SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Explainable Machine Learning for Modeling Online Impulse-Buying Tendency: A Generational Comparison of Gen Z and Millennials

Yashawant Pathak[1*], Bhupendra Kumar[2], Suhel Ahmed Khan[3], Jagdish Makhijani[4]

[1]Amity University Madhya Pradesh, Gwalior 474005, Madhya Pradesh, India. ORCID: 0009-0000-4385-0281,
E-mail: yashpathak21@gmail.com
[2]Amity University Madhya Pradesh, Gwalior 474005, Madhya Pradesh, India. ORCID: 0000-0002-9609-2311,
E-mail: bkumar@gwa.amity.edu
[3]Rustamji Institute of Technology, BSF Academy, Gwalior 475005, Madhya Pradesh, India. ORCID: 0009-0009-5385-8709,
E-mail: suhel@rjit.ac.in
[4]Rustamji Institute of Technology, BSF Academy, Gwalior 475005, Madhya Pradesh, India. ORCID: 0000-0002-3601-3319,
E-mail: drjagdish@rjit.ac.in
*Corresponding author: Yashawant Pathak, Email: yashpathak21@gmail.com

Abstract
Online impulse-buying tendency is widely studied through explanatory models, but less is known about how well it can be predicted across generational cohorts. This study analyzed 540 adult online shoppers in India (299 Gen Z; 241 Millennials) using a nine-item buying-impulsiveness measure. Internal consistency was good (Cronbach's α = .820), and sampling adequacy was high (KMO = .903), and cohort means did not differ significantly (p = .202). Five regression algorithms were evaluated with nested cross-validation. Ridge regression performed best in the pooled sample (R^2 = .155; RMSE = .647; MAE = .524), while a behavioral-only feature set improved R^2 to .171. SHAP attribution identified discount responsiveness, unplanned browsing, and recommendation exploration as the leading predictive signals. Cohort-specific explanations showed similar core predictors but different secondary rankings. The results favor parsimonious, interpretable prediction over unnecessary model complexity in survey-based consumer analytics.

Keywords: online impulse buying; buying impulsiveness; Generation Z; Millennials; machine learning; explainable artificial intelligence; SHAP

Introduction

Digital retail has compressed the interval between seeing a product and buying it. A consumer can encounter a recommendation, compare alternatives, receive a discount prompt, and complete payment within the same mobile session. This compressed decision path matters for impulse buying because the phenomenon is defined less by the product purchased than by the immediacy and reduced deliberation surrounding the decision (Rook, 1987; Rook & Fisher, 1995). Meta-analytic evidence also shows that impulsive purchasing arises from a mixture of personal tendencies, situational conditions, affect, self-control, and marketing stimuli rather than from a single universal trigger (Amos et al., 2014; Iyer et al., 2020).

Online environments add their own cues to this process. Website characteristics and interface quality can intensify an urge to buy impulsively (Parboteeah et al., 2009; Wells et al., 2011), while virtual atmospherics and mobile browsing can influence spontaneous purchase responses (Floh & Madlberger, 2013; Zheng et al., 2019). More recent work has extended this line of inquiry to online convenience, video commerce, AI-enabled retail services, and conversational agents (Lina et al., 2022; Ngo et al., 2024; Ooi et al., 2025; Shamim & Misra, 2025). These studies make clear that digital impulse buying is embedded in an increasingly adaptive shopping environment.

Generational comparison adds a second layer of complexity. Indian evidence indicates that Gen Z and Millennials differ in shopping orientations, online enthusiasm, exploration, and information use (Agrawal, 2022; Thangavel et al., 2021). Direct research on online impulse buying likewise suggests that generational differences can occur in selected relationships without appearing uniformly across an entire behavioral model (Cavazos-Arroyo & Máynez-Guaderrama, 2022). This makes a simple assumption such as "Gen Z is more impulsive" too coarse for empirical work.

Most studies in this area are explanatory: they specify a theoretical model and test whether particular paths or group differences are statistically significant. Prediction asks a different question—whether information available about a new consumer is sufficient to estimate an outcome outside the sample used to fit the model (Shmueli, 2010). That distinction is relevant to e-commerce analytics, where machine-learning methods are increasingly used for conversion, cart completion, segmentation, and buying-intention problems (Bastos & Bernardes, 2024; Chen, 2025; Esmeli & Gokce, 2025; Kotravel Bharathi & Elakkiyan, 2026). Recent work in IJAIML has also used AI and deep learning for retail consumer-behavior analysis, reinforcing the journal-level relevance of transparent predictive methods in this domain (Saritha et al., 2026).

Model complexity, however, is not synonymous with predictive quality. Survey data are modest in size compared with clickstream or transaction logs and often combine ordinal, categorical, and self-reported features. Flexible algorithms may therefore fit sample-specific variation that does not generalize. A credible comparison should place regularized linear models, kernel methods, bagged trees, and boosting algorithms under the same leakage-resistant validation design rather than assuming that the most complex learner must win (Breiman, 2001; Chen & Guestrin, 2016; Cortes & Vapnik, 1995; Friedman, 2001; Hoerl & Kennard, 1970).

Interpretability is equally important. SHapley Additive exPlanations (SHAP) assign feature-level contributions to model predictions and can be summarized globally or within subgroups (Lundberg & Lee, 2017). Explainable machine learning has already been used to interpret e-commerce conversion, post-cart behavior, and consumer segmentation (Bastos & Bernardes, 2024; Esmeli & Gokce, 2025; Wen et al., 2024). The present study extends that approach to a validated online buying-impulsiveness score and asks whether Gen Z and Millennial consumers display similar average scores but different predictive structures. The paper therefore benchmarks five regression algorithms with nested cross-validation, tests whether broader demographic and shopping-context information improves on a smaller behavioral feature set, and interprets the leading model with SHAP. The emphasis is deliberately predictive rather than causal. Figure 1 summarizes the analytical framework.

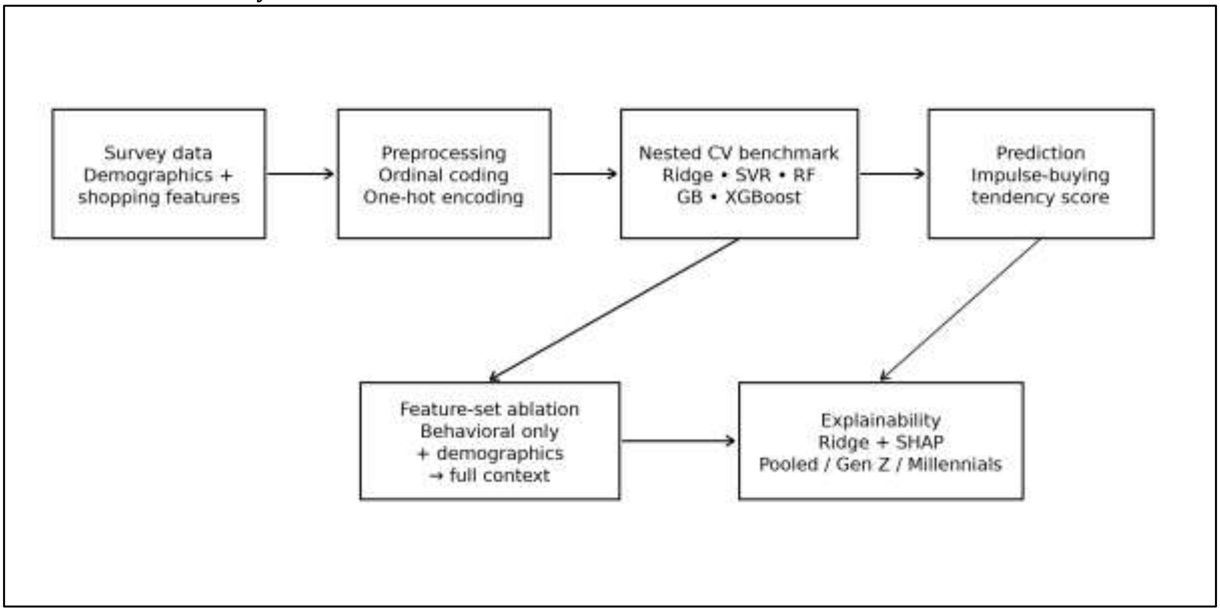

Figure 1. Explainable machine-learning framework used in the study.

1.1 Research Questions

RQ1. Do Gen Z and Millennial consumers differ in their average online impulse-buying tendency?

RQ2. How accurately can online impulse-buying tendency be predicted from observable shopping and demographic features?

RQ3. Does nonlinear model complexity improve out-of-sample performance relative to regularized linear prediction?

RQ4. Does adding demographic and shopping-context information improve generalization beyond behavioral shopping features?

RQ5. Which features make the largest contributions to model predictions, and how do their attribution patterns differ between Gen Z and Millennials?

2. Related Literature

2.1 Online impulse buying and buying-impulsiveness measurement

Rook (1987) shifted impulse-buying research away from treating every unplanned purchase as equivalent and emphasized the subjective experience of a sudden, compelling urge. Rook and Fisher (1995) subsequently operationalized buying impulsiveness as a tendency to buy spontaneously and with limited reflection. Beatty and Ferrell (1998) placed this tendency within a process that also includes available resources, browsing, affect, shopping enjoyment, urge, and purchase. Verplanken and Herabadi (2001) further distinguished cognitive and affective aspects of impulse-buying tendency.

These distinctions are important for the present design. The dependent variable is a psychometric score adapted from the Rook–Fisher buying-impulsiveness scale, not a transaction log of observed impulse purchases. The paper therefore uses the term "impulse-buying tendency" when referring to the modeled target. A separate self-reported count category for recent unplanned purchases is used only as an external behavioral check. Meta-analyses support this distinction between trait-like impulsiveness, situational antecedents, and purchase outcomes (Amos et al., 2014; Iyer et al., 2020).

Digital commerce changes exposure to stimuli rather than eliminating the underlying psychological process. Website and atmospheric cues can influence impulse-buying urge (Floh & Madlberger, 2013; Parboteeah et al., 2009; Wells et al., 2011), while mobile and social-media environments increase opportunities for browsing, product discovery, and rapid purchase completion (Djafarova & Bowes, 2021; Lina et al., 2022; Zheng et al., 2019). Video-commerce research similarly identifies time pressure, quantity pressure, economic benefits, social influence, and audiovisual presentation as relevant stimuli (Ngo et al., 2024).

AI is now part of that retail environment as well. Indian evidence links AI-enabled service quality and personalization with perceived value and subsequent impulse-buying processes (Shamim & Misra, 2025), while chatbot research shows that conversational and social cues can influence trust, rapport, and impulse buying (Ooi et al., 2025). These studies examine AI as a retail stimulus. By contrast, the present paper uses AI/ML as the analytical method for modeling consumer tendency.

2.2 Gen Z and Millennials in online shopping

Generational cohorts are useful as descriptive groupings, but their boundaries should not be treated as natural laws. This study follows the widely used operational boundary that places Millennials in 1981–1996 and Gen Z from 1997 onward within the adult eligible sample (Dimock, 2019). The cohort variable is used for comparison rather than as a deterministic explanation of individual behavior.

Research in India provides a clear reason to make that comparison. Thangavel et al. (2021) found that Millennials and Gen Z differ in online decision-making styles, and Agrawal (2022) reported differences in exploration, enjoyment, deal hunting, and information use. In Central India, consumer-trait research has also linked impulsive purchasing with buying tendency and interpersonal mechanisms (Atulkar & Kesari, 2018).

Evidence specific to generational impulse buying is more nuanced. Cavazos-Arroyo and Máynez-Guaderrama (2022) found cohort differences in selected pathways rather than a universal difference across all relationships. This supports a prediction-oriented possibility: the two generations may have similar average tendency scores while the information that best predicts those scores differs.

2.3 Machine learning and explainable consumer analytics

Ridge regression provides a regularized linear benchmark that is useful when predictors are correlated or numerous relative to the sample size (Hoerl & Kennard, 1970). Support vector regression extends margin-based learning to continuous outcomes (Cortes & Vapnik, 1995). Random forest reduces the instability of individual trees through bootstrap aggregation and randomized feature selection (Breiman, 2001), whereas gradient boosting and XGBoost build additive ensembles sequentially, with XGBoost adding regularization and implementation optimizations (Chen & Guestrin, 2016; Friedman, 2001).

Recent consumer-analytics studies show why benchmarking matters. Bastos and Bernardes (2024) used explainable machine learning to identify drivers of online purchasing, Esmeli and Gokce (2025) compared multiple algorithms for post-cart behavior, and Wen et al. (2024) combined multi-algorithm prediction with SHAP for consumer segmentation. Chen (2025) and Kotravel Bharathi and Elakkiyan (2026) provide further evidence of growing ML use in online-shopping prediction.

SHAP addresses the interpretability problem by expressing a prediction as a baseline plus feature-specific contributions (Lundberg & Lee, 2017). In the present study, SHAP is applied to the best-performing model family rather than used as a substitute for predictive validation. Nested cross-validation separates model selection from outer-fold performance estimation, reducing optimistic bias from tuning on the same observations used for evaluation (Cawley & Talbot, 2010; Kohavi, 1995; Varma & Simon, 2006).

2.4 Research gap

The literature leaves a narrow but useful gap. Online impulse-buying studies are still dominated by explanatory models, whereas e-commerce ML studies usually predict conversion, cart completion, purchase intention, or segmentation rather than a psychometrically screened buying-impulsiveness score. Generational research, meanwhile, generally emphasizes mean or path differences. This paper links the three strands by combining psychometric screening, nested predictive benchmarking, feature-set ablation, and cohort-specific model explanation.

3. Materials and Methods

3.1 Design, setting, and participants

A cross-sectional online survey was conducted in India, predominantly in Central India, between March and July 2026. The questionnaire was administered through Google Forms. The survey link was circulated online, and the author also visited colleges and other institutions to invite eligible respondents from both Gen Z and Millennial cohorts to participate through the online form. Respondents were eligible if they were 18 years or older, were online shoppers, and had purchased at least one product online during the previous month. Electronic informed consent was obtained before the substantive questions were displayed. A total of 540 complete eligible responses were retained (299 Gen Z and 241 Millennials). Because individual selection probabilities were not known, the dataset is treated as a non-probability online sample.

3.2 Measures and target construction

The survey recorded birth year, gender, education, occupation, personal income/allowance, purchase frequency, monthly online spending, browsing duration, number of shopping platforms regularly used, primary shopping device, product category, payment method, responsiveness to discounts or limited-time offers, unplanned browsing, exploration of platform recommendations, and social-media product discovery. Ordinal features were coded in their natural order; nominal variables were one-hot encoded inside the modeling pipeline.

Online buying impulsiveness was measured with nine items adapted to the online-shopping context from Rook and Fisher (1995). Responses ranged from 1 (strongly disagree) to 5 (strongly agree). The planning-oriented item was reverse scored as IB8R = 6 − IB8. The continuous target was the arithmetic mean of IB1–IB7, IB8R, and IB9. A higher score represents stronger online impulse-buying tendency.

A separate question recorded the approximate number of unplanned online purchases during the previous 30 days. Because that measure overlaps conceptually with the modeled target, it was excluded from the predictor set and retained only as an external behavioral check. This conservative choice avoids target leakage.

3.3 Data screening and psychometric assessment

The analytical file was checked for missing observations, duplicate respondent IDs, duplicate full records, invalid scale values, inconsistency between birth year and derived generation, and invariant responding across all nine target items. No case met a prespecified exclusion condition, so all 540 eligible observations were retained. Internal consistency was evaluated with Cronbach's alpha and corrected item–total correlations. Kaiser–Meyer–Olkin (KMO) sampling adequacy, Bartlett's test of sphericity, and correlation-matrix eigenvalues were used as additional checks on whether a single composite score was reasonable.

3.4 Predictive features and preprocessing

The primary model excluded respondent ID, the nine target items, the reverse-scored duplicate item, the composite score itself, and the recent-unplanned-purchase validation item. Birth year was retained as a numeric variable in pooled prediction; generation was derived from birth year for subgroup analyses rather than duplicated as a second pooled predictor. Numeric and ordinal features were median-imputed when necessary and standardized. Nominal variables were imputed with the most frequent category and one-hot encoded. Every preprocessing operation was fitted within the relevant training fold.

3.5 Nested model benchmarking

Five regression algorithms were benchmarked: ridge regression, support vector regression (SVR), random forest regression, gradient boosting regression, and XGBoost (Breiman, 2001; Chen & Guestrin, 2016; Cortes & Vapnik, 1995; Friedman, 2001; Hoerl & Kennard, 1970). The outer loop used five shuffled folds (random seed 42) for performance estimation, and the inner loop used three shuffled folds for hyperparameter selection. Inner-loop selection minimized RMSE. Model comparison used outer-fold R^2, MAE, and RMSE. This nested structure ensures that preprocessing and tuning decisions are made without access to the corresponding outer test fold.

The search spaces are reported in Table 1. Analyses were implemented in Python 3.13.5 using pandas 2.2.3, NumPy 2.3.5, SciPy 1.17.0, scikit-learn 1.8.0, XGBoost 3.1.3, and SHAP 0.50.0. Scikit-learn provided the preprocessing and validation framework (Pedregosa et al., 2011).

Table 1. Hyperparameter search spaces used in the nested inner loop.

Model	Search space
Ridge regression	alpha ∈ {0.01, 0.1, 1, 10, 100}
SVR	C ∈ {0.1, 1, 10}; epsilon ∈ {0.05, 0.1, 0.2}; gamma ∈ {scale, 0.01, 0.1}
Random forest	n_estimators = 200; max_depth ∈ {None, 5}; min_samples_leaf ∈ {2, 5}; max_features = sqrt
Gradient boosting	n_estimators ∈ {100, 200}; learning_rate ∈ {0.03, 0.05}; max_depth ∈ {1, 2}; min_samples_leaf ∈ {3, 10}

Model	Search space
XGBoost	n_estimators ∈ {100, 200}; max_depth ∈ {2, 3}; learning_rate ∈ {0.03, 0.05}; subsample ∈ {0.8, 1.0}; colsample_bytree ∈ {0.8, 1.0}; reg_lambda ∈ {1, 5}

3.6 Feature-set ablation and cohort analysis

To test whether broader respondent information improved prediction, the best-performing model family was re-estimated with three nested feature sets: (1) behavioral shopping features only; (2) behavioral plus demographic features; and (3) the full set including device, category, and payment context. The same outer and inner folds were used. Ridge models were also fitted separately within Gen Z and Millennial subsets to examine cohort-specific predictive structure. Because subgroup samples differ in size and outcome variance, differences in their cross-validated metrics are interpreted descriptively rather than as formal evidence that one generation is intrinsically more predictable.

3.7 Explainability

SHAP values were computed for the fitted ridge models using a linear SHAP explainer. For one-hot encoded categorical variables, absolute contributions were aggregated back to the original questionnaire-variable family. Mean absolute SHAP values were used for global ranking. Separate Gen Z and Millennial models were interpreted to compare attribution patterns. These attributions describe how a fitted predictive model uses the available features; they are not causal effects (Lundberg & Lee, 2017).

3.8 Statistical analysis

Descriptive statistics summarized the sample and target distribution. Welch's independent-samples t test compared Gen Z and Millennials because it does not require equal variances, and Cohen's d summarized effect size. Spearman correlations were used for ordinal shopping features. The recent unplanned-purchase item was correlated with the target only as a behavioral validity check. All inferential tests were two-sided with $\alpha = .05$; they supplement rather than replace out-of-sample model evaluation.

4. Results

4.1 Sample and scale properties

The sample comprised 299 Gen Z respondents (55.4%) and 241 Millennials (44.6%). Women accounted for 52.4% and men for 42.6% of the sample; the remaining respondents selected another gender option or preferred not to state it. Undergraduate and postgraduate degree holders formed the largest education groups. Table 2 gives the principal sample profile.

Table 2. Selected respondent characteristics (N = 540).

Variable	Category	n	%
Generation	Gen Z	299	55.4
	Millennials	241	44.6
Gender	Female	283	52.4
	Male	230	42.6
	Other / self-describe	14	2.6
	Prefer not to say	13	2.4
Education	Undergraduate degree	218	40.4
	Postgraduate degree	175	32.4
	Diploma / certificate	66	12.2
	Higher secondary / school	39	7.2
	Doctoral degree	25	4.6
	Other	17	3.1

The nine-item target showed good internal consistency ($\alpha = 0.820$); reliability was similar in Gen Z ($\alpha = 0.818$) and Millennials ($\alpha = 0.822$). Corrected item–total correlations ranged from .442 to .562, and no item deletion improved alpha beyond the complete scale. KMO was 0.903, Bartlett's test was significant, $\chi^2(36) = 1104.09$, $p < .001$, and the first correlation-matrix eigenvalue was 3.699, with all remaining eigenvalues below 1. These results support use of a single composite tendency score (Tables 3 and 4).

Table 3. Scale-level psychometric checks.

Statistic	Overall	Gen Z	Millennials
Cronbach's α	0.820	0.818	0.822
KMO	0.903	—	—
Bartlett χ^2 (df = 36)	1104.09, p < .001	—	—

Statistic	Overall	Gen Z	Millennials
First eigenvalue	3.699	—	—

Table 4. Item diagnostics for the nine-item target.

Item	Mean	SD	Corrected item–total r	α if deleted
IB1	2.985	1.136	0.547	0.798
IB2	2.998	1.152	0.517	0.802
IB3	2.959	1.116	0.562	0.797
IB4	2.939	1.084	0.493	0.805
IB5	2.957	1.079	0.485	0.806
IB6	2.872	1.081	0.521	0.802
IB7	2.948	1.102	0.559	0.797
IB8R	2.976	1.142	0.442	0.811
IB9	2.944	1.100	0.548	0.798

4.2 Generational mean comparison

Gen Z had a mean score of 2.988 (SD = .697), compared with 2.910 (SD = .727) for Millennials. The mean difference was small and not statistically significant, Welch's $t(504.46) = 1.276$, $p = 0.202$, 95% CI [-0.043, 0.200], Cohen's $d = 0.111$. The substantial overlap visible in Figure 2 argues against a simple claim that one cohort is uniformly more impulse-prone.

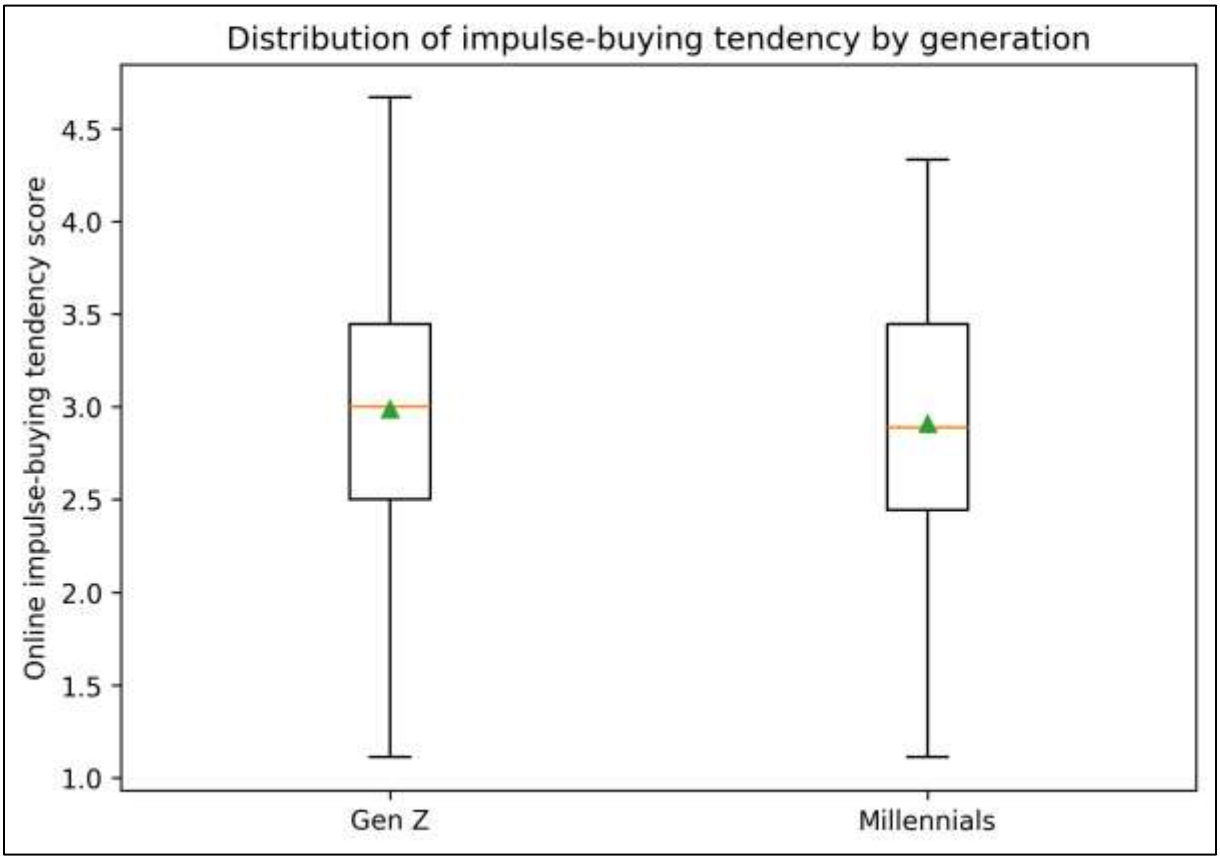

Figure 2. Distribution of online impulse-buying tendency scores by generation.

4.3 Behavioral associations

The external unplanned-purchase indicator correlated with the composite tendency score (Spearman's ρ = .337, $p < .001$). Among variables eligible for prediction, discount responsiveness showed the strongest bivariate association (ρ = .268), followed by recommendation exploration (ρ = .218) and unplanned browsing (ρ = .215). Monthly spending and purchase frequency were nearly unrelated to the target. The full set of exploratory correlations is reported in Table 5.

Table 5. Spearman associations with impulse-buying tendency.

Feature	ρ overall	p	ρ Gen Z	p	ρ Millennials	p
Purchase frequency	0.039	.362	0.007	.897	0.075	.246
Monthly spending	-0.006	.892	-0.042	.484	0.038	.561
Browsing time	0.128	.003	0.133	.022	0.125	.052
Platform count	-0.075	.082	-0.076	.188	-0.071	.274
Discount responsiveness	0.268	< .001	0.329	< .001	0.207	.001
Unplanned browsing	0.215	< .001	0.190	.001	0.238	< .001
Recommendation exploration	0.218	< .001	0.225	< .001	0.215	.001
Social-media discovery	0.097	.024	0.052	.373	0.145	.025
Unplanned purchases (validation only)	0.337	< .001	0.343	< .001	0.327	< .001

4.4 Nested predictive benchmark

Ridge regression produced the best pooled generalization: mean outer-fold R^2 = .155, MAE = .524, and RMSE = .647. SVR ranked second by R^2 (.136), followed by XGBoost (.112), gradient boosting (.111), and random forest (.076). The advantage of the regularized linear model is modest but consistent with the main technical conclusion: added nonlinear flexibility did not improve prediction for this survey dataset (Table 6; Figure 3).

Table 6. Pooled nested cross-validation performance (mean ± SD across five outer folds).

Model	R^2	MAE	RMSE
Ridge regression	0.155 ± 0.071	0.524 ± 0.030	0.647 ± 0.043
Support vector regression	0.136 ± 0.105	0.526 ± 0.025	0.653 ± 0.039
XGBoost	0.112 ± 0.053	0.541 ± 0.037	0.664 ± 0.049
Gradient boosting	0.111 ± 0.042	0.543 ± 0.039	0.665 ± 0.049
Random forest	0.076 ± 0.034	0.555 ± 0.036	0.678 ± 0.046

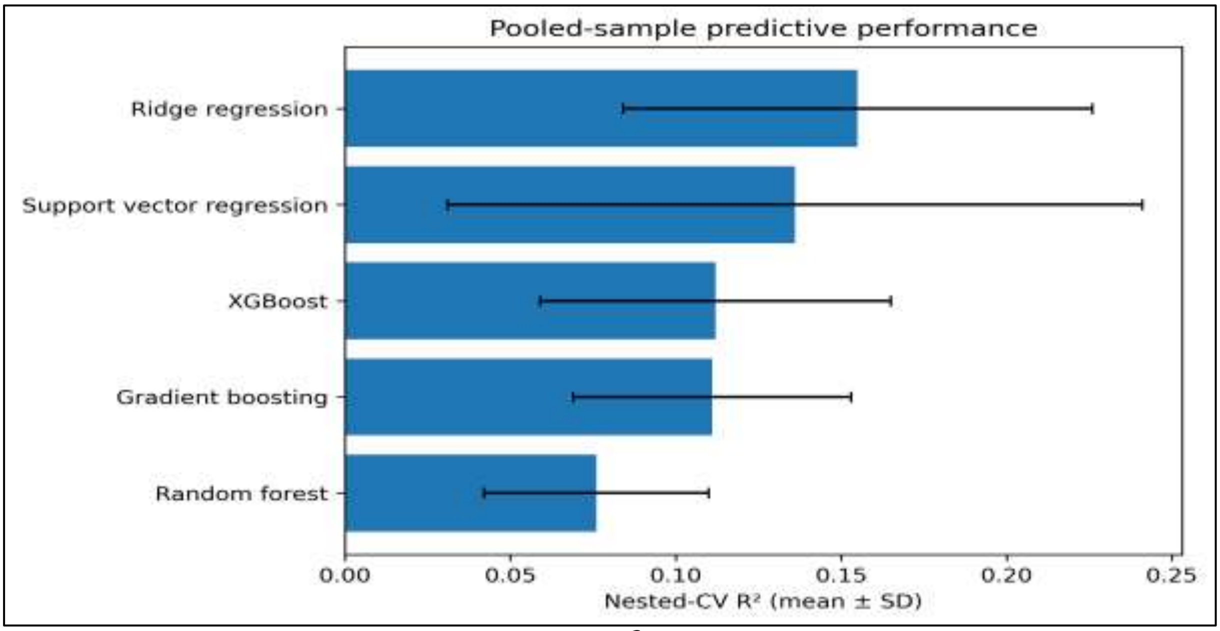

Figure 3. Pooled-sample nested cross-validation R^2 by algorithm. Error bars show SD across the five outer folds.

4.5 Feature-set ablation

The ablation experiment strengthened that conclusion. Restricting ridge regression to the eight behavioral shopping variables increased mean R^2 to .171 and reduced RMSE to .641. Adding demographics reduced R^2 to .155, and the complete feature set remained at .155. Thus, the broader feature space added descriptive information but no measurable predictive benefit under the locked validation procedure (Table 7; Figure 4).

Table 7. Ridge feature-set ablation under nested validation.

Feature set	R^2	MAE	RMSE
Behavioral features only	0.171 ± 0.072	0.522 ± 0.035	0.641 ± 0.045
Behavioral + demographic	0.155 ± 0.067	0.524 ± 0.032	0.647 ± 0.044
Full feature set	0.155 ± 0.071	0.524 ± 0.030	0.647 ± 0.043

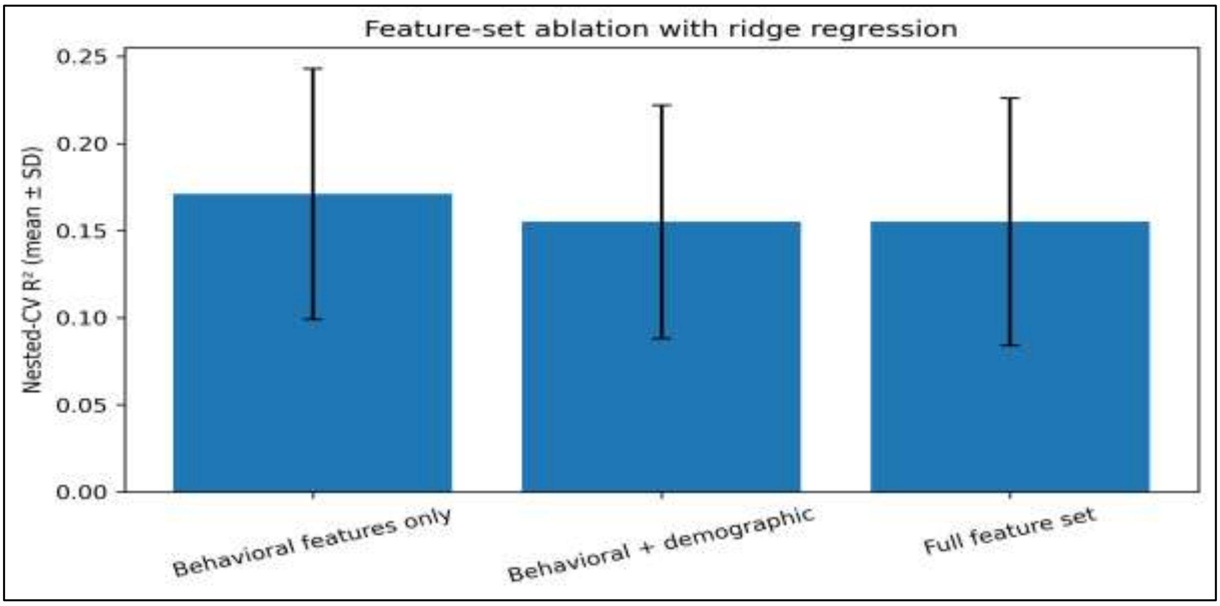

Figure 4. Feature-set ablation for ridge regression under the same nested validation protocol.

4.6 Cohort-specific ridge models

When the same ridge procedure was fitted separately by cohort, mean R^2 was .135 for Gen Z and .120 for Millennials; corresponding RMSE values were .635 and .676. These metrics should not be treated as a formal head-to-head test because the subgroup distributions and sample sizes differ. They do show, however, that the regularized model retained nonzero out-of-sample signal in both groups (Table 8).

Table 8. Cohort-specific ridge performance.

Dataset	n	R^2	MAE	RMSE
Pooled	540	0.155 ± 0.071	0.524 ± 0.030	0.647 ± 0.043
Gen Z	299	0.135 ± 0.067	0.516 ± 0.041	0.635 ± 0.044
Millennials	241	0.120 ± 0.065	0.553 ± 0.066	0.676 ± 0.061

4.7 SHAP attribution

The pooled ridge explanation placed discount responsiveness first (mean |SHAP| = .132), followed by unplanned browsing (.128) and recommendation exploration (.108). Browsing time and social-media discovery formed a second tier, while birth year contributed comparatively little (Table 9; Figure 5).

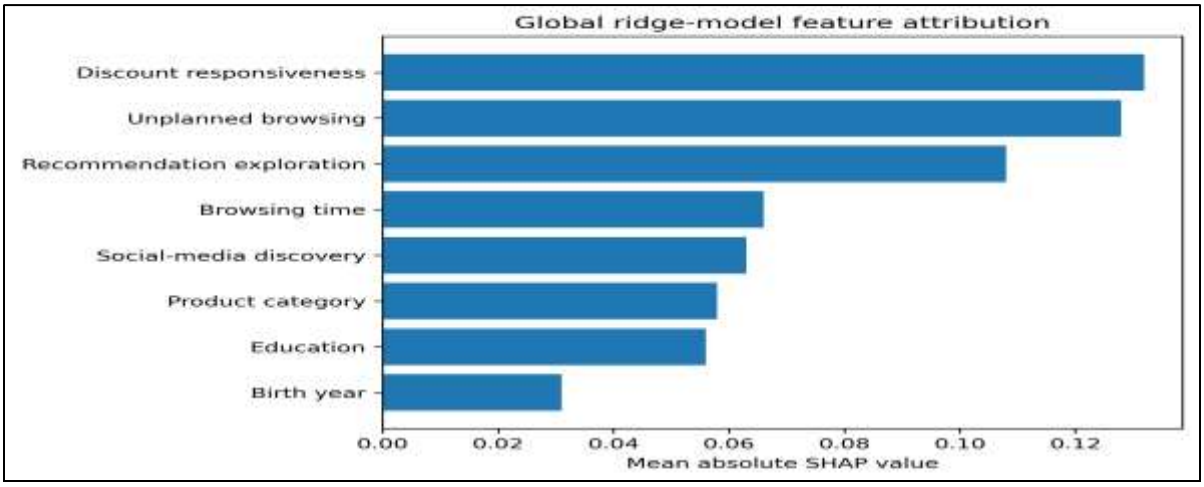

Figure 5. Global mean absolute SHAP attribution from the pooled ridge model.

The cohort models retained the same broad behavioral core but changed its order. Discount responsiveness was strongest for Gen Z (.138), whereas unplanned browsing was strongest for Millennials (.117). Social-media discovery also had a larger mean absolute attribution among Millennials (.070) than Gen Z (.039). Because these are model attributions from separately fitted cross-sectional samples, the differences are best viewed as descriptive signals for future hypothesis testing rather than causal generational effects (Figure 6).

Table 9. Mean absolute SHAP attribution for selected original variables.

Feature	**Pooled**	**Gen Z**	**Millennials**
Discount responsiveness	0.132	0.138	0.088
Unplanned browsing	0.128	0.095	0.117
Recommendation exploration	0.108	0.097	0.087
Social-media discovery	0.063	0.039	0.070
Browsing time	0.066	0.059	0.052
Education	0.056	0.041	0.057
Product category	0.058	0.045	0.038

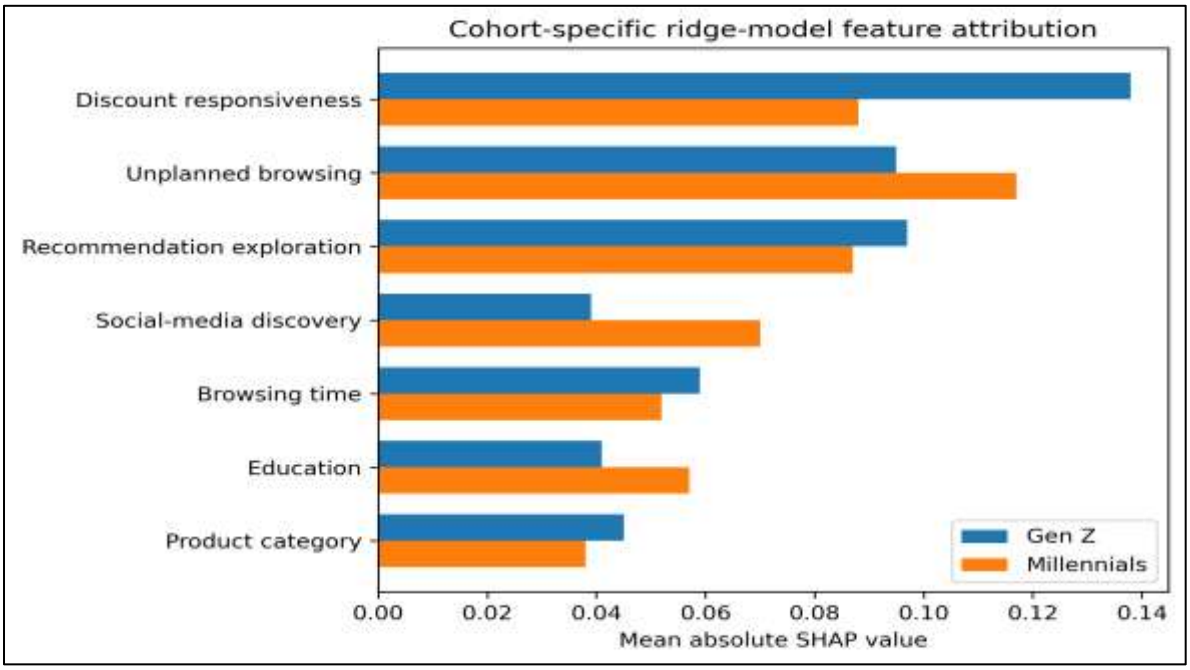

Figure 6. Mean absolute SHAP attribution from separately fitted Gen Z and Millennial ridge models.

5. Discussion

5.1 Comparable means, different predictive emphasis

The absence of a significant mean difference is a useful result in its own right. Gen Z scored only .079 points above Millennials, with a small standardized effect. This does not contradict prior cohort research; rather, it fits evidence that generations can differ in shopping styles and in selected structural relationships without differing on every downstream outcome (Agrawal, 2022; Cavazos-Arroyo & Máynez-Guaderrama, 2022; Thangavel et al., 2021). Treating generation as a deterministic proxy for impulsiveness would therefore be difficult to justify from these data.

The model explanations tell a more specific story. Discount responsiveness, unplanned browsing, and recommendation exploration carried the largest predictive contributions. These features are consistent with the established role of browsing and situational cues in impulse-buying processes (Beatty & Ferrell, 1998; Floh & Madlberger, 2013; Iyer et al., 2020; Parboteeah et al., 2009). Recent digital studies also highlight the relevance of convenience, promotional pressure, social influence, and personalized environments (Lina et al., 2022; Ngo et al., 2024; Shamim & Misra, 2025).

5.2 Why the simpler model won

The best model was not the most complex one. Ridge regression outperformed tree ensembles and XGBoost under nested validation, and the behavior-only feature set outperformed larger input sets. This is not evidence against nonlinear machine learning. Session-level and transaction-level e-commerce datasets can contain much richer nonlinear structure and far larger sample sizes (Bastos & Bernardes, 2024; Esmeli & Gokce, 2025). In a 540-person survey, however, regularization appears to have offered a better bias–variance trade-off. The result reinforces the value of benchmarking rather than selecting an algorithm because of its reputation or novelty.

The modest R^2 also deserves a restrained interpretation. Buying impulsiveness is partly psychological, and the current predictors intentionally emphasize observable shopping routines instead of a large battery of latent antecedents. The model therefore captures a useful but limited slice of the behavior. Prediction and explanation should not be conflated: a variable can be theoretically meaningful without producing a large gain in out-of-sample accuracy (Shmueli, 2010).

5.3 Generational interpretation and practical relevance

The subgroup attributions suggest that the core predictive features are shared, but their priorities are not identical. Gen Z showed a stronger attribution for discount responsiveness, which is compatible with prior reports of greater deal-hunting and exploratory orientation in this cohort (Agrawal, 2022). Millennials showed comparatively larger attributions for unplanned browsing and social-media discovery. These differences are exploratory and should be replicated with a second sample before being treated as stable cohort characteristics.

For e-commerce analytics, the practical implication is not to target a generation simply because it is assumed to be more impulsive. Behavioral context is more informative than the cohort label itself. From a consumer-protection perspective, the same result also argues for caution: explainable models should be used to audit promotional systems and identify potentially manipulative patterns, not merely to intensify pressure on consumers who appear susceptible to spontaneous purchasing.

5.4 Contribution

The contribution is methodological as well as substantive. The paper separates three questions that are often merged in consumer-behavior studies: whether group means differ, whether an individual score can be predicted out of sample, and which observed variables a predictive model actually uses. By combining psychometric screening, nested model selection, ablation, and SHAP, the analysis shows that these questions can lead to different conclusions. Here, the cohorts look similar in average tendency, prediction is modest but reproducible, and the strongest predictive information is concentrated in a small set of shopping behaviors.

6. Conclusion

Using survey data from 540 adult online shoppers in India, this study evaluated whether online impulse-buying tendency can be modeled with interpretable machine learning and whether predictive patterns differ between Gen Z and Millennials. The nine-item target showed good reliability and unidimensional adequacy. Mean tendency scores were similar across cohorts. Under nested cross-validation, ridge regression generalized better than SVR, random forest, gradient boosting, and XGBoost, and a behavior-only feature set performed better than broader demographic-context models. SHAP identified discount responsiveness, unplanned browsing, and recommendation exploration as the principal predictive signals. The findings favor parsimonious modeling over unnecessary algorithmic complexity. They also show why generational comparison should move beyond a simple ranking of mean impulsiveness. Similar group averages can coexist with different feature-priority patterns. Because the data are cross-sectional and self-reported, these patterns should be interpreted as predictive associations rather than causal mechanisms.

6.1 Limitations and future research

The sample is non-probability and concentrated in Central India, which limits population-level generalization. The target is a self-reported tendency scale rather than verified transaction behavior, and the design is cross-sectional. Formal measurement-invariance testing was not conducted, although reliability was similar across cohorts and the overall scale showed a strong one-factor pattern. Predictive performance was modest, which is consistent with the intentionally compact feature set. Future studies should replicate the model with independent samples, clickstream or transaction data, longitudinal designs, prespecified interaction tests, and richer psychological features such as self-control, affect, hedonic motivation, and situational urgency.

References

1. Agrawal, D. K. (2022). Determining behavioural differences of Y and Z generational cohorts in online shopping. International Journal of Retail & Distribution Management, 50(7), 880–895. https://doi.org/10.1108/IJRDM-12-2020-0527
2. Amos, C., Holmes, G. R., & Keneson, W. C. (2014). A meta-analysis of consumer impulse buying. Journal of Retailing and Consumer Services, 21(2), 86–97. https://doi.org/10.1016/j.jretconser.2013.11.004
3. Atulkar, S., & Kesari, B. (2018). Impulse buying: A consumer trait prospective in context of Central India. Global Business Review, 19(2), 477–493. https://doi.org/10.1177/0972150917713546
4. Bastos, J. A., & Bernardes, M. I. (2024). Understanding online purchases with explainable machine learning. Information, 15(10), 587. https://doi.org/10.3390/info15100587
5. Beatty, S. E., & Ferrell, M. E. (1998). Impulse buying: Modeling its precursors. Journal of Retailing, 74(2), 169–191. https://doi.org/10.1016/S0022-4359(99)80092-X
6. Breiman, L. (2001). Random forests. Machine Learning, 45, 5–32. https://doi.org/10.1023/A:1010933404324
7. Cavazos-Arroyo, J., & Máynez-Guaderrama, A. I. (2022). Antecedents of online impulse buying: An analysis of gender and Centennials' and Millennials' perspectives. Journal of Theoretical and Applied Electronic Commerce Research, 17(1), 122–137. https://doi.org/10.3390/jtaer17010007
8. Cawley, G. C., & Talbot, N. L. C. (2010). On over-fitting in model selection and subsequent selection bias in performance evaluation. Journal of Machine Learning Research, 11, 2079–2107.
9. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785–794). Association for Computing Machinery. https://doi.org/10.1145/2939672.2939785
10. Chen, X. (2025). Consumer online shopping behavior prediction based on machine learning algorithm. Procedia Computer Science, 262, 1395–1401. https://doi.org/10.1016/j.procs.2025.05.187
11. Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20, 273–297. https://doi.org/10.1007/BF00994018
12. Dimock, M. (2019, January 17). Defining generations: Where Millennials end and Generation Z begins. Pew Research Center.
13. Djafarova, E., & Bowes, T. (2021). 'Instagram made me buy it': Generation Z impulse purchases in fashion industry. Journal of Retailing and Consumer Services, 59, 102345. https://doi.org/10.1016/j.jretconser.2020.102345
14. Esmeli, R., & Gokce, A. (2025). An analysis of consumer purchase behavior following cart addition in e-commerce utilizing explainable artificial intelligence. Journal of Theoretical and Applied Electronic Commerce Research, 20(1), 28. https://doi.org/10.3390/jtaer20010028
15. Floh, A., & Madlberger, M. (2013). The role of atmospheric cues in online impulse-buying behavior. Electronic Commerce Research and Applications, 12(6), 425–439. https://doi.org/10.1016/j.elerap.2013.06.001
16. Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. The Annals of Statistics, 29(5), 1189–1232. https://doi.org/10.1214/aos/1013203451
17. Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. Technometrics, 12(1), 55–67. https://doi.org/10.1080/00401706.1970.10488634
18. Iyer, G. R., Blut, M., Xiao, S. H., & Grewal, D. (2020). Impulse buying: A meta-analytic review. Journal of the Academy of Marketing Science, 48(3), 384–404. https://doi.org/10.1007/s11747-019-00670-w
19. Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In Proceedings of the 14th International Joint Conference on Artificial Intelligence (Vol. 2, pp. 1137–1143).
20. Kotravel Bharathi, C. K., & Elakkiyan, K. (2026). Leveraging artificial intelligence for predictive

modelling of consumer buying intentions on e-commerce platforms. Discover Artificial Intelligence, 6, 108. https://doi.org/10.1007/s44163-025-00815-7
21. Lina, Y., Hou, D., & Ali, S. (2022). Impact of online convenience on Generation Z online impulsive buying behavior: The moderating role of social media celebrity. Frontiers in Psychology, 13, 951249. https://doi.org/10.3389/fpsyg.2022.951249
22. Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In Advances in Neural Information Processing Systems 30 (pp. 4765–4774).
23. Ngo, T. T. A., Nguyen, H. L. T., Nguyen, H. P., Mai, H. T. A., Mai, T. H. T., & Hoang, P. L. (2024). A comprehensive study on factors influencing online impulse buying behavior: Evidence from Shopee video platform. Heliyon, 10(15), e35743. https://doi.org/10.1016/j.heliyon.2024.e35743
24. Ooi, M.-Y., Lo, P.-S., Aw, E. C.-X., Dastane, O., & Tan, G. W.-H. (2025). From chat to cart: How AI boosts online impulse buying. Journal of Computer Information Systems. Advance online publication. https://doi.org/10.1080/08874417.2025.2576214
25. Parboteeah, D. V., Valacich, J. S., & Wells, J. D. (2009). The influence of website characteristics on a consumer's urge to buy impulsively. Information Systems Research, 20(1), 60–78. https://doi.org/10.1287/isre.1070.0157
26. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825–2830.
27. Rook, D. W. (1987). The buying impulse. Journal of Consumer Research, 14(2), 189–199. https://doi.org/10.1086/209105
28. Rook, D. W., & Fisher, R. J. (1995). Normative influences on impulsive buying behavior. Journal of Consumer Research, 22(3), 305–313. https://doi.org/10.1086/209452
29. Saritha, G., Kumar, Y., Laxmaiah, B., CK, R., Sumathi, S., & Swapna, P. K. (2026). Enhancing consumer behavior analysis in retail using hybrid deep learning models. International Journal of Artificial Intelligence and Machine Learning, 6(4s), 779–791. https://doi.org/10.51483/IJAIML.6.4s.2026.779-791
30. Shamim, K., & Misra, S. C. (2025). Leveraging AI to drive online impulse buying: A SOBC perspective. International Journal of Retail & Distribution Management, 53(10–11), 1040–1056. https://doi.org/10.1108/IJRDM-07-2024-0334
31. Shmueli, G. (2010). To explain or to predict? Statistical Science, 25(3), 289–310. https://doi.org/10.1214/10-STS330
32. Thangavel, P., Pathak, P., & Chandra, B. (2021). Millennials and Generation Z: A generational cohort analysis of Indian consumers. Benchmarking: An International Journal, 28(7), 2157–2177. https://doi.org/10.1108/BIJ-01-2020-0050
33. Varma, S., & Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. BMC Bioinformatics, 7, 91. https://doi.org/10.1186/1471-2105-7-91
34. Verplanken, B., & Herabadi, A. (2001). Individual differences in impulse buying tendency: Feeling and no thinking. European Journal of Personality, 15(S1), S71–S83. https://doi.org/10.1002/per.423
35. Wells, J. D., Parboteeah, D. V., & Valacich, J. S. (2011). Online impulse buying: Understanding the interplay between consumer impulsiveness and website quality. Journal of the Association for Information Systems, 12(1), 32–56. https://doi.org/10.17705/1jais.00254
36. Wen, Z., Lin, W., & Liu, H. (2024). Consumer segmentation and decision: Explainable machine learning insights. Journal of Computer Information Systems, 66(4), 422–438. https://doi.org/10.1080/08874417.2024.2386540
37. Zheng, X., Men, J., Yang, F., & Gong, X. (2019). Understanding impulse buying in mobile commerce: An investigation into hedonic and utilitarian browsing. International Journal of Information Management, 48, 151–160. https://doi.org/10.1016/j.ijinfomgt.2019.02.010