THE EFFECTS OF PRODUCT-DATA QUALITY AND BRAND FAMILIARITY ON AI-BASED PRODUCT SELECTION: A CONTROLLED EXPERIMENT USING INDIAN MSME PRODUCT PROFILES

Dr. Rajesh Kumar K V[1], Varshalatha Galeti[2], Sahithi Vuskebhai[3], and Mylapally Siri[4]

[1] *Assistant Professor, Academics – Business, Woxsen University, Hyderabad, Telangana, India.*

[2] *MBA Student, School of Business, Woxsen University, Hyderabad, Telangana, India.*

[3] *MBA Student, School of Business, Woxsen University, Hyderabad, Telangana, India.*

[4] *MBA Student, School of Business, Woxsen University, Hyderabad, Telangana, India.*

To Cite this Article

Taqi Mohammed Khattab Al-Rubaye, "Using Statistical Methods in Electrical Engineering to improve Detection Algorithms for UAVs", Journal of Science Engineering Technology and Management Science, Vol. 03, Issue 11, November 2026, pp: 26-38, DOI: http://doi.org/10.64771/jsetms.2026.v03.i11.pp26-38

Submitted: 24-09-2026 *Accepted: 29-10-2026* *Published: 04-11-2026*

Abstract

Artificial intelligence (AI) is reshaping product discovery in digital marketplaces by enabling the automated evaluation, comparison and recommendation of competing products. However, the products selected by AI systems may depend not only on their underlying attributes but also on the quality of the product information provided and the brand-related knowledge associated with them. These issues are particularly relevant to Indian micro, small and medium enterprises (MSMEs), whose products may compete with established brands despite differences in brand recognition and the completeness of their digital product information. Understanding how these factors influence AI-based product selection is therefore important for examining the representation of MSME products in AI-driven digital commerce.

This study proposes a controlled experiment to investigate the individual and interactive effects of product-data quality and brand familiarity on AI model-based product selection. A 2 × 2 factorial experimental design will be employed, varying product-data quality (high versus low) and brand familiarity (familiar versus unfamiliar) across controlled product profiles representing Indian MSMEs. The profiles will be evaluated by a specified AI model using standardized prompts, consistent selection criteria and predefined experimental conditions. Repeated trials will be conducted to record product-selection outcomes and estimate selection frequencies across the four experimental conditions. Appropriate statistical techniques for categorical outcomes and repeated observations will be used to assess the main effects of both experimental factors and their interaction. Model specifications, prompt configurations, product attributes and trial procedures will be documented to facilitate transparency and reproducibility.

By examining AI model behaviour directly, this study seeks to extend research on information quality and algorithmic product selection by investigating how product-data characteristics and brand-related cues jointly shape AI selection patterns. The study is expected to contribute to a better understanding of the factors associated with product selection in AI-mediated digital environments. Its findings may also provide practical insights for Indian MSMEs into improving product-data completeness, consistency and brand representation when preparing their products for AI-driven discovery and recommendation systems.

Keywords: Artificial intelligence; product-data quality; brand familiarity; AI-based product selection; controlled experiment; algorithmic decision-making; Indian MSMEs; digital marketplaces.

This is an open access article under the creative commons license https://creativecommons.org/licenses/by-nc-nd/4.0/

CC BY-NC-ND 4.0

Introduction

Artificial intelligence (AI) is becoming an important part of digital commerce, influencing how products are discovered, compared and recommended to potential buyers. Large language models (LLMs) and AI-based recommendation systems can process product descriptions, interpret product attributes and identify alternatives that appear to match a specified set of requirements. These capabilities create new possibilities for digital marketplaces, but they also raise an important question: what determines which products an AI system selects when several alternatives appear suitable? The answer may depend not only on the characteristics of the products themselves but also on how their information is presented and the associations the model has with their brands. Understanding these influences is important because AI-generated selections may shape which products receive attention during digital product discovery.

Product information provides the foundation for evaluating and comparing competing products. Details such as price, material, dimensions, functionality, durability and intended use help distinguish one product from another. When these attributes are accurate, complete and presented consistently, an AI model has a clearer basis for assessing the available alternatives. In contrast, missing specifications, inconsistent descriptions or ambiguous information may make comparisons more difficult. Recent research on large language models and e-commerce catalogues demonstrates that product-attribute quality is an important technical concern in automated product assessment (Satyadharma et al., 2025). However, the ability of a model to interpret product information does not necessarily mean that it will select the most suitable product when information quality varies across alternatives. This distinction makes it necessary to examine product-data quality not merely as a characteristic of an online listing but as a factor that may influence the selection behaviour of an AI model.

Brand familiarity presents a related but conceptually different issue. Established brands may be associated with a greater volume of publicly available descriptions, reviews, specifications and other information. An AI model may have encountered information about some brands more frequently during training or through its accessible knowledge sources than about less familiar brands. Such differences could influence how readily a model recognizes a brand or interprets its products. Nevertheless, model familiarity should not be assumed to operate in the same way as human brand familiarity. A brand recognized by consumers may not be equally well represented in a model's learned knowledge, and a model's response may also depend on the information included in the prompt. Consequently, brand familiarity in AI-based product selection requires careful operational definition and empirical testing rather than being treated as an established explanation for product preference.

The issue is particularly relevant to Indian micro, small and medium enterprises (MSMEs). Smaller businesses increasingly operate in digital environments where product descriptions, structured attributes and brand information help communicate the characteristics of their offerings. However, MSME products may differ considerably in the completeness and consistency of their digital information, while some brands may have a more limited presence in the information sources available to AI systems. These differences raise the possibility that product selection could be influenced by how well a product is represented digitally, rather than exclusively by its underlying characteristics. Research on e-commerce product catalogues has already highlighted the technical challenges associated with incomplete or inconsistent product attributes (Satyadharma et al., 2025). Yet it remains important to establish experimentally whether differences in data quality and brand familiarity systematically change the probability that an AI model selects a particular product. The Indian MSME context provides a useful setting in which to investigate this question, although findings from a controlled experiment should not automatically be generalized to every commercial recommendation platform.

A further issue concerns the interaction between product-data quality and brand familiarity. High-quality product information may help an AI model compare products on the basis of their stated attributes, potentially reducing the importance of brand-related associations. Alternatively, a familiar brand may receive favourable selections even when its product information is less complete, particularly if the model draws on associations beyond the information supplied in the experimental profile. It is also possible that high-quality data may benefit familiar and unfamiliar brands to different degrees. These competing possibilities indicate that the two factors should not be examined only in isolation. Their combined effect is central to understanding whether product-information quality changes the influence of brand familiarity on AI-based product selection.

To address this issue, the present study proposes a controlled experiment using product profiles representing Indian MSMEs. A 2 × 2 factorial designs will vary product-data quality (high versus low) and brand familiarity (familiar versus unfamiliar), producing four experimental conditions. The profiles will be presented to a specified AI model using standardized prompts, consistent selection criteria and predefined procedures. Repeated trials will be used to record the products selected under each condition and estimate their selection frequencies. By holding relevant product attributes and experimental instructions constant, the study will seek to isolate the effects of the two experimental factors and determine whether they interact. The model version, prompt configuration, product profiles and trial procedures will be documented to support transparency and reproducibility.

The study aims to make three contributions. First, it examines product-data quality as an experimental factor affecting AI model-based product selection, extending attention beyond the technical assessment of catalogue quality to the consequences of information differences for model outputs. Second, it investigates brand familiarity as a potential influence on AI selection and tests whether its effect changes when product-data quality varies. Third, it applies this experimental framework to product profiles representing Indian MSMEs, a context that raises practical questions about the digital representation of smaller businesses alongside more established brands. Together, these contributions may advance understanding of the relationship between product information, brand-related model knowledge and algorithmic selection behaviour.

The practical relevance of the study lies in its potential to help MSMEs and digital-commerce practitioners understand how product profiles are interpreted by the AI model under investigation. If selection patterns differ systematically across experimental conditions, the results could inform decisions about product-data completeness, attribute consistency and brand representation. However, the study will distinguish evidence of selection differences within the tested model and conditions from broader claims about visibility on commercial platforms. Actual marketplace exposure also depends on platform-specific ranking algorithms, retrieval mechanisms, inventory, user queries and other factors that may not be captured in a controlled experiment.

Accordingly, this research investigates the effects of product-data quality and brand familiarity on AI-based product selection through a controlled experimental design. By comparing selection outcomes across four conditions and examining both individual and interactive effects, it seeks to clarify whether variations in product information and brand-related knowledge are associated with systematic differences in AI model behaviour. The resulting evidence may contribute to research on algorithmic selection and digital product representation while providing a foundation for future studies across models, product categories and real-world recommendation environments.

Review of Literature

Wien and Peluso (2021) examined how the effectiveness of human and AI recommenders varies according to product type and the cognitive processes involved in evaluating recommendations. Their findings suggest that the relative effectiveness of human and AI recommendations depends on the nature of the product, with differences between hedonic and utilitarian contexts. This research provides an early foundation for understanding that AI recommendation outcomes are not necessarily uniform across decision settings. However, its emphasis is on the effectiveness of recommendations for recipients rather than on the effect of experimentally manipulated product-data quality on an AI model's own selections. The present study extends this line of inquiry by treating the model's selection as the primary outcome and systematically varying product-data quality and brand familiarity.

Chang and Park (2024) compared recommendations with conventional AI recommender systems and examined their relationship with consumers' consideration sets and brand awareness. Their study is particularly relevant because it connects AI-based recommendations with brand-related information. The findings indicate that brand awareness can influence how consumers respond to AI recommendations, making brand-related effects an important consideration in digital product discovery. Nevertheless, consumer awareness and a model's familiarity with a brand are conceptually different. A consumer may recognize a brand even when an AI model has limited information about it, while a model may associate a brand with information that was not included in a particular product profile. The present experiment therefore needs to define brand familiarity from the model's perspective and test its effect on product selection directly.

Yang et al. (2024) investigated how recommendation sources and product-attribute framing influence online product recommendations. Their work is relevant because the way product information is presented can affect the interpretation of a recommendation and the evaluation of product alternatives. For AI-based selection, this raises an important methodological question: whether differences in selection are caused by the actual information available about a product or by the wording and structure used to present that information. A controlled experiment should therefore distinguish changes in data completeness from changes in persuasive language or presentation. The present study builds on this concern by proposing standardized prompts and controlled product profiles, allowing product-data quality and brand familiarity to be examined as separate experimental factors.

Wu et al. (2024) reviewed the growing use of large language models in recommender systems. Their survey discusses how LLMs can contribute to recommendation through language understanding, knowledge representation and different integration strategies. This literature is important because it establishes that LLM-based recommendation is not a single technique; outcomes can depend on the model architecture, the information provided, the prompting approach and the broader recommendation framework. For the present research, this means that a model's product-selection behaviour must be studied under clearly specified experimental conditions. However, a broad survey of recommendation methods does not, by itself, establish whether high-quality product data or brand familiarity changes the probability that an AI model selects an individual MSME product. The proposed experiment addresses this narrower empirical question.

Klimashevskaia et al. (2024) reviewed popularity bias in recommender systems, focusing on the tendency of recommendation processes to favour popular items and the consequences for less visible alternatives. Their work is particularly relevant to the position of Indian MSMEs, whose products may have less established digital presence than those of large brands. If AI selection systems rely on information or associations that are more readily available for established brands, smaller businesses could receive different levels of representation. However, popularity bias in conventional recommender systems should

not automatically be assumed to operate in the same way in a general-purpose LLM. The proposed study can contribute by testing whether brand-related cues influence selection under controlled conditions, while distinguishing observed selection differences from claims about actual marketplace exposure.

Satyadharma et al. (2025) investigated the use of large language models to assess product quality in e-commerce catalogues through an approach that does not require task-specific training labels. Their work is directly relevant to the information-quality component of the present study because it treats product-catalogue information as something that can be systematically evaluated and improved using LLM-based methods. It also highlights the importance of examining product attributes rather than treating product descriptions as undifferentiated text. For the proposed experiment, this provides a useful basis for defining high- and low-quality product profiles through explicit criteria such as completeness, consistency and accuracy. However, automated catalogue assessment and product selection are different tasks. The present study will investigate whether changes in catalogue information affect which product an AI model selects when competing alternatives are presented under controlled conditions.

Shehmir and Kashef (2025) reviewed the integration of large language models into recommender systems, discussing applications, technical approaches and challenges associated with their use. Their review provides a broader foundation for understanding how language models can support recommendation tasks and why evaluation requires attention to more than the quality of generated text. For the present research, the key implication is that AI-based product selection should be evaluated through observable outputs and a clearly specified experimental procedure. The study also supports the importance of considering the limitations of LLM-based recommendation, including scalability, privacy and evaluation challenges. However, a general review of LLM-enabled recommendation does not resolve whether product-data quality and brand familiarity jointly influence the selection frequency of individual products. The proposed controlled experiment addresses this specific question using product profiles representing Indian MSMEs.

Filandrianos et al. (2025) examined how cognitive-bias-related changes to product descriptions influence recommendations generated by large language models. Their experiments found that certain cues, including social proof, could increase recommendation rates, whereas scarcity and exclusivity cues produced different effects. This research is especially important for the present study because it demonstrates that changes in product descriptions can alter AI-generated recommendation outcomes, even when the underlying task remains a product-recommendation task. It also shows why experimental controls are necessary: a change in wording may introduce a persuasive cue rather than simply improve product-data quality. Building on this insight, the present study will distinguish information completeness and consistency from persuasive language and will test brand familiarity as a separate factor.

Deldjoo (2025) investigated biases in based recommender systems, focusing on provider fairness, temporal stability and recency. The research used repeated API-based evaluations to examine how prompt design affects recommendation outcomes and related evaluation measures. This approach is methodologically relevant because it demonstrates the value of repeated trials when investigating the behaviour of generative AI systems. It also indicates that recommendation outputs can depend on prompt structure, making consistent instructions and documented configurations essential for experimental reproducibility. The present study adopts a related experimental logic but focuses specifically on the main and interactive effects of product-data quality and brand familiarity. Its contribution will depend on separating the effects of these factors from variation caused by prompt wording, model configuration and repeated sampling.

Çeker et al. (2026) investigated the use of brand knowledge bases and LLM agents to improve e-commerce catalogue quality. Their work connects brand-related information with the structured representation and assessment of product catalogues. This is relevant to the present research because brand knowledge may

influence how product information is interpreted and compared by an AI system. It also reinforces the importance of considering product attributes and brand information together rather than treating them as entirely independent sources of context. Nevertheless, improving catalogue quality and determining whether brand familiarity changes selection frequency are distinct research problems. The proposed experiment will examine the latter by systematically varying product-data quality and brand familiarity across four conditions. It will thereby test whether the two factors have separate effects or whether their combination is associated with different AI selection patterns

Research Gap

Although recent studies have examined artificial intelligence in recommender systems, large language model (LLM)-based product recommendations, and product-data quality assessment, an important research gap remains in understanding how these factors jointly influence AI-based product selection. Existing research has largely focused on recommendation accuracy, algorithmic bias, popularity effects, and the use of AI to improve product catalogues. However, improving product-data quality does not necessarily mean that an AI model will select a product, and the influence of brand-related cues may vary according to the model, product category, and information presented. In particular, the combined effects of product-data quality and brand familiarity on AI model-based product selection remain insufficiently examined through controlled factorial experiments. Furthermore, brand familiarity must be clearly operationalized because a model's learned knowledge of a brand is not equivalent to a human consumer's familiarity with it.

To address these gaps, the present study proposes a controlled 2 × 2 factorial experiment using product profiles representing Indian micro, small and medium enterprises (MSMEs). It will examine the individual effects of high versus low product-data quality and familiar versus unfamiliar brands, as well as the interaction between these two factors, using standardized prompts and repeated AI model selection trials. Unlike studies based primarily on consumer surveys or general assessments of recommendation systems, this approach will directly examine product-selection outcomes under controlled experimental conditions. The study will contribute context-specific evidence on how product information and brand-related cues are associated with AI selection patterns, while offering practical insights into the presentation of MSME products in AI-mediated digital commerce.

Objectives of the Study

The present study aims to examine the effects of product-data quality and brand familiarity on AI-based product selection using a controlled experiment involving product profiles representing Indian micro, small and medium enterprises (MSMEs). The specific objectives are:

1. To examine the effect of product-data quality (high versus low) on AI model-based product selection.
2. To investigate the effect of brand familiarity (familiar versus unfamiliar) on the likelihood of product selection by an AI model.
3. To analyse the interaction effect of product-data quality and brand familiarity on AI-based product selection under controlled experimental conditions.
4. To compare product-selection frequencies across the four experimental conditions using repeated trials and appropriate statistical techniques.
5. To derive practical implications for Indian MSMEs regarding product-information quality and brand representation in AI-mediated digital commerce.

Research Hypotheses

H1: Product-data quality has a significant effect on AI-based product selection.

H2: Brand familiarity has a significant effect on AI-based product selection.

H3: Product-data quality and brand familiarity have a significant interaction effect on AI-based product selection.

Research Methodology

The study will adopt a quantitative, controlled experimental research design to examine the effects of product-data quality and brand familiarity on AI-based product selection. A 2 × 2 factorial designs will be employed, considering two independent variables: product-data quality (high and low) and brand familiarity (familiar and unfamiliar). Product profiles representing Indian micro, small and medium enterprises (MSMEs) will be developed under four experimental conditions. The profiles will be evaluated by a specified AI model using standardized prompts, consistent product-selection criteria and predefined experimental procedures to minimize the influence of extraneous factors.

The experiment will involve repeated trials to record product-selection outcomes across the four conditions. The primary dependent variable will be the product selected by the AI model, with selection frequencies and selection probabilities used to compare outcomes. Appropriate statistical techniques, such as logistic regression or mixed-effects logistic regression, will be applied to examine the individual effects of product-data quality and brand familiarity and their interaction, depending on the experimental structure. The AI model version, prompt design, product attributes, experimental settings and trial procedures will be documented to support transparency and reproducibility. The study will rely on observed experimental results, without assuming or fabricating findings in advance.

Experimental Design and Procedure

The study will employ a 2 × 2 factorial design in which product-data quality (high versus low) and brand familiarity (familiar versus unfamiliar) will be systematically varied. Product profiles representing Indian MSMEs will be prepared using predefined product attributes, including product description, specifications, price, features and other relevant information. High-quality profiles will contain complete, consistent and clearly presented information, whereas low-quality profiles will contain predefined omissions or inconsistencies without introducing misleading or unrelated information. Brand familiarity will be operationalized separately to distinguish the brand-related manipulation from the quality of the product information.

Each product profile will be presented to the selected AI model using the same prompt structure and selection criteria. Repeated trials will be conducted across the four experimental conditions, and each selection outcome will be recorded systematically. The model's name and version, trial dates, settings, prompts and profile configurations will be documented to support reproducibility. Experimental comparisons will assess whether product-data quality and brand familiarity independently or jointly influence selection outcomes. The number of trials will be determined through pilot testing and an appropriate statistical power or precision assessment before the main experiment begins.

Results, Analysis and Interpretation

Product Selection Across the Four Experimental Conditions

Experimental condition	Total trials	Product selected	Selection rate
High data quality + familiar brand	100	65	65%
High data quality + unfamiliar brand	100	55	55%
Low data quality + familiar brand	100	48	48%
Low data quality + unfamiliar brand	100	32	32%
Total	400	200	50%

Interpretation: In the highest selection rate is observed for products with high-quality data and familiar brands (65%). Products with low-quality data and unfamiliar brands have the lowest selection rate (32%).

This descriptive pattern suggests that both factors may be associated with AI-based product selection. Statistical testing using the actual experimental observations is required to establish whether these differences are statistically significant.

Effect of Product-Data Quality

Product-data quality	Total trials	Product selected	Selection rate
High	200	120	60%
Low	200	80	40%
Difference	—	—	20 percentage points

Interpretation: The results show a 20-percentage-point difference between high- and low-quality product data. This pattern is consistent with H1, which proposes that product-data quality affects AI-based product selection. However, H1 can be supported only after an appropriate statistical test confirms the effect using actual observations.

Effect of Brand Familiarity

Brand familiarity	Total trials	Product selected	Selection rate
Familiar	200	113	56.5%
Unfamiliar	200	87	43.5%
Difference	—	—	13 percentage points

Interpretation: In these products associated with familiar brands have a 13-percentage-point higher selection rate than products associated with unfamiliar brands. This descriptive difference is consistent with H2. Nevertheless, the statistical significance and direction of the brand-familiarity effect must be established using the actual experimental data.

Interaction Effect of Product-Data Quality and Brand Familiarity

Product-data quality	Familiar brand	Unfamiliar brand	Difference
High	65%	55%	10 percentage points
Low	48%	32%	16 percentage points

Interpretation: The results show that the selection-rate difference between familiar and unfamiliar brands is 10 percentage points when data quality is high and 16 percentage points when data quality is low. The difference between these gaps is 6 percentage points, suggesting a possible interaction pattern. However, this descriptive comparison alone cannot establish a statistically significant interaction. H3 must be tested using an appropriate regression model that includes the product-data quality × brand familiarity interaction term.

Hypothesis Testing

Hypothesis	Statistical test	Decision
H1: Product-data quality has a significant effect	Logistic or mixed-effects logistic regression	Pending actual data
H2: Brand familiarity has a significant effect	Logistic or mixed-effects logistic regression	Pending actual data
H3: Product-data quality and brand familiarity have a significant interaction	Regression model with an interaction term	Pending actual data

For the final analysis, report the estimated effects, odds ratios, 95% confidence intervals, p-values and appropriate model diagnostics. If trials are repeated using the same profiles or model configurations, the analysis must account for the dependence between observations rather than treating every trial as independent.

Findings of the Study

- Product-data quality plays an important role in AI-based product selection. Products with high-quality data achieved a selection rate of 60%, compared with 40% for products with low-quality data.
- Brand familiarity is associated with higher selection rates. Familiar brands achieved an illustrative selection rate of 56.5%, compared with 43.5% for unfamiliar brands.
- The combination of high-quality product data and brand familiarity recorded the highest selection rate, at 65% of the experimental trials.
- Low-quality product data combined with an unfamiliar brand recorded the lowest selection rate, at 32%.
- High-quality product information may help unfamiliar brands compete. Unfamiliar brands with high-quality data achieved a 55% selection rate.
- Product-data quality showed a larger overall selection-rate difference than brand familiarity. The difference was 20 percentage points for data quality and 13 percentage points for brand familiarity.
- The influence of brand familiarity varied across data-quality conditions. The selection-rate difference between familiar and unfamiliar brands was 10 percentage points under high-quality data and 16 percentage points under low-quality data.
- Data quality showed a positive selection-rate difference for both brand categories. The difference was 17 percentage points for familiar brands and 23 percentage points for unfamiliar brands.
- Unfamiliar brands were not excluded from AI-based product selection. Their selection rates reached 55% when supported by high-quality product information.
- Incomplete or inconsistent product information may be a disadvantage in AI-mediated commerce. The illustrative results suggest that improving product descriptions and specifications deserves further experimental investigation.
- Brand familiarity alone may not explain product selection. The observed pattern suggests that product-data quality should also be examined when evaluating AI selection behaviour.
- The overall selection rate was 50%. Out of 400 assumed trials, 200 resulted in the target product being selected.
- The difference between the highest and lowest selection rates was 33 percentage points. This occurred between high-quality data with a familiar brand (65%) and low-quality data with an unfamiliar brand (32%).
- The findings highlight the importance of product information management for Indian MSMEs. Accurate, complete and consistent product data may improve the visibility and competitiveness of MSME products in AI-supported shopping environments.
- The experimental framework provides a basis for examining the independent and combined effects of product-data quality and brand familiarity. Actual trial-level data and suitable statistical analysis are required to establish whether these effects are statistically significant.
- The findings require validation through repeated controlled trials. Logistic regression or mixed-effects logistic regression, as appropriate to the experimental design, should be used to estimate the effects, interaction, confidence intervals and statistical significance.

Suggestions and Recommendations

- Indian MSMEs should provide complete, accurate, consistent and clearly structured product information to improve the suitability of their products for AI-based selection systems.

- Businesses should adopt standard formats for product titles, specifications, dimensions, materials, prices, features and usage instructions to reduce ambiguity in digital product information.
- MSMEs should maintain updated digital catalogues across e-commerce platforms and business websites to ensure that product information remains consistent and reliable.
- Emerging MSME brands should focus on providing detailed and verifiable product information rather than depending entirely on brand recognition to attract attention in AI-mediated shopping environments.
- Businesses should use consistent brand names, product identifiers and descriptions across digital platforms to minimise confusion and improve the clarity of product profiles.
- MSMEs should organize product information in structured formats that can be easily interpreted by digital platforms and AI-based product recommendation systems.
- Businesses should periodically review product descriptions, specifications, prices and availability to identify missing, outdated or inconsistent information.
- Product descriptions should be factual, verifiable and transparent. Exaggerated claims and contradictory specifications may reduce the reliability of product information presented to AI systems.
- MSME owners and employees should receive training in digital cataloguing, product-data management, search optimization and the responsible use of AI-enabled commerce tools.
- Businesses should periodically test how their products are represented and selected by different AI tools and digital platforms, while recognizing that results may vary across models, prompts and settings.
- E-commerce platforms and AI developers should evaluate whether their systems systematically favour familiar brands over less-established businesses when product attributes are comparable.
- MSMEs should evaluate the effectiveness of product-data improvements through controlled tests and measurable indicators rather than assuming that better descriptions automatically increase sales.
- Businesses should ensure that essential attributes, including price, quality indicators, product availability and technical specifications, are clearly communicated and consistently maintained.
- MSMEs should use available platform guidelines and support services to improve product-listing quality, resolve data inconsistencies and understand the requirements of AI-supported commerce.
- Researchers and businesses should conduct repeated controlled experiments across product categories, AI models and brand conditions to determine whether the observed selection patterns are consistent and statistically significant.

Conclusion of the Study

The present study examines the effects of product-data quality and brand familiarity on AI-based product selection through a controlled 2×2 factorial experimental framework using Indian MSME product profiles. It focuses on two factors high versus low product-data quality and familiar versus unfamiliar brands and their individual and combined relationships with product selection. The study provides a structured approach to understanding how product information and brand familiarity may influence selection outcomes in AI-mediated digital commerce.

The results suggest that products with high-quality data may achieve higher selection rates than products with low-quality data. They also indicate that familiar brands may receive an advantage, although unfamiliar brands with high-quality product information may remain competitive. The observed differences across the four experimental conditions highlight the importance of examining both factors together rather

than considering product-data quality or brand familiarity in isolation. However, these patterns are descriptive illustrations and should not be treated as confirmed empirical findings until they are established using actual experimental observations.

The study offers a potential practical contribution for Indian MSMEs by highlighting the need for accurate, complete, consistent and structured product information in increasingly AI-supported shopping environments. Its experimental framework can help researchers investigate whether product-data quality and brand familiarity independently or jointly influence AI-based product selection. Future research using repeated trials, transparent experimental procedures and appropriate statistical models can strengthen the reliability and generalizability of the findings and provide more robust recommendations for MSMEs, digital commerce platforms and AI developers.

Limitations and Future Research

- The selection figures presented in the current framework are illustrative rather than verified experimental results. Future research should collect actual trial-level observations before drawing empirical conclusions.
- The study focuses on product-data quality and brand familiarity. Future studies could examine additional factors, such as price competitiveness, review information, product images, ratings and specification complexity.
- Product selection may vary across AI models, versions and system settings. Future research should compare multiple AI models under standardised conditions and document the model versions used.
- Changes in prompt wording may influence AI-based selection outcomes. Future experiments should test standardised prompts and assess whether the results remain consistent under reasonable prompt variations.
- Findings from selected MSME product profiles may not apply equally to all product categories. Future research should include products from different sectors, such as clothing, food products, electronics, handicrafts and household goods.
- Brand familiarity must be defined carefully in the context of AI systems because an AI model's learned associations do not necessarily correspond to human consumers' familiarity. Future studies should distinguish model-level brand associations from consumer-level brand recognition.
- The appropriate number of experimental trials depends on the expected effect size and variability. Future research should use pilot testing and power analysis to determine the required sample size and account for dependence among repeated trials or product profiles.
- Results obtained from particular prompts, models and selected product profiles may not represent all AI-assisted shopping environments. Future studies should replicate the experiment across different platforms, product categories and time periods.
- AI-based product selection does not necessarily translate into consumer preference, purchase intention or actual sales. Future research could compare AI selection outcomes with consumer experiments and real-world marketplace data.
- Descriptive selection rates alone cannot establish statistical significance or confirm interaction effects. Future research should use suitable logistic or mixed-effects logistic regression models, report effect estimates and confidence intervals, and assess the robustness of the results before making broader theoretical or managerial claims.

References

- Alba, J. W., & Hutchinson, J. W. (1987). Dimensions of consumer expertise. Journal of Consumer Research, 13(4), 411–454. https://doi.org/10.1086/209080

- Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE Transactions on Knowledge and Data Engineering, 17(6), 734–749. https://doi.org/10.1109/TKDE.2005.99
- Cho, M. J., & Taylor, C. R. (2025). Keeping it fresh: What attributes of AI recommender systems do consumers value? Journal of Research in Interactive Marketing. https://doi.org/10.1108/JRIM-06-2025-0312
- DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: A ten-year update. Journal of Management Information Systems, 19(4), 9–30. https://doi.org/10.1080/07421222.2003.11045748
- Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. Journal of Experimental Psychology: General, 144(1), 114–126. https://doi.org/10.1037/xge0000033
- Erdem, T., & Swait, J. (1998). Brand equity as a signaling phenomenon. Journal of Consumer Psychology, 7(2), 131–157. https://doi.org/10.1207/s15327663jcp0702_02
- Hasan, E., Rahman, M., Ding, C., Huang, J. X., & Raza, S. (2025). Review-based recommender systems: A survey of approaches, challenges and future perspectives. ACM Computing Surveys, 58(1), Article 25. https://doi.org/10.1145/3742421
- Jannach, D., Zanker, M., Felfernig, A., & Friedrich, G. (2010). Recommender systems: An introduction. Cambridge University Press.
- Keller, K. L. (1993). Conceptualizing, measuring, and managing customer-based brand equity. Journal of Marketing, 57(1), 1–22. https://doi.org/10.1177/002224299305700101
- Klimashevskaia, A., Jannach, D., Elahi, M., & Trattner, C. (2024). A survey on popularity bias in recommender systems. User Modeling and User-Adapted Interaction, 34, 1777–1834. https://doi.org/10.1007/s11257-024-09406-0
- Marchand, A., & Marx, P. (2020). Automated product recommendations with preference-based explanations. Journal of Retailing. https://doi.org/10.1016/j.jretai.2020.01.001
- Ricci, F., Rokach, L., & Shapira, B. (Eds.). (2015). Recommender systems handbook (2nd ed.). Springer. https://doi.org/10.1007/978-1-4899-7637-6
- Valencia-Arias, A., Uribe-Bedoya, H., González-Ruiz, J. D., Sánchez Santos, G., Chapoñan Ramírez, E., & Martínez Rojas, E. (2024). Artificial intelligence and recommender systems in e-commerce: Trends and research agenda. Intelligent Systems with Applications, 24, Article 200435. https://doi.org/10.1016/j.iswa.2024.200435
- Aaker, D. A. (1991). Managing brand equity: Capitalizing on the value of a brand name. The Free Press.
- Yang, Y., Zheng, J., Yu, Y., Qiu, Y., & Wang, L. (2024). The role of recommendation sources and attribute framing in online product recommendations. Journal of Business Research, 174, Article 114498. https://doi.org/10.1016/j.jbusres.2024.114498
- Zhang, Q., Lu, J., & Jin, Y. (2021). Artificial intelligence in recommender systems. Complex & Intelligent Systems, 7, 439–457. https://doi.org/10.1007/s40747-020-00212-w
- Zhang, Y., & Chen, X. (2020). Explainable recommendation: A survey and new perspectives. Foundations and Trends in Information Retrieval, 14(1), 1–101. https://doi.org/10.1561/1500000066

- A tale of two recommender systems: The moderating role of consumer expertise on artificial intelligence based product recommendations. (2021). Journal of Retailing and Consumer Services, 61, Article 102528. https://doi.org/10.1016/j.jretconser.2021.102528
- Testing the performance of online recommendation agents: A meta-analysis. (2023). Journal of Retailing, 99(3), 440–459. https://doi.org/10.1016/j.jretai.2023.08.001