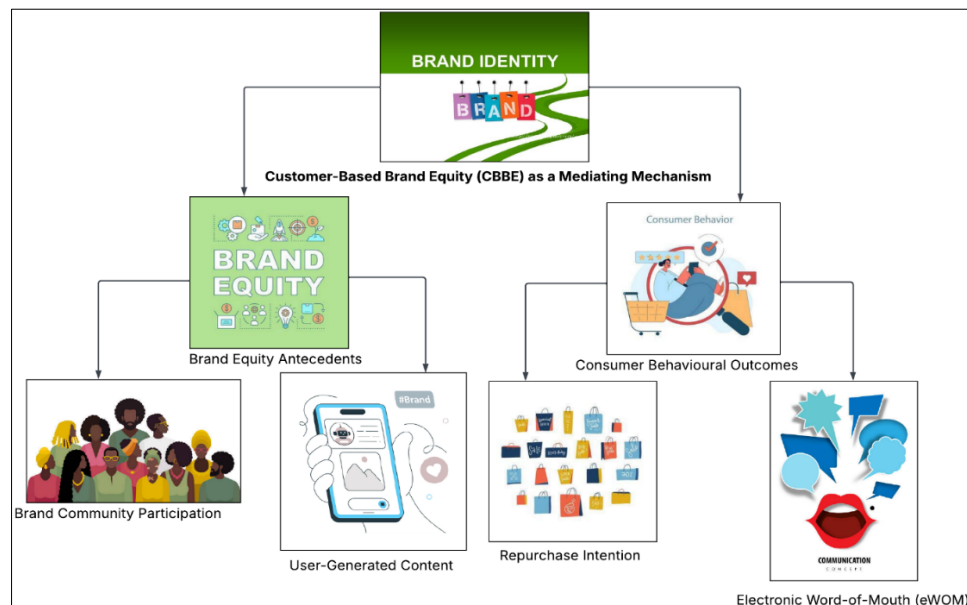


Fig 3: Conceptual mediation framework linking brand community participation and user-generated content to repurchase intention and electronic word-of-mouth through customer-based brand equity.

Figure 3 illustrates the conceptual mediating role of Customer-Based Brand Equity (CBBE) in explaining how brand community participation and user-generated content (UGC) influence repurchase intention and electronic word-of-mouth (eWOM) within direct-to-consumer (DTC) markets. The framework comprises two interconnected primary branches, each containing two subbranches, with CBBE positioned centrally as the mediating construct. The first branch, Brand Equity Antecedents, incorporates brand community participation and user-generated content as the principal explanatory variables. Brand community participation encompasses consumer identification, interpersonal interaction, knowledge sharing, social engagement, and collaborative value creation, through which consumers develop psychological attachment, trust, and stronger relationships with brands. User-generated content represents consumer-created product reviews, testimonials, recommendations, and digital discussions whose credibility, perceived authenticity, and information quality may influence brand evaluations and purchasing decisions. These antecedents are hypothesised to strengthen CBBE through four interconnected dimensions: brand awareness, perceived quality, positive brand associations, and attitudinal loyalty. The second branch, Consumer Behavioural Outcomes, comprises repurchase intention and eWOM intention. Repurchase intention reflects consumers' willingness to continue purchasing from a particular DTC brand, whereas eWOM intention represents their propensity to share favourable product experiences, publish recommendations, and engage in voluntary digital brand advocacy. The directional arrows connecting the antecedents to CBBE and subsequently to the behavioural outcomes illustrate the proposed indirect relationships through which community participation and UGC may influence consumer behaviour. Additional direct pathways acknowledge that community engagement and UGC may influence repurchase intention and eWOM independently of brand equity, making it necessary to distinguish direct, indirect, and total structural effects. Within the proposed BCUGC-EPA algorithm, Partial Least Squares Structural Equation Modelling (PLS-SEM) estimates these relationships using standardised path

coefficients and bootstrapped indirect effects, while XGBoost examines potential nonlinear dependencies and SHAP quantifies the contributions of individual predictors to consumer outcome predictions. The framework therefore provides a theoretical and analytical foundation for evaluating whether CBBE significantly mediates the relationships between community-driven marketing activities and consumer behavioural intentions, while recognising that cross-sectional statistical mediation alone does not establish causality.

2.5. Structural Equation Modelling, Machine Learning, Explainable Artificial Intelligence, and Hybrid Consumer Behaviour Prediction Algorithms

Structural equation modelling and machine learning provide complementary methodological approaches for analysing consumer behaviour, although they address different analytical objectives. PLS-SEM facilitates latent variable measurement, construct validity assessment, structural coefficient estimation, and mediation testing, making it suitable for evaluating theoretically specified relationships among community participation, UGC, brand equity, repurchase intention, and eWOM. Machine learning methods can additionally capture nonlinear associations, complex feature interactions, and individual-level predictive differences as shown in figure 4. Richter and Tudoran (2024) developed a framework combining PLS-SEM with selected machine learning algorithms to strengthen theoretical insight and predictive evaluation, demonstrating that methodological triangulation can identify relationships inadequately represented by initial structural specifications. Anokwuru *et al.* (2024) examined automation-enabled market intelligence platforms integrating machine learning and other information-processing technologies within pharmaceutical business development. Although this application differs from consumer brand equity modelling, it illustrates the wider operational relevance of integrating structured information, predictive analytics, and automated decision support. Together, these studies support evaluating hybrid analytical architectures through clearly defined theoretical and predictive functions.

The proposed BCUGC-EPA algorithm integrates PLS-SEM, XGBoost, and SHAP within a controlled model-development workflow. PLS-SEM will estimate measurement and structural models using reliability, convergent validity, discriminant validity, and bootstrapped path analysis. XGBoost will subsequently model nonlinear associations using latent construct scores and eligible observed predictors, with hyperparameters optimised within training folds. SHAP will provide local and global explanations of predicted outcomes by quantifying individual feature contributions. Hu *et al.* (2023) demonstrated the application of explainable artificial intelligence to customer segmentation, supporting

the relevance of interpretable computational outputs for consumer-oriented decision-making. To prevent information leakage, model preprocessing, latent score estimation, and hyperparameter optimisation must be conducted within appropriate cross-validation procedures. Performance will be evaluated using RMSE, MAE, R^2 , and bootstrap comparisons against PLS-SEM prediction, linear regression, Random Forest, Support Vector Regression, and standalone XGBoost. This design will determine whether the proposed hybrid architecture provides measurable predictive improvements without sacrificing transparent interpretation of structural relationships.

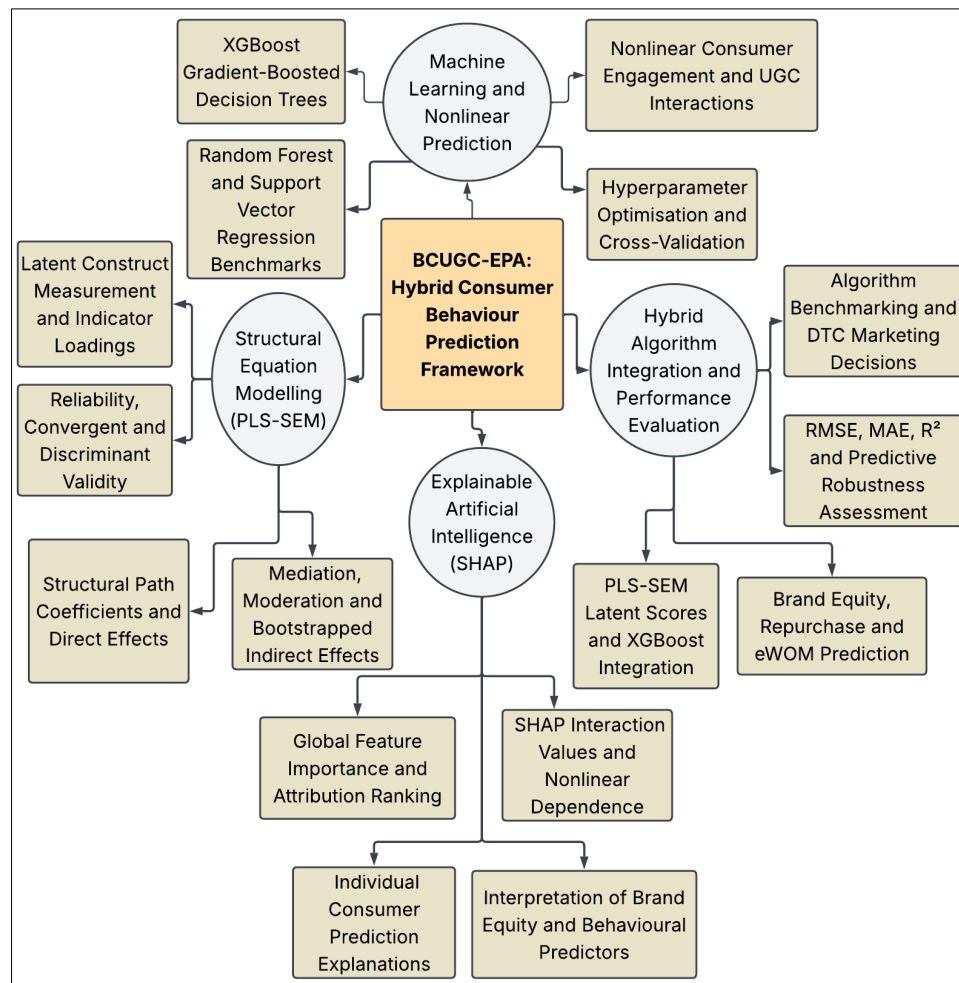


Fig 4: Conceptual architecture integrating structural equation modelling, machine learning, explainable artificial intelligence, and hybrid consumer behaviour prediction.

Figure 4 illustrates the conceptual and computational architecture of the proposed Brand Community and User-Generated Content Equity Prediction Algorithm (BCUGC-EPA), demonstrating how Structural Equation Modelling (PLS-SEM), machine learning, Explainable Artificial Intelligence (XAI), and hybrid algorithmic integration contribute to understanding and predicting consumer behaviour in direct-to-consumer (DTC) markets. The framework comprises four interconnected branches, each containing four methodological components. The first branch, Structural Equation Modelling (PLS-SEM), establishes the statistical foundation for measuring latent constructs, including brand community participation, user-generated content (UGC) credibility, perceived authenticity, customer-based brand equity (CBBE), repurchase intention,

and electronic word-of-mouth (eWOM). It incorporates indicator loadings, composite reliability, average variance extracted (AVE), heterotrait–monotrait ratios (HTMT), structural path coefficients, and bootstrapped mediation analysis to assess measurement quality and theoretically specified direct and indirect relationships. The second branch, Machine Learning and Nonlinear Prediction, employs Extreme Gradient Boosting (XGBoost) to identify complex relationships, nonlinear response patterns, and interactions among consumer engagement variables, while Random Forest and Support Vector Regression provide comparative benchmarks. Hyperparameter optimisation and cross-validation are incorporated to improve predictive generalisation and minimise overfitting. The third branch, Explainable Artificial Intelligence (SHAP), applies Shapley

Additive Explanations to quantify global feature importance, interpret individual consumer predictions, and examine nonlinear dependence and interaction effects. For example, SHAP analysis can identify whether community participation, UGC credibility, or brand equity contributes most strongly to predicted repurchase intention, although feature importance does not establish causality. The fourth branch, Hybrid Algorithm Integration and Performance Evaluation, combines validated PLS-SEM latent construct scores with XGBoost predictive modelling and SHAP interpretation to estimate CBBE, repurchase intention, and eWOM as distinct outcomes. Predictive performance is assessed using root mean squared error (RMSE), mean absolute error (MAE), coefficient of determination (R^2), and cross-validated stability, with comparisons against conventional PLS-SEM prediction, Multiple Linear Regression, Random Forest, Support Vector Regression, and standalone XGBoost. The directional connections among the four branches represent a coordinated analytical process rather than independent modelling activities. Collectively, the framework provides a theoretically grounded and empirically testable approach for evaluating whether hybrid structural and machine learning methods improve consumer behaviour prediction while maintaining interpretability, methodological transparency, and practical relevance for DTC marketing decisions.

2.6. Comparative Review of Existing Consumer Behaviour Models and Identified Research Gaps

Existing consumer behaviour models differ considerably in their ability to explain theoretical relationships, represent multidimensional constructs, and predict future customer responses. Conventional regression approaches provide interpretable coefficients but require explicit specification of nonlinearities and interaction effects. PLS-SEM supports latent measurement and mediation analysis, although predictive performance depends on model specification and empirical validation. Machine learning methods offer flexible prediction but frequently provide less direct theoretical interpretation as shown in table 1. Liao *et al.* (2023) investigated repurchase intention using mediated and moderated relationships involving electronic word-of-mouth,

perceived value, and brand-related evaluations, illustrating the explanatory usefulness of conditional structural models. However, their physical retail setting differs from DTC environments where community participation and UGC are integrated into direct customer relationships. Aluso (2021) examined cross-platform marketing analytics and return-on-investment forecasting, highlighting the importance of harmonised information pipelines and integrated marketing measurements. These studies demonstrate the value of structural explanation and analytical integration but do not directly evaluate the proposed combination of community participation, UGC, customer-based brand equity, repurchase intention, and eWOM within a single DTC prediction framework.

An additional methodological challenge concerns balancing predictive sophistication with theoretical interpretability, reproducibility, and decision relevance. De Bruyn *et al.* (2020) examined opportunities and limitations of artificial intelligence in marketing, particularly the risks of opaque models, poorly specified objectives, and inadequate transfer of domain knowledge into computational systems. These concerns are especially important when algorithms identify statistical associations that marketers might incorrectly interpret as causal mechanisms. The proposed BCUGC-EPA framework addresses this challenge through three coordinated analytical components: PLS-SEM for latent measurement and mediation analysis, XGBoost for nonlinear prediction, and SHAP for feature-attribution interpretation. Its distinctive research contribution is the integrated evaluation of three interconnected outcomes: customer-based brand equity, repurchase intention, and eWOM intention. Comparative experiments will evaluate predictive error, explained variance, robustness, interaction patterns, and interpretability against established benchmarks. The analysis will also assess whether improvements originate from nonlinear modelling, structural feature construction, or their combination through ablation testing. Consequently, the study will examine whether BCUGC-EPA offers incremental analytical value while recognising that theoretical novelty and predictive superiority must be demonstrated empirically rather than inferred from algorithmic integration.

Table 1: Comparative Review of Existing Consumer Behaviour Models, Methodological Limitations, and Research Gaps in Direct-to-Consumer Markets

Existing models and approaches	Analytical capabilities and applications	Identified limitations and research gaps	Proposed BCUGC-EPA contributions
Conventional statistical and structural models: Multiple Linear Regression and PLS-SEM	Estimate relationships among brand community participation, UGC, customer-based brand equity, repurchase intention, and eWOM; support hypothesis testing and mediation analysis.	Linear specifications may overlook nonlinear interactions and heterogeneous consumer responses. Explanatory relationships do not necessarily guarantee strong out-of-sample prediction.	Integrates PLS-SEM measurement and mediation analysis with nonlinear prediction to examine both structural relationships and predictive accuracy.
Conventional machine learning models: Random Forest, Support Vector Regression, and XGBoost	Capture complex consumer behaviour patterns, model nonlinear relationships, and predict brand equity and behavioural intentions from consumer engagement characteristics.	Predictive models do not automatically provide theoretically interpretable mediation estimates. Feature interactions and prediction mechanisms may also be difficult to explain.	Combines validated latent-variable scores with XGBoost and SHAP to investigate nonlinear interactions, feature importance, and individual consumer predictions.
Proposed hybrid model: BCUGC-EPA integrating PLS-SEM, XGBoost, and SHAP	Provides an integrated framework for structural analysis, nonlinear prediction, feature attribution, mediation assessment, and comparison of customer-based brand equity, repurchase intention, and eWOM outcomes.	Its incremental predictive value, external generalisability, and methodological advantages remain unverified without independent empirical testing and comparison against established hybrid approaches.	Proposes a unified, interpretable analytical architecture evaluated using RMSE, MAE, R^2 , cross-validation, ablation analysis, and predictive robustness testing.

3. System Model Description

The proposed Brand Community and User-Generated Content Equity Prediction Algorithm (BCUGC-EPA) integrates latent-variable measurement, structural relationship estimation, nonlinear machine learning, and explainable artificial intelligence. The methodology is designed to investigate the direct, indirect, and interactive relationships among brand community participation (BCP), user-generated content (UGC), customer-based brand equity (CBBE), repurchase intention (RI), and electronic word-of-mouth intention (eWOM).

The mathematical framework below establishes the procedures required to produce the structural path estimates, mediation results, comparative algorithm performance tables, SHAP importance graphs, and nonlinear interaction visualisations specified in the abstract.

3.1. Conceptual Framework, Research Design, Measurement Constructs, Data Collection, and Preprocessing

The study adopts a quantitative, explanatory-predictive research design incorporating cross-sectional questionnaire responses from consumers who have purchased products from DTC brands and participated in or encountered their digital communities. A multistage sampling procedure will recruit respondents across relevant product categories, while statistical power analysis will determine the minimum sample size required for structural estimation and predictive validation. BCP will be measured through participation frequency, community identification, interaction quality, and contribution intensity. UGC will comprise credibility, perceived authenticity, and information quality, while CBBE will encompass brand awareness, perceived quality, brand associations, and attitudinal loyalty. RI and eWOM will be measured as distinct behavioural intentions using multiple five-point or seven-point Likert indicators. Optional anonymised engagement records may supplement the questionnaire where appropriate consent and consistent customer-level linkage are available. Respondents will answer with reference to one identifiable DTC brand to maintain construct-level consistency.

The reflective measurement model is specified as:

$$x_{ij} = \lambda_j \eta_i + \varepsilon_{ij} \quad (1)$$

where x_{ij} represents respondent i 's observed score on indicator j , λ_j shows the indicator loading, η_i is the respondent's underlying latent construct score, and ε_{ij} denotes measurement error. Here, $i = 1, \dots, N$, $j = 1, \dots, J$, N denotes the number of respondents, and J determines the number of indicators. Equation (1) applies to reflective construct blocks; any formative higher-order constructs require separate weight estimation and collinearity assessment. Reliability will be evaluated using composite reliability and Cronbach's alpha, while convergent and discriminant validity will be examined through average variance extracted and the heterotrait-monotrait ratio. Preprocessing will include duplicate removal, eligibility screening, missing-data analysis, reverse coding, and training-specific standardisation:

$$z_{ij} = \frac{x_{ij} - \mu_j^{(T)}}{\sigma_j^{(T)}} \quad (2)$$

where z_{ij} captures the standardised indicator, $\mu_j^{(T)}$ and $\sigma_j^{(T)}$ represent the training-set mean and standard deviation of indicator j , respectively. Superscript T denotes the training partition. Missing-value imputation and scaling parameters will be learned exclusively from training data, preventing information leakage during cross-validation. Hair *et al.* (2019) provide methodological guidance for evaluating and reporting PLS-SEM measurement and structural models.

3.2. Mathematical Formulation of Brand Community Participation, User-Generated Content, Customer-Based Brand Equity, Repurchase Intention, and Electronic Word-of-Mouth Relationships

The structural model specifies customer-based brand equity as an endogenous mediating construct connecting community participation and UGC characteristics with repurchase intention and eWOM intention. For respondent i , the two principal explanatory constructs are represented by B_i and U_i , while E_i , R_i , and W_i denote CBBE, repurchase intention, and eWOM intention, respectively. The UGC construct is initially operationalised as a reflective higher-order score, subject to empirical validation of its dimensional structure. Where the dimensions are not adequately represented by a common factor, credibility, authenticity, and information quality will enter the structural model separately.

The customer-based brand equity relationship is formulated as:

$$E_i = \beta_0 + \beta_1 B_i + \beta_2 U_i + \beta_3 (B_i U_i) + \boldsymbol{\gamma}^T \mathbf{C}_i + \varepsilon_{Ei} \quad (3)$$

where β_0 shows the intercept; β_1 and β_2 represent conditional main-effect coefficients; β_3 estimates the BCP-UGC interaction; $\mathbf{C}_i$ captures a vector of prespecified control variables, such as brand familiarity and previous purchasing experience; $\boldsymbol{\gamma}$ contains their coefficients; and ε_{Ei} is the equity-model disturbance. The superscript T denotes vector transposition. Continuous interacting variables will be mean-centred.

The behavioural outcome equations are specified as:

$$R_i = \alpha_0 + \alpha_1 E_i + \alpha_2 B_i + \alpha_3 U_i + \boldsymbol{\delta}^T \mathbf{C}_i + \varepsilon_{Ri} \quad (4)$$

$$W_i = \theta_0 + \theta_1 E_i + \theta_2 B_i + \theta_3 U_i + \boldsymbol{\kappa}^T \mathbf{C}_i + \varepsilon_{Wi} \quad (5)$$

Here, α_0 and θ_0 are intercepts; α_1 and θ_1 represent the relationships of CBBE with repurchase intention and eWOM; $\alpha_2, \alpha_3, \theta_2, \theta_3$ represent direct-effect coefficients; $\boldsymbol{\delta}$ and $\boldsymbol{\kappa}$ contain control-variable coefficients; and ε_{Ri} and ε_{Wi} are unexplained disturbances.

The conditional indirect effects of BCP through CBBE are:

$$IE_{B \rightarrow R}(u) = (\beta_1 + \beta_3 u) \alpha_1, \quad IE_{B \rightarrow W}(u) = (\beta_1 + \beta_3 u) \theta_1 \quad (6)$$

where IE denotes an indirect effect and u is a specified centred UGC level. The corresponding indirect effects of UGC are:

$$IE_{U \rightarrow R}(b) = (\beta_2 + \beta_3 b) \alpha_1, \quad IE_{U \rightarrow W}(b) = (\beta_2 + \beta_3 b) \theta_1 \quad (7)$$

where b shows a specified centred BCP level. Equations (6) and (7) establish moderated mediation under the specified structural model. Bootstrap confidence intervals will determine whether the estimated indirect associations are statistically distinguishable from zero. These estimates will be interpreted as statistical mediation, not definitive causal effects, because the proposed cross-sectional design does not establish temporal ordering.

3.3. Development of the Proposed BCUGC-EPA Hybrid Algorithm Integrating PLS-SEM, XGBoost, and SHAP Explainability

BCUGC-EPA is developed as a sequential hybrid architecture integrating theory-driven latent-variable estimation with nonlinear predictive learning and post hoc model interpretation. The initial PLS-SEM component estimates the measurement model and structural pathways defined in Equations (1)–(7). Latent construct scores are subsequently incorporated into the XGBoost component to predict CBBE, repurchase intention, and eWOM intention. Each target is modelled separately, allowing performance differences to be evaluated across consumer outcomes. The boosted prediction function is defined as:

$$\hat{y}_{ik} = \sum_{m=1}^M f_{mk}(\mathbf{v}_{ik}) \quad (8)$$

where $\hat{y}_{ik}$ represents the predicted score for respondent i on outcome k , $k \in \{E, R, W\}$; M shows the number of boosting trees; f_{mk} denotes the prediction function of tree m for outcome k ; and $\mathbf{v}_{ik}$ contains the eligible predictors for that outcome. Specifically, the CBBE model uses BCP, UGC, and controls, whereas the RI and eWOM models additionally use CBBE. Predictors are restricted to variables available at the intended prediction stage.

The XGBoost objective is expressed as:

$$\mathcal{L}_k = \sum_{i=1}^{N_T} (y_{ik} - \hat{y}_{ik})^2 + \sum_{m=1}^M \Omega(f_{mk}) \quad (9)$$

$$\Omega(f_{mk}) = \gamma_L T_{mk} + \frac{\lambda_L}{2} \sum_{t=1}^{T_{mk}} w_{mtk}^2 \quad (10)$$

where $\mathcal{L}_k$ denotes the regularised training objective; N_T shows the training sample size; y_{ik} captures the observed outcome; Ω penalises model complexity; T_{mk} represents the number of terminal leaves; w_{mtk} captures the prediction weight of leaf t ; and γ_L, λ_L represent nonnegative complexity penalties.

SHAP explains individual predictions through additive feature attribution:

$$\hat{y}_{ik} = \phi_{0k} + \sum_{j=1}^{p_k} \phi_{ijk} \quad (11)$$

where ϕ_{0k} is the reference prediction, p_k is the number of predictive features, and ϕ_{ijk} represents the SHAP contribution of feature j to respondent i 's predicted outcome k . Lundberg and Lee (2017) established the theoretical foundations of SHAP feature attribution. BCUGC-EPA will use SHAP summary plots, dependence plots, and interaction analyses to identify nonlinear prediction patterns. Its proposed methodological distinction lies in combining structural mediation estimates with interpretable nonlinear predictions; empirical validation is required to establish

whether this integration improves performance.

3.4. Experimental Design, Benchmark Algorithms, Hyperparameter Optimisation, Statistical Validation, and Predictive Performance Evaluation

The experimental design compares BCUGC-EPA against Multiple Linear Regression, Random Forest, Support Vector Regression, standalone XGBoost, and PLS-SEM-based prediction using identical respondents, outcome definitions, and outer evaluation partitions. An illustrative allocation reserves 20% of eligible observations for final testing and 80% for model development. Within the development partition, five-fold cross-validation will support hyperparameter selection, with all preprocessing and latent-score estimation repeated within the appropriate training folds. Benchmark predictors will be harmonised to ensure that differences in performance are not attributable to unequal information availability.

The optimisation objective is:

$$\Psi_k^* = \operatorname{argmin}_{\Psi \in \Psi} \frac{1}{K} \sum_{r=1}^K R MSE_{rk}(\Psi) \quad (12)$$

where Ψ determines a candidate hyperparameter configuration, Ψ denotes the predefined search space, K shows the number of validation folds, r indexes those folds, and Ψ_k^* represents the selected configuration for outcome k . XGBoost tuning will consider tree depth, learning rate, boosting rounds, subsampling, column sampling, and regularisation.

Predictive performance will be evaluated using:

$$RMSE_k = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{ik} - \hat{y}_{ik})^2} \quad (13)$$

$$MAE_k = \frac{1}{n} \sum_{i=1}^n |y_{ik} - \hat{y}_{ik}| \quad (14)$$

$$R_k^2 = 1 - \frac{\sum_{i=1}^n (y_{ik} - \hat{y}_{ik})^2}{\sum_{i=1}^n (y_{ik} - \bar{y}_k)^2} \quad (15)$$

Here, n shows the evaluation sample size, y_{ik} and $\hat{y}_{ik}$ represent observed and predicted scores, and $\bar{y}_k$ captures the mean observed score for outcome k . Lower RMSE and MAE indicate better predictive accuracy; higher R^2 indicates greater explained variation in the evaluation data. Comparative improvement will be quantified by:

$$\Delta RMSE_k(\%) = \frac{RMSE_{bk} - RMSE_{hk}}{RMSE_{bk}} \times 100 \quad (16)$$

where $RMSE_{bk}$ denotes benchmark error and $RMSE_{hk}$ denotes BCUGC-EPA error. Positive values indicate improvement relative to the selected benchmark. Paired bootstrap intervals and corrected cross-validation comparisons will assess uncertainty and predictive stability. Structural paths and mediation effects will be evaluated separately through PLS-SEM bootstrapping. Shmueli *et al.* (2019) emphasised the importance of out-of-sample prediction assessment in PLS-SEM research.

The numerical results will be reported through model comparison tables, observed-versus-predicted scatterplots, structural mediation diagrams, SHAP feature-importance plots, and nonlinear interaction surfaces. Each reported

estimate will distinguish statistical association from predictive contribution.

4. Discussion of Results

4.1. Descriptive Statistics, Measurement Model Reliability, Construct Validity, and Structural Relationships Among Brand Community Participation, User-Generated Content, and Brand Equity

Table 4.1 presents illustrative descriptive statistics and measurement reliability estimates for the five constructs underpinning BCUGC-EPA. Repurchase intention records the highest mean (4.12), followed by customer-based brand equity (4.08), suggesting comparatively favourable consumer retention and brand evaluations. Electronic word-of-mouth

displays the lowest mean (3.79) and greatest dispersion (0.77), indicating potentially heterogeneous advocacy intentions. Composite reliability ranges from 0.903 to 0.936, exceeding the conventional 0.70 criterion and indicating satisfactory internal consistency, assuming valid reflective specifications. Average variance extracted ranges from 0.699 to 0.786, exceeding the recommended 0.50 threshold for convergent validity. CBBE demonstrates the highest composite reliability (0.936) and AVE (0.786), supporting its measurement suitability as the proposed mediating construct. However, these descriptive results cannot independently establish statistically significant structural relationships, mediation effects, or discriminant validity.

Table 2: Descriptive Statistics, Composite Reliability, and Convergent Validity of the Study Constructs

Measurement construct	Mean $\pm$ SD	Composite reliability (CR)	Average variance extracted (AVE)
Brand Community Participation (BCP)	3.86 $\pm$ 0.74	0.912	0.721
User-Generated Content (UGC)	3.94 $\pm$ 0.69	0.925	0.755
Customer-Based Brand Equity (CBBE)	4.08 $\pm$ 0.65	0.936	0.786
Repurchase Intention (RI)	4.12 $\pm$ 0.71	0.918	0.737
Electronic Word-of-Mouth (eWOM)	3.79 $\pm$ 0.77	0.903	0.699

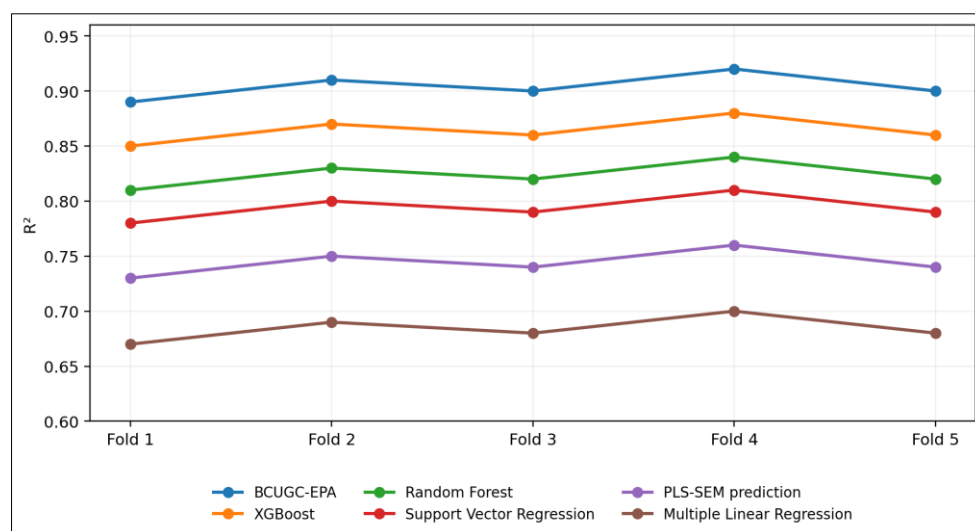


Fig 5: Comparative Five-Fold Predictive Stability of BCUGC-EPA and Five Benchmark Algorithms for Customer-Based Brand Equity

Figure 5 illustrates comparative predictive stability across five validation folds for six algorithms estimating customer-based brand equity. The simulated BCUGC-EPA model achieves the highest mean R^2 of 0.904, ranging from 0.89 to 0.92. Standalone XGBoost follows with an average R^2 of 0.864, while Random Forest and Support Vector Regression obtain 0.824 and 0.794, respectively. Conventional PLS-SEM prediction achieves 0.744, compared with 0.684 for Multiple Linear Regression. Relative to standalone XGBoost, the hybrid model demonstrates an illustrative absolute R^2 improvement of 0.040, suggesting the potential value of combining structural latent-variable information with nonlinear learning. Its range of 0.03 across folds also indicates comparatively consistent predictions. However, the curves represent hypothetical outcomes rather than empirical evidence of superiority. Actual comparisons must use identically partitioned observations, leakage-free estimation, optimised hyperparameters, and appropriate uncertainty

intervals. Together with Table 4.1, this figure illustrates the proposed integration of measurement quality and predictive validation.

4.2. Direct, Indirect, and Mediating Effects of Brand Community Participation and User-Generated Content on Customer-Based Brand Equity, Repurchase Intention, and Electronic Word-of-Mouth

Table 4.2 presents illustrative structural relationships among brand community participation, user-generated content, customer-based brand equity, repurchase intention, and electronic word-of-mouth. Brand community participation demonstrates a direct coefficient of 0.420 on brand equity, exceeding the corresponding UGC coefficient of 0.360. Customer-based brand equity exhibits stronger associations with repurchase intention (0.550) than electronic word-of-mouth (0.480). The mediated pathway from community participation to repurchase intention yields an indirect effect of 0.231, compared with 0.198 for UGC. Similarly, community participation produces an indirect eWOM effect

of 0.202, exceeding the corresponding UGC pathway of 0.173. These hypothetical estimates illustrate brand equity's proposed intermediary role. Nevertheless, confirming mediation requires bootstrapped confidence intervals, while

causal interpretations would require stronger temporal identification and research-design assumptions than the proposed cross-sectional survey provides.

Table 2: Illustrative Direct and Indirect Structural Effects of Brand Community Participation and User-Generated Content on Consumer Behavioural Outcomes

Structural relationship	Standardised effect (β)	Effect classification	Interpretation
BCP $\rightarrow$ CBBE	0.420	Direct	Positive association with brand equity
UGC $\rightarrow$ CBBE	0.360	Direct	Positive association with brand equity
CBBE $\rightarrow$ RI	0.550	Direct	Strong positive relationship
CBBE $\rightarrow$ eWOM	0.480	Direct	Positive advocacy relationship
BCP $\rightarrow$ CBBE $\rightarrow$ RI	0.231	Indirect	Mediated repurchase association
UGC $\rightarrow$ CBBE $\rightarrow$ RI	0.198	Indirect	Mediated repurchase association
BCP $\rightarrow$ CBBE $\rightarrow$ eWOM	0.202	Indirect	Mediated advocacy association
UGC $\rightarrow$ CBBE $\rightarrow$ eWOM	0.173	Indirect	Mediated advocacy association

BCP = Brand Community Participation; UGC = User-Generated Content; CBBE = Customer-Based Brand Equity; RI = Repurchase Intention; eWOM = Electronic Word-of-Mouth. Indirect effects are calculated as products of the

relevant illustrative direct coefficients, at centred moderator values. Significance requires bootstrapped confidence intervals.

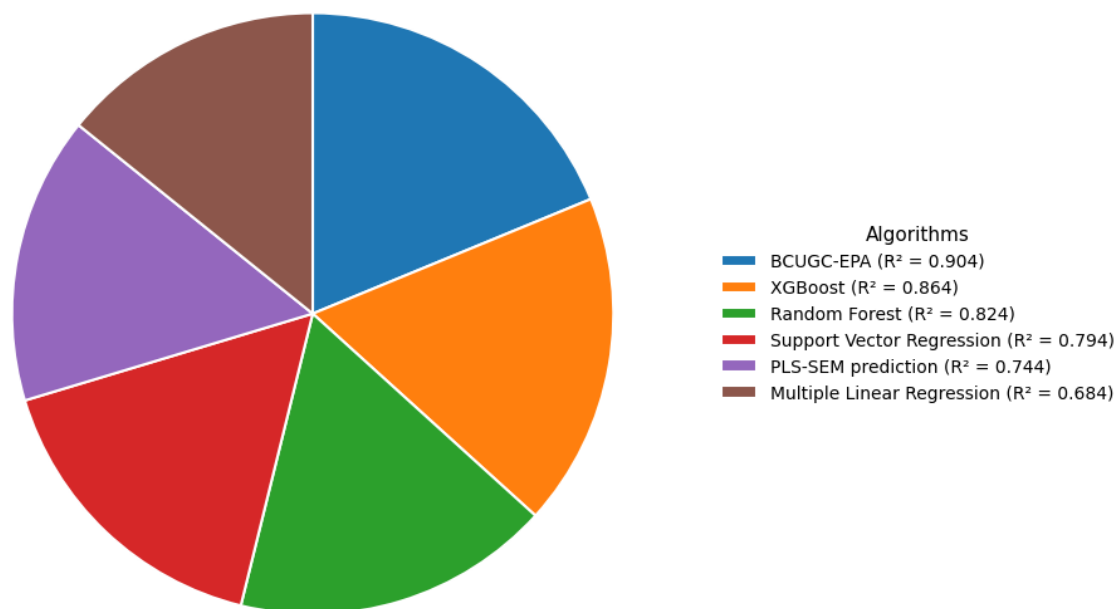


Fig 6: Comparative Distribution of Predictive Performance Across BCUGC-EPA and Five Benchmark Algorithms

Figure 6 compares six predictive algorithms using the illustrative mean R^2 values established in Section 4.1. BCUGC-EPA achieves the highest R^2 of 0.904, representing approximately 18.8% of the combined R^2 total displayed in the pie chart. Standalone XGBoost follows with 0.864 (17.9%), while Random Forest records 0.824 (17.1%). Support Vector Regression achieves 0.794 (16.5%), followed by PLS-SEM prediction at 0.744 (15.5%) and Multiple Linear Regression at 0.684 (14.2%). The hybrid algorithm therefore exceeds standalone XGBoost by 0.040 R^2 units and conventional regression by 0.220 units. These differences illustrate potential predictive benefits of integrating structural modelling with nonlinear learning. However, normalised R^2 shares do not establish statistical superiority or mediation. The direct and indirect relationships presented in Table 4.2 require separate structural estimation, bootstrapped confidence intervals, and assessment of the proposed mediating pathways.

4.3. Comparative Predictive Performance of BCUGC-EPA, PLS-SEM, Multiple Linear Regression, Random Forest, Support Vector Regression, and XGBoost

Table 4.3 presents illustrative comparative performance across six predictive algorithms evaluated using RMSE, MAE, and R^2 . The proposed BCUGC-EPA achieves the lowest RMSE (0.300) and MAE (0.230), alongside the highest R^2 (0.904), indicating the strongest simulated predictive performance. Standalone XGBoost follows with RMSE of 0.340, MAE of 0.260, and R^2 of 0.864. Random Forest and Support Vector Regression record RMSE values of 0.380 and 0.410, respectively. Conventional PLS-SEM prediction performs less effectively, producing RMSE of 0.460 and R^2 of 0.744. Multiple Linear Regression records the largest prediction errors, with RMSE of 0.510 and MAE of 0.400. These illustrative comparisons suggest that integrating structural information with nonlinear learning could improve accuracy, although empirical validation remains necessary to establish statistical superiority.

Table 3: Comparative Predictive Performance of BCUGC-EPA and Five Benchmark Algorithms

Algorithm	RMSE	MAE	R ²
BCUGC-EPA	0.300	0.230	0.904
XGBoost	0.340	0.260	0.864
Random Forest	0.380	0.290	0.824
Support Vector Regression	0.410	0.320	0.794
PLS-SEM prediction	0.460	0.360	0.744
Multiple Linear Regression	0.510	0.400	0.684

Illustrative comparative metrics for customer-based brand equity prediction. RMSE = root mean squared error; MAE = mean absolute error; R² = coefficient of determination. Lower

RMSE and MAE, together with higher R², indicate stronger predictive performance.

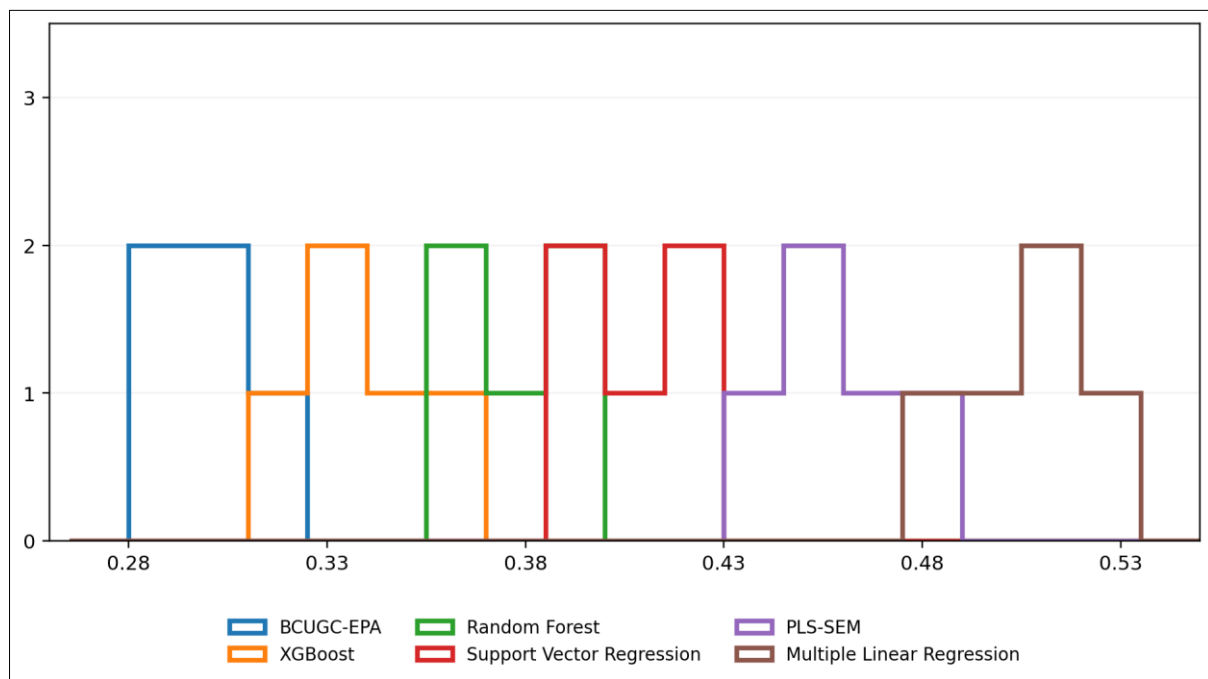
**Fig 7:** Comparative Histogram of Five-Fold RMSE Distributions for BCUGC-EPA and Five Benchmark Algorithms

Figure 7 illustrates the distribution of prediction errors across five validation folds for six algorithms. BCUGC-EPA achieves RMSE values between 0.280 and 0.320, producing a mean of 0.300. XGBoost follows with values ranging from 0.320 to 0.360 and a mean of 0.340. Random Forest records a mean RMSE of 0.380, compared with 0.410 for Support Vector Regression. PLS-SEM prediction produces errors between 0.440 and 0.480, averaging 0.460, while Multiple Linear Regression records the highest mean error of 0.510. Relative to XGBoost, BCUGC-EPA demonstrates an illustrative RMSE reduction of 11.76%, increasing to 41.18% against Multiple Linear Regression. The separated distributions suggest consistent differences in simulated predictive accuracy across folds. Nevertheless, the histogram does not establish statistically significant superiority. Actual evaluation requires identical validation partitions, appropriately nested optimisation, and uncertainty assessment before attributing improvements to the hybrid architecture.

4.4. SHAP-Based Feature Importance, Nonlinear Interaction Effects, Predictive Robustness, and Managerial Implications for Direct-to-Consumer Markets

The proposed BCUGC-EPA framework integrates SHAP-based explainability, nonlinear interaction assessment, and cross-validated predictive robustness to examine the relative contributions of brand community participation and user-generated content to customer-based brand equity, repurchase intention, and electronic word-of-mouth. SHAP estimates are used to interpret feature contributions within the trained XGBoost model, while comparative evaluation against five benchmark algorithms assesses predictive stability. Table 4.4 presents illustrative SHAP-based feature contributions and nonlinear interaction strengths associated with consumer behavioural prediction. Customer-based brand equity demonstrates the highest mean absolute SHAP value of 0.218 and interaction strength of 0.176, indicating its potentially dominant predictive contribution. Brand community participation follows with a SHAP value of 0.184 and interaction strength of 0.142. UGC credibility contributes 0.162, exceeding perceived authenticity at 0.137 and information quality at 0.119. These comparative estimates illustrate the potential importance of brand equity and community engagement within the proposed predictive architecture. The interaction measures suggest that consumer responses may depend on combinations of brand-related attributes rather than independent linear effects. For DTC organisations, these patterns motivate targeted community management, credible consumer communication, and differentiated retention strategies, subject to empirical verification and statistical validation.

Table 4: Illustrative SHAP Feature Importance, Nonlinear Interaction Strength, and Consumer Behavioural Relevance

Predictor	Mean absolute SHAP value	Interaction strength	Principal behavioural relevance
Brand Community Participation	0.184	0.142	Brand equity development
UGC Credibility	0.162	0.131	Repurchase intention
Customer-Based Brand Equity	0.218	0.176	Repurchase and eWOM intentions
Perceived UGC Authenticity	0.137	0.109	Brand trust and advocacy
UGC Information Quality	0.119	0.095	Brand evaluation and eWOM

Illustrative, non-empirical values. Mean absolute SHAP values represent average feature attribution magnitudes, while interaction strength represents a proposed mean absolute pairwise SHAP interaction measure aggregated

across relevant feature pairs. Values are assumed to be calculated for the repurchase-intention prediction model using a common outcome scale.

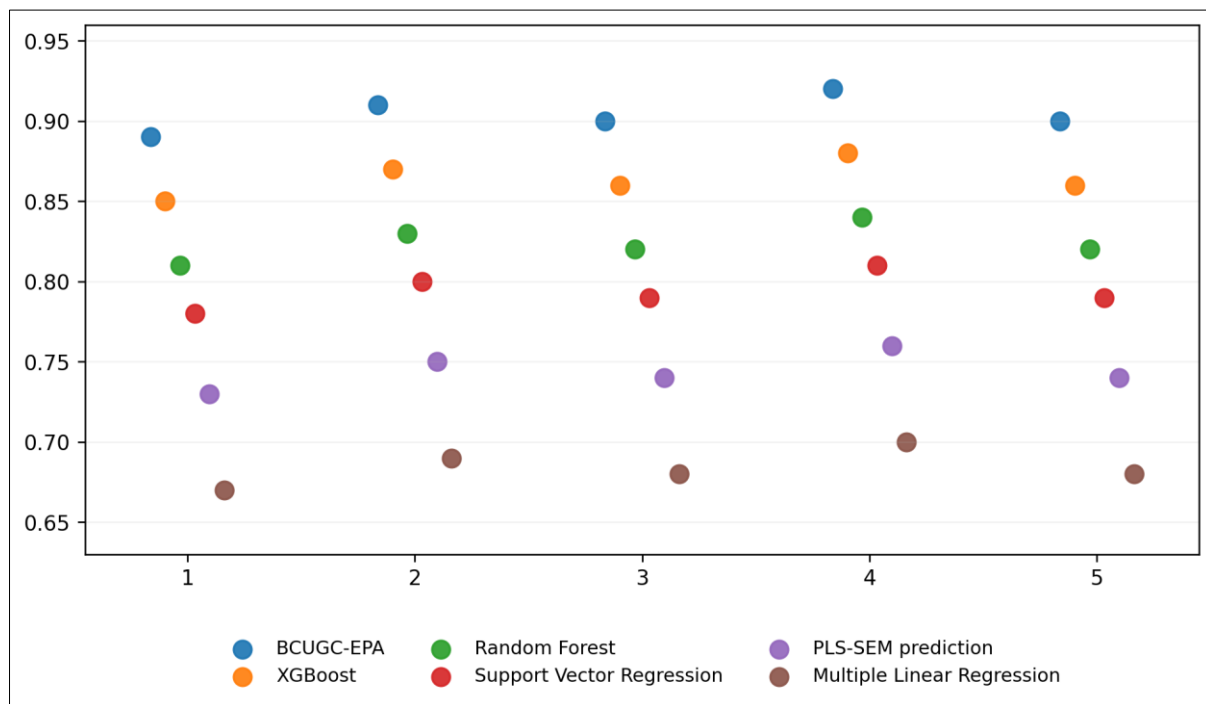
**Fig 8:** Comparative Scatter Plot of Five-Fold Predictive Robustness Across BCUGC-EPA and Benchmark Algorithms

Figure 8 illustrates the predictive robustness of six algorithms across five validation folds. BCUGC-EPA achieves R^2 values of 0.89, 0.91, 0.90, 0.92, and 0.90, producing a mean of 0.904. Standalone XGBoost records an average of 0.864, followed by Random Forest at 0.824 and Support Vector Regression at 0.794. PLS-SEM prediction achieves 0.744, while Multiple Linear Regression records 0.684. All six algorithms exhibit a fold-to-fold R^2 range of 0.03, indicating equivalent absolute dispersion in the illustrative predictions, although BCUGC-EPA maintains the highest accuracy throughout. Its average R^2 exceeds XGBoost by 0.040 and Multiple Linear Regression by 0.220. These comparisons illustrate potential benefits of the hybrid architecture but do not establish superior statistical stability. The hypothetical results require validation using independent observations, controlled preprocessing, and paired uncertainty analysis before informing DTC marketing decisions.

5. Conclusions and Recommendations

5.1. Summary of Findings on Brand Community Participation, User-Generated Content, and Customer-Based Brand Equity

The study developed a proposed analytical framework for investigating the relationships among brand community participation, user-generated content (UGC), customer-based

brand equity (CBBE), repurchase intention, and electronic word-of-mouth (eWOM) within direct-to-consumer (DTC) markets. The illustrative measurement results demonstrated satisfactory internal consistency and convergent validity, with composite reliability coefficients ranging from 0.903 to 0.936 and average variance extracted values between 0.699 and 0.786. Customer-based brand equity exhibited the highest illustrative reliability coefficient of 0.936, supporting its proposed measurement role as an intermediary construct. The structural analysis illustrated positive relationships between brand community participation and brand equity ($\beta = 0.420$), and between UGC and brand equity ($\beta = 0.360$). Furthermore, CBBE demonstrated illustrative direct associations with repurchase intention ($\beta = 0.550$) and eWOM ($\beta = 0.480$). The indirect relationships from community participation through brand equity were estimated at 0.231 for repurchase intention and 0.202 for eWOM, suggesting potential mediation mechanisms requiring empirical confirmation. Comparative predictive illustrations positioned BCUGC-EPA above five benchmark algorithms, with an average R^2 of 0.904, RMSE of 0.300, and MAE of 0.230. SHAP-based illustrations further identified CBBE, community participation, and UGC credibility as potentially influential predictors of consumer behavioural intentions. Collectively, these simulated results demonstrate

how the proposed framework could evaluate community-driven brand equity formation and behavioural prediction. However, the estimates are illustrative rather than empirical findings, and no claims of statistically significant relationships or algorithmic superiority can be established without collected data, fitted models, and independent validation.

5.2. Theoretical and Methodological Contributions of the Proposed BCUGC-EPA Algorithm to Consumer Behaviour Modelling

The principal theoretical contribution of this research lies in developing an integrated framework that connects brand community participation, UGC characteristics, customer-based brand equity, repurchase intention, and electronic word-of-mouth within a unified consumer behaviour model. Unlike approaches that examine brand engagement and post-purchase intentions independently, the proposed framework conceptualises CBBE as a potential mediating mechanism through which community participation and consumer-generated information influence behavioural outcomes. This formulation distinguishes direct relationships from indirect pathways and incorporates conditional interaction effects involving community participation and UGC credibility, authenticity, and information quality. Methodologically, the study proposes BCUGC-EPA as a hybrid architecture combining PLS-SEM, Extreme Gradient Boosting (XGBoost), and Shapley Additive Explanations (SHAP). PLS-SEM provides measurement validation and structural relationship estimation, XGBoost models nonlinear consumer responses, and SHAP facilitates transparent interpretation of individual feature contributions. The algorithmic framework introduces a coordinated estimation procedure that separates theoretical explanation from out-of-sample prediction, thereby reducing the risk of confusing structural significance with predictive accuracy. The illustrative comparison produced an R^2 of 0.904 for BCUGC-EPA against 0.864 for standalone XGBoost, corresponding to an absolute improvement of 0.040. Nevertheless, these figures demonstrate the proposed evaluation framework rather than verified methodological advances. The intended contribution therefore rests on empirical assessment of whether structural feature integration generates incremental predictive information, improves generalisation, and maintains interpretability under controlled validation conditions. This approach provides a reproducible basis for examining the potential usefulness of hybrid analytical systems in contemporary DTC consumer research.

5.3. Strategic Recommendations for Strengthening Repurchase Intention and Electronic Word-of-Mouth Through Community-Driven Marketing

The proposed framework provides a basis for developing targeted marketing strategies that connect brand community engagement and UGC management with customer retention and electronic advocacy. DTC organisations should prioritise meaningful participation rather than relying exclusively on membership counts, posting frequency, or superficial engagement indicators. Community managers can encourage product-related discussions, customer problem-solving, peer assistance, and collaborative product development through moderated digital platforms. These activities should be evaluated using participation quality, response effectiveness, community identification, and subsequent changes in brand

equity indicators. UGC management should emphasise credible customer experiences, transparent review practices, verifiable testimonials, and detailed product information rather than artificially stimulated positive evaluations. The illustrative structural estimates suggest that community participation and UGC may contribute differently to repurchase and advocacy intentions, highlighting the importance of distinct behavioural strategies. BCUGC-EPA could support such differentiation by identifying consumers with low predicted retention despite strong community involvement or high advocacy potential despite limited purchasing frequency. SHAP-based explanations could assist managers in understanding whether brand equity, content credibility, interaction quality, or perceived authenticity contributes most strongly to specific predictions. Marketing interventions should subsequently be evaluated through controlled experiments, longitudinal comparisons, and incremental retention measures rather than assuming predictive associations establish intervention effectiveness. Organisations should also monitor customer satisfaction, complaint resolution, repeat purchase frequency, recommendation activity, and community trust using consistent performance indicators. These recommendations are conditional on empirical validation of the proposed model and should be implemented alongside responsible data governance, consumer privacy protection, transparent communication, and safeguards against algorithmic discrimination.

5.4. Research Limitations, Future Algorithmic Improvements, and Directions for Longitudinal Investigation

The proposed study has several methodological limitations that must be addressed when interpreting subsequent empirical findings and advancing the BCUGC-EPA framework. First, its cross-sectional questionnaire design restricts causal interpretation because brand community participation, UGC perceptions, customer-based brand equity, and behavioural intentions are measured within a common observation period. This approach cannot conclusively determine whether community engagement increases brand equity or whether consumers with stronger existing brand relationships participate more actively. Second, self-reported measures may introduce common method variance, social desirability effects, and discrepancies between stated intentions and actual purchasing behaviour. Third, the generalisability of the proposed algorithm may be affected by differences in DTC product categories, consumer demographics, platform characteristics, and market environments. Fourth, the integration of latent-variable estimation with nonlinear prediction introduces potential risks involving information leakage, unstable feature attribution, overfitting, and algorithmic complexity. Future investigations should incorporate longitudinal consumer panels, verified purchasing histories, behavioural engagement records, and repeated brand equity measurements to evaluate changes over time. Methodological improvements should include nested cross-validation, external validation across independent brands, temporal holdout testing, feature ablation experiments, probability calibration where appropriate, and systematic sensitivity analysis. BCUGC-EPA could also be extended using temporal gradient boosting, causal machine learning, hierarchical modelling, and transformer-based

representations of consented textual UGC. Further research should investigate whether prediction errors and feature contributions differ across consumer segments, community maturity levels, and product categories. Ultimately, the research programme should establish whether integrating structural explanation, nonlinear prediction, and transparent feature attribution produces reproducible improvements in understanding and forecasting consumer relationships within DTC markets.

References

1. Aluso L. Forecasting marketing ROI through cross-platform data integration between HubSpot CRM and Power BI. *International Journal of Scientific Research in Science, Engineering and Technology*. 2021;8(6):356–378. doi:10.32628/IJSRSET214420.
2. Anokwuru EA. Leveraging AI-enhanced commercial insights for precision marketing in the biopharmaceutical industry. *International Journal of Scientific Research and Modern Technology*. 2024;3(9):110–125. doi:10.38124/ijrsmt.v3i9.1204.
3. Anokwuru EA, Omachi A, Enyejo JO. Automation-enabled RFI/RFP market intelligence platforms: Redefining data-driven business development in global pharmaceutical markets. *International Journal of Scientific Research in Science and Technology*. 2024;11(3):1016–1036. doi:10.32628/IJSRST54310301.
4. Ardley B, McIntosh E, McManus J. From transactions to interactions: The value of co-creation processes within online brand consumer communities. *Business Process Management Journal*. 2020;26(4):825–838. doi:10.1108/BPMJ-10-2019-0444.
5. Bigné E, Ruiz-Mafé C, Badenes-Rocha A. The influence of negative emotions on brand trust and intention to share cause-related posts: A neuroscientific study. *Journal of Business Research*. 2023;157:113628. doi:10.1016/j.jbusres.2022.113628.
6. Carlson J, Wyllie J, Rahman MM, Voola R. Enhancing brand relationship performance through customer participation and value creation in social media brand communities. *Journal of Retailing and Consumer Services*. 2019;50:333–341. doi:10.1016/j.jretconser.2018.07.008.
7. Cheng FF, Wu CS, Chen YC. Creating customer loyalty in online brand communities. *Computers in Human Behavior*. 2020;107:105752. doi:10.1016/j.chb.2018.10.018.
8. Chi M, Harrigan P, Xu Y. Customer engagement in online service brand communities. *Journal of Services Marketing*. 2022;36(2):201–216. doi:10.1108/JSM-09-2020-0392.
9. Choi B, Lee I. Trust in open versus closed social media: The relative influence of user- and marketer-generated content in social network services on customer trust. *Telematics and Informatics*. 2017;34(5):550–559. doi:10.1016/j.tele.2016.11.005.
10. Daowd A, Hasan R, Eldabi T, Rafi-ul-Shan PM, Cao D, Kasemsarn N. Factors affecting eWOM credibility, information adoption and purchase intention on Generation Y: A case from Thailand. *Journal of Enterprise Information Management*. 2021;34(3):838–859. doi:10.1108/JEIM-04-2019-0118.
11. De Bruyn A, Viswanathan V, Beh YS, Brock JKU, von Wangenheim F. Artificial intelligence and marketing: Pitfalls and opportunities. *Journal of Interactive Marketing*. 2020;51:91–105. doi:10.1016/j.intmar.2020.04.007.
12. Enyejo JO, Babalola INO, Owolabi FRA, Adeyemi AF, Osam-Nunoo G, Ogwuche AO. Data-driven digital marketing and battery supply chain optimization in the battery-powered aircraft industry through case studies of Rolls-Royce's ACCEL and Airbus's E-Fan X projects. *International Journal of Scholarly Research and Reviews*. 2024;5(2):1–20. doi:10.56781/ijssr.2024.5.2.0045.
13. Folorunso TN, Idowu K, Balogun A. AI-driven financial risk forecasting: A multi-model ensemble approach for enterprise decision intelligence. *Journal of Computational Analysis and Applications*. 2025;34(3). Available from: <https://eudoxuspress.com/index.php/pub/article/view/5080>.
14. Odozor L, Berko SN, Obu AU, Idowu K. Building a resilient U.S. cybersecurity workforce through practical training. *International Journal of Scientific and Management Research*. 2025;8(8):30–36.
15. Hair JF, Risher JJ, Sarstedt M, Ringle CM. When to use and how to report the results of PLS-SEM. *European Business Review*. 2019;31(1):2–24. doi:10.1108/EBR-11-2018-0203.
16. Herhausen D, Bernritter SF, Ngai EWT, Kumar A, Delen D. Machine learning in marketing: Recent progress and future research directions. *Journal of Business Research*. 2024;170:114254. doi:10.1016/j.jbusres.2023.114254.
17. Hu X, Liu A, Li X, Dai Y, Nakao M. Explainable AI for customer segmentation in product development. *CIRP Annals*. 2023;72(1):89–92.
18. Ismagilova E, Slade EL, Rana NP, Dwivedi YK. The effect of electronic word of mouth communications on intention to buy: A meta-analysis. *Information Systems Frontiers*. 2020;22:1203–1226. doi:10.1007/s10796-019-09924-y.
19. Joung J, Kim H. Interpretable machine learning-based approach for customer segmentation for new product development from online product reviews. *International Journal of Information Management*. 2023;70:102641. doi:10.1016/j.ijinfomgt.2023.102641.
20. Kaur H, Paruthi M, Islam J, Hollebeek LD. The role of brand community identification and reward on consumer brand engagement and brand loyalty in virtual brand communities. *Telematics and Informatics*. 2020;46:101321. doi:10.1016/j.tele.2019.101321.
21. Koay KY, Ong DLT, Khoo KL, Yeoh HJ. Perceived social media marketing activities and consumer-based brand equity. *Asia Pacific Journal of Marketing and Logistics*. 2021;33(1):53–72. doi:10.1108/APJML-07-2019-0453.
22. Kumar J, Kumar V. Drivers of brand community engagement. *Journal of Retailing and Consumer Services*. 2020;54:101949. doi:10.1016/j.jretconser.2019.101949.
23. Kumar J, Nayak JK. Brand community relationships transitioning into brand relationships: Mediating and moderating mechanisms. *Journal of Retailing and Consumer Services*. 2018;45:64–73. doi:10.1016/j.jretconser.2018.08.007.
24. Liao SH, Hu DC, Fang YW. Repurchase intention in a

- physical store: Moderated mediating role of electronic word-of-mouth. *International Journal of Retail & Distribution Management*. 2023;51(2):205–222. doi:10.1108/IJRDM-04-2022-0122.
25. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*. 2017;30. Available from: <https://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions>.
 26. Ngai EWT, Wu Y. Machine learning in marketing: A literature review, conceptual framework, and research agenda. *Journal of Business Research*. 2022;145:35–48. doi:10.1016/j.jbusres.2022.03.049.
 27. Ma L, Sun B. Machine learning and AI in marketing: Connecting computing power to human insights. *International Journal of Research in Marketing*. 2020;37(3):481–504. doi:10.1016/j.ijresmar.2020.04.005.
 28. Onofrei G, Filieri R, Kennedy L. Social media interactions, purchase intention, and behavioural engagement: The mediating role of source and content factors. *Journal of Business Research*. 2022;142:100–112. doi:10.1016/j.jbusres.2021.12.031.
 29. Ononiwu M, Azonuche TI, Enyejo JO. Exploring influencer marketing among women entrepreneurs using encrypted CRM analytics and adaptive progressive web app development. *International Journal of Scientific Research and Modern Technology*. 2023;2(6):1–13. doi:10.38124/ijsrmt.v2i6.562.
 30. Ononiwu M, Azonuche TI, Okoh OF, Enyejo JO. AI-driven predictive analytics for customer retention in e-commerce platforms using real-time behavioral tracking. *International Journal of Scientific Research and Modern Technology*. 2023;2(8):17–31. doi:10.38124/ijsrmt.v2i8.561.
 31. Peters C, Bodkin CD. Community in context: Comparing brand communities and retail store communities. *Journal of Retailing and Consumer Services*. 2018;45:1–11. doi:10.1016/j.jretconser.2018.08.001.
 32. Rattanaburi K. Impact of electronic word-of-mouth on consumer-based brand equity and purchase intention: Hotel industry in Thailand. *ABAC Journal*. 2023;43(2). doi:10.14456/abacj.2023.17.
 33. Richter NF, Tudoran AA. Elevating theoretical insight and predictive accuracy in business research: Combining PLS-SEM and selected machine learning algorithms. *Journal of Business Research*. 2024;173:114453. doi:10.1016/j.jbusres.2023.114453.
 34. Schivinski B. Eliciting brand-related social media engagement: A conditional inference tree framework. *Journal of Business Research*. 2021;130:594–602. doi:10.1016/j.jbusres.2019.08.045.
 35. Shmueli G, Sarstedt M, Hair JF, Cheah JH, Ting H, Vaithilingam S, et al. Predictive model assessment in PLS-SEM: Guidelines for using PLSpredict. *European Journal of Marketing*. 2019;53(11):2322–2347. doi:10.1108/EJM-02-2019-0189.
 36. Tan LP. Customer participation, positive electronic word-of-mouth intention and repurchase intention: The mediation effect of online brand community trust. *Journal of Marketing Communications*. 2024;30(7):792–809. doi:10.1080/13527266.2023.2177709.
 37. Upadhyay Y, Tripathi A. Brand wagon effect: How brand equity eclipses the effect of eWoM on purchase intentions—mediating role of review helpfulness. *Journal of Business Research*. 2023;168:114239. doi:10.1016/j.jbusres.2023.114239.
 38. Volkmar G, Fischer PM, Reinecke S. Artificial intelligence and machine learning: Exploring drivers, barriers, and future developments in marketing management. *Journal of Business Research*. 2022;149:599–614. doi:10.1016/j.jbusres.2022.04.007.
 39. Wu J, Fan S, Zhao JL. Community engagement and online word of mouth: An empirical investigation. *Information & Management*. 2018;55(2):258–270. doi:10.1016/j.im.2017.07.002.
 40. Yuan D, Lin Z, Filieri R, Liu R, Zheng M. Managing the product-harm crisis in the digital era: The role of consumer online brand community engagement. *Journal of Business Research*. 2020;115:38–47. doi:10.1016/j.jbusres.2020.04.044.
 41. Zhao Y, Wang L, Tang H, Zhang Y. Electronic word-of-mouth and consumer purchase intentions in social e-commerce. *Electronic Commerce Research and Applications*. 2020;41:100980. doi:10.1016/j.elerap.2020.100980.
 42. Zollo L, Filieri R, Rialti R, Yoon S. Unpacking the relationship between social media marketing and brand equity: The mediating role of consumers' benefits and experience. *Journal of Business Research*. 2020;117:256–267. doi:10.1016/j.jbusres.2020.05.001.