Analysis of Demographic and Consumer Behavior Factors on Satisfaction with AI Technology Usage in Digital Retail Using the Random Forest Algorithm

Eko Priyanto[1,*], Ahmad Saekhu[2], Priyo Agung Prasetyo[3,]

[1,2,3]*Magister of Computer Science, Universitas Amikom Purwokerto, Jawa Tengah, Indonesia*

(Received: May 15, 2024; Revised: July 1, 2024; Accepted: September 15, 2024; Available online: December 1, 2024)

Abstract

The rapid integration of artificial intelligence (AI) into digital retail has reshaped consumer interactions, enabling personalized services and operational enhancements. This study investigates the demographic and behavioral factors influencing consumer satisfaction with AI technologies in digital retail, using the Random Forest classification algorithm for predictive modeling. After comprehensive preprocessing and hyperparameter tuning through grid search cross-validation, the Random Forest model achieved an overall accuracy of 83%. While the model showed strong performance for predicting satisfied consumers yielding a precision of 0.84, recall of 0.97, and F1-score of 0.90, it performed poorly in identifying dissatisfied users, with a recall of only 0.27 and F1-score of 0.39, highlighting a class imbalance issue. Feature importance analysis revealed that experiential factors, particularly enhanced AI experience and preference for online services, significantly influenced satisfaction levels, whereas demographic variables such as age and gender had limited predictive value. These findings emphasize the need for digital retailers to focus on user-centric design and service personalization, rather than demographic segmentation alone, to enhance customer satisfaction and loyalty. Furthermore, the study contributes methodologically by demonstrating the effectiveness of Random Forest in handling complex consumer datasets and theoretically by validating TAM and Customer Satisfaction Theory in the context of AI adoption. Despite limitations related to class imbalance and sector-specific data, this research offers actionable insights for retailers, marketers, and system developers aiming to improve AI-driven service quality and consumer engagement. Future studies are encouraged to address these limitations through the inclusion of emotional and contextual variables and by expanding the analysis to other industries for broader applicability.

Keywords: Artificial Intelligence, Consumer Satisfaction, Digital Retail, Random Forest, Technology Adoption

1. Introduction

The advancement of artificial intelligence (AI) technology has profoundly transformed the digital retail sector, generating new efficiencies and elevating customer experiences. AI enables businesses to analyze vast datasets of consumer behavior, thereby facilitating personalized shopping journeys and optimizing critical operational processes such as supply chain logistics and inventory management. These technological enhancements contribute not only to increased sales and cost reductions but also to the emergence of a new retail paradigm commonly referred to as “new retail.” This model integrates online and offline services using technologies such as AI and cloud computing to create seamless and consistent customer experiences [1].

Within the context of new retail, AI plays a central role in driving customer satisfaction and loyalty. By enabling intelligent service interactions and delivering highly personalized product recommendations, AI directly influences consumer purchase intentions [2]. However, the adoption of AI also presents challenges for retailers, such as navigating data privacy regulations and overcoming barriers to technological integration. These issues necessitate careful strategy and infrastructure planning to fully leverage AI's potential while maintaining consumer trust [3].

Understanding the key factors that influence user satisfaction with AI applications is vital for improving user engagement and encouraging sustained adoption. Research highlights that satisfaction is significantly shaped by communication clarity and the overall quality of interactions with AI platforms, such as chatbots and virtual assistants

***Corresponding author: Eko Priyanto (eko.p@umnu.ac.id)**
DOI: https://doi.org/10.47738/ijaim.v4i4.91

[4]. Moreover, satisfaction is closely linked to personalization, trust in technology, and perceived ease of use dimensions that are central to the user’s willingness to continue engaging with AI [5]. Users tend to express higher satisfaction when they feel in control of their personal data and are provided transparency about algorithmic decision-making [6].

Other behavioral aspects, such as cognitive engagement and the holistic user experience, also play significant roles in shaping consumer perceptions of AI. Studies have shown that AI systems delivering service quality comparable to human interactions can significantly enhance satisfaction and foster loyalty, particularly in customer-centric sectors like retail and banking [7]. These insights underscore the importance of developing AI applications that are not only technically proficient but also user-centric in their design and execution.

In examining satisfaction with AI in digital retail, demographic factors must also be considered, as they greatly influence consumer behavior and technology usage. Generation Z, often described as digital natives, exhibit a high level of digital literacy and fluency with emerging technologies. They frequently rely on digital tools for education, shopping, and social interaction, positioning them as a key target demographic for businesses operating in digital spaces [8]. Their familiarity and comfort with technology make them more receptive to AI-powered platforms.

In contrast, older populations frequently encounter challenges with digital technologies due to lower familiarity, reduced digital literacy, and perceived usability barriers. This digital divide is further widened by socio-economic factors such as education, income, and geographic location, resulting in unequal access to and usage of AI technologies across different segments of society [9]. Emotional engagement also plays a role users are more inclined to adopt technologies that evoke enjoyment or enhance their overall experience [10].

Demographic attributes such as age and digital proficiency are particularly influential in shaping user satisfaction with AI in digital retail environments. Studies have demonstrated that younger consumers, especially Millennials and Generation Z, are generally more satisfied with AI-driven services due to their higher comfort levels and openness to technological innovation [11]. On the other hand, older individuals may exhibit skepticism or discomfort, which can result in lower satisfaction rates.

In addition to demographics, trust in AI systems, perceived service quality, and overall system usability are critical determinants of user satisfaction. High-quality service interactions, user-centered design, and efficient performance are perceived as valuable by consumers, leading to greater satisfaction and increased loyalty [12]. AI’s ability to deliver tailored, preference-based experiences further elevates satisfaction, as modern consumers increasingly seek personalization in their digital interactions [13].

To analyze and predict the factors that affect consumer satisfaction with AI, the Random Forest algorithm has emerged as an effective analytical tool. As an ensemble learning method, Random Forest enhances predictive accuracy by combining the outputs of multiple decision trees, thereby reducing variance and minimizing overfitting [14]. Its ability to manage high-dimensional and complex datasets makes it particularly well-suited for identifying key variables that influence user satisfaction [15].

Furthermore, Random Forest is especially capable of handling imbalanced datasets, which are common in consumer research where satisfied and dissatisfied users may not be equally represented [16]. Its robustness allows it to reveal complex, non-linear relationships between variables such as user demographics, interaction experiences, and satisfaction levels insights that may be overlooked by simpler models.

The algorithm’s effectiveness has been demonstrated across domains such as healthcare and job satisfaction, where it has successfully identified nuanced relationships between user characteristics and perceived service quality [15]. Additionally, Random Forest models can integrate both qualitative feedback and quantitative metrics, offering a comprehensive approach to understanding satisfaction dynamics. Techniques such as grid search for parameter optimization further enhance model accuracy and reliability.

This study is significant in its potential to inform strategic decision-making for digital retail businesses seeking to enhance their customer engagement and technological adoption strategies. By identifying the demographic and behavioral factors influencing AI satisfaction, businesses can tailor their marketing and service designs more precisely to meet the expectations of different consumer segments. For instance, personalized communication and improved interaction quality with AI platforms have been shown to directly enhance customer retention and loyalty [4].

AI-powered tools can also support smarter marketing efforts by leveraging customer data analytics to build customized experiences. This capability is particularly important in the post-pandemic retail landscape, where direct-to-consumer models are gaining prominence [17]. Additionally, AI facilitates more efficient inventory and demand management, improving service responsiveness and contributing to elevated customer satisfaction [16].

By applying robust analytical models like Random Forest, retailers can generate accurate predictions about customer satisfaction and make informed, data-driven adjustments to their strategies. These insights not only enable better targeting and personalization but also drive continuous development of AI tools in alignment with user expectations [18]. Such capabilities offer a competitive edge in an increasingly digital and customer-centric marketplace.

In summary, this study contributes to both theoretical and practical understandings of AI adoption and digital consumer behavior. It integrates insights from prior literature on AI's operational, experiential, and demographic influences while offering actionable frameworks for businesses aiming to improve satisfaction through AI. Furthermore, it underscores the value of machine learning techniques particularly Random Forest in capturing the complexity of consumer experiences and guiding AI system optimization in digital retail environments.

2. Literature Review

2.1. Previous studies on AI user satisfaction

User satisfaction with AI technologies in digital retail has become a focal point of contemporary research, with scholars emphasizing the psychological and perceptual factors that shape consumer experiences. Lopes et al. [19] highlight the role of psychosocial influences, such as perceived behavioral control and social norms, in driving consumer acceptance of AI-enabled services. Their study shows that when users perceive AI as useful and observe its acceptance in their social circles, engagement and satisfaction increase. This is supported by Ciuchita et al. [20], who found that favorable perceptions of AI in advertising enhance consumers' overall attitudes toward AI in retail, establishing a strong link between consumer perception and satisfaction.

Another dominant theme in the literature is the significance of personalization through AI integration. Srivastava and Pal [21] emphasize how in-store AI features like chatbots and facial recognition enhance shopping experiences by offering real-time, tailored interactions. Likewise, Raji et al. [13] reinforce that AI-driven personalization increases customer satisfaction by aligning product recommendations and services with individual preferences. These findings point to personalization as a critical factor in effective AI adoption strategies. Generational differences also emerge as influential variables, Guerra-Tamez et al. [22] report that Generation Z due to their frequent exposure to digital technologies—tend to respond more positively to AI interactions, which in turn influences their purchasing behavior and trust in brands.

Beyond generational trends, psychological and behavioral factors play a crucial role in AI acceptance and satisfaction. Arachchi and Samarasinghe [23] identify perceived ease of use, enjoyment, and usefulness as key drivers of satisfaction among various age groups. Their research aligns with Dutta [24] findings, which advocate for the strategic implementation of generative AI to personalize services, streamline operations, and enhance customer engagement. Collectively, the literature suggests that for retailers to successfully harness AI technologies, they must address both the technical and experiential dimensions of user interaction recognizing that satisfaction is shaped not only by functionality but also by how customers perceive, engage with, and emotionally respond to AI systems.

2.2. Influence of Demographic Factors on Technology Adoption

Demographic factors such as age, education, income, and gender play a significant role in shaping individuals' willingness and ability to adopt new technologies. Age and education, in particular, have been shown to influence adoption behaviors across various sectors. For instance, Prakash et al. [25] observed that younger farmers with higher education levels are more likely to embrace modern agricultural technologies, underscoring the role of education as a key enabler of technology acceptance. Bashir et al. [26] similarly found that individuals aged 25 to 35 were more inclined to adopt vermicomposting techniques, driven by economic considerations and openness to innovation. Income also emerges as a critical determinant, with Pandita et al. [27] demonstrating that higher-income individuals are more likely to adopt electric vehicles, while Emon [28] found a similar pattern in the adoption of solar technology. These findings suggest that financial capability often determines access to and adoption of advanced technological solutions.

Gender and cultural influences further shape adoption patterns. Wexler and Fan [29] reported that women tend to express lower comfort levels with emerging technologies such as automated vehicles, indicating a potential gender-based aversion to unfamiliar systems. Similarly, Sallam et al. [30] found that perceptions of usefulness and ease of use vary across genders, particularly in educational technology contexts. Cultural and social dimensions also influence adoption, as Ikoyo-Eweto et al. [31] emphasized the role of social structures and access to extension services in determining the speed and extent of agricultural technology uptake. These findings highlight that technology adoption is not solely a matter of access or utility, but also deeply embedded in social, cultural, and demographic contexts. Understanding these influences is crucial for designing inclusive strategies that promote broader and more equitable technology adoption.

2.3. Consumer Behavior in the Digital Retail Context

Consumer behavior in the digital retail environment is shaped by a dynamic combination of technological, social, demographic, and psychological factors. AI technologies, particularly chatbots, have been shown to influence consumer satisfaction and purchasing behavior. M. Lee and Park [4] found that effective communication through AI shopping chatbots enhances trust and satisfaction, which in turn fosters continued usage. Similarly, ElSayad and Mamdouh [32] emphasized that trust and perceived usefulness of AI platforms are key predictors of purchase intentions, particularly among younger consumers who are more receptive to digital innovations. In addition, social media has emerged as a major influence on digital shopping behavior. During the COVID-19 pandemic, Sumi and Ahmed [33] observed that consumer attitudes shifted significantly due to increased social media engagement, with peer influence and promotional content shaping purchase decisions. Murshed and Uğurlu [34] further advocate for integrating social media into marketing strategies to effectively drive online shopping engagement.

Demographic and cultural factors also play a critical role in shaping consumer behavior in digital retail. Research by Sudirjo and Tjahyadi [35] revealed that younger Indonesian consumers prioritize usability and trust in online platforms, reflecting the generational shift toward digital adoption. Rakate and Gaikwad [36] supported this by noting how different demographic groups respond uniquely to fashion and technology trends, underscoring the need for targeted marketing strategies. Moreover, psychological factors such as hedonic and impulsive buying behaviors are increasingly relevant in the digital context. Miah et al. [37] emphasized how the pandemic altered decision-making patterns and consumer habits. Together, these findings illustrate the complexity of digital consumer behavior and highlight the importance for retailers to develop adaptive, data-informed strategies that address both technological and human factors.

2.4. Use of Random Forest Algorithm in Consumer and Technology Data Analysis

The Random Forest algorithm has been widely recognized for its effectiveness in analyzing consumer behavior and technology-related data, particularly in enhancing user satisfaction and predictive decision-making. Luo [15] demonstrated its utility in forecasting customer satisfaction by employing information gain to select relevant attributes from large-scale consumer datasets. The study highlights Random Forest's ability to build reliable and accurate predictive models in complex environments with multiple variables. Supporting this, Huang [38] showed that Random Forest outperformed logistic regression models when predicting purchase behavior based on user data from the Tianchi platform, underlining its potential for improving marketing personalization and inventory optimization. Additionally, Kumar and Kumar [39] found that Random Forest consistently provided higher classification accuracy compared to traditional decision trees in human activity pattern analysis, reinforcing its value in user behavior modeling.

Another key advantage of Random Forest is its robustness in handling incomplete datasets, a common challenge in real-world consumer data analysis. Ma et al. [40] reported that the algorithm maintained strong predictive accuracy despite missing values, reducing the need for extensive data preprocessing. This makes it particularly useful for applications where data integrity cannot always be guaranteed. Beyond consumer analytics, the algorithm has also been successfully applied in financial forecasting. Liu [41] utilized Random Forest in enterprise risk management, achieving high clustering accuracy and predictive reliability, which illustrates the algorithm's versatility across domains. Overall, these studies establish Random Forest as a powerful and adaptable tool capable of uncovering nuanced patterns in consumer behavior and technological trends. Its scalability, accuracy, and ability to process large, messy datasets make it especially valuable for businesses seeking data-driven insights to enhance customer experience and strategic decision-making.

2.5. Supporting Theories: Technology Acceptance Model (TAM) and Customer Satisfaction Theory

Understanding consumer behavior in the context of AI adoption within digital retail is effectively guided by two foundational theories: the Technology Acceptance Model (TAM) and Customer Satisfaction Theory. The TAM framework asserts that perceived ease of use and perceived usefulness are the primary drivers of an individual's intention to adopt new technologies [24]. In retail environments, these constructs are especially relevant as consumers increasingly engage with AI-powered platforms, such as chatbots and generative AI tools. When these technologies are perceived as intuitive and helpful in enhancing shopping efficiency or personalization, user acceptance tends to rise. Dutta [24] supports this by illustrating how generative AI, when integrated into omnichannel strategies, satisfies evolving consumer expectations and strengthens adoption intent by delivering valuable, user-friendly experiences. This underscores the TAM principle that usability and perceived benefits are central to digital technology acceptance.
Complementing TAM, Customer Satisfaction Theory posits that satisfaction arises when the actual experience of a product or service meets or exceeds consumer expectations. In the context of digital retail, this often relates to service quality, such as the responsiveness and reliability of AI systems. Irshad et al. [42] demonstrated that improvements in service delivery and corporate social responsibility (CSR) efforts positively affect satisfaction and purchase intentions, linking emotional responses to long-term behavioral outcomes. Moreover, satisfaction is intertwined with perceived value and brand loyalty, as shown by Zhang and Ahmad [43], who identified that CSR initiatives and consumer identity significantly influence purchasing behavior. Collectively, these theories provide a comprehensive understanding of how technological features and service experiences jointly shape consumer attitudes, guiding strategic efforts to enhance satisfaction, loyalty, and sustained use of AI technologies in digital retail.

3. Methodology

To ensure the robustness and validity of the predictive model, this study follows a structured methodology consisting of several key stages, ranging from data acquisition to final model evaluation. The complete workflow is illustrated in Figure 1.

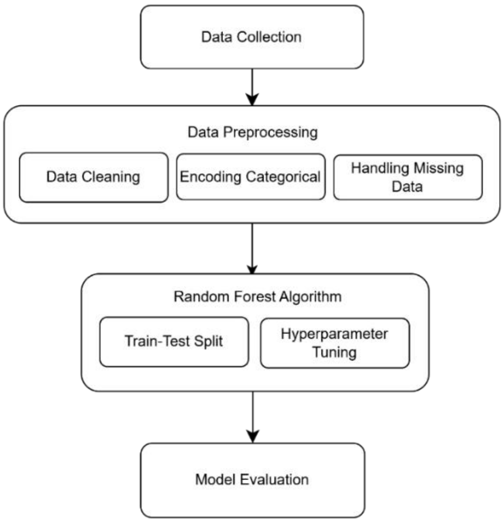

Figure 1. Research Methodology

3.1. Data Collection

The dataset used in this study was sourced from Kaggle, a well-known open data platform providing diverse datasets for research. Specifically, the "AI in Retail Dataset" contains comprehensive information about consumers' demographic characteristics and their behaviors related to the use of AI technologies within digital retail environments.

ISSN 2776-8007

Key demographic variables included in the dataset are Country, Gender, Age, Education, and Living Region, while consumer behavior variables encompass online consumer status, preferences for online services, payment methods, and usage of AI tools such as chatbots and virtual assistants. The target variable, AI Satisfaction, reflects consumers' level of satisfaction with AI technologies in retail settings. This dataset was originally compiled from survey data collected across multiple regions and is publicly accessible on Kaggle, enabling a robust analysis of factors influencing AI satisfaction. Prior to analysis, the data underwent cleaning procedures to handle missing values and categorical data encoding to prepare it for machine learning modeling, specifically using the Random Forest algorithm.

3.2.Data Preprocessing

Before conducting the analysis, the raw dataset underwent several preprocessing steps to ensure data quality and suitability for the Random Forest model. Initially, the dataset was carefully examined for inconsistencies, duplicate entries, and irrelevant columns. Records containing incomplete or invalid responses were removed to maintain data integrity, and column names were standardized by trimming any leading or trailing whitespace. Since many variables were categorical such as Gender, Education, Country, and AI Tool Usage, they were converted into numerical format using label encoding, which assigns unique integers to each category while preserving the categorical information. Missing data were addressed based on their nature and extent; records with only a few missing values were removed, while features with more significant missing data were handled using imputation techniques, replacing missing categorical values with the mode and numerical values with the median. These preprocessing steps were essential to prepare the dataset accurately and enable the Random Forest algorithm to effectively learn patterns related to consumer satisfaction with AI technologies in digital retail.

3.3. Random Forest Algorithm

Random Forest is an ensemble learning algorithm widely used for classification and regression tasks. It operates by constructing multiple decision trees during training, where each tree is trained on a bootstrap sample (random subset with replacement) of the data. At each split in a tree, only a random subset of features is considered, which introduces diversity among trees and reduces correlation. The final prediction is made by aggregating the predictions of all trees, typically through majority voting for classification problems. Mathematically, if $T_1, T_2 ..., T_n$ are the individual decision trees, the Random Forest prediction $\hat{y}$ for an input x is:

$$\hat{y} = mode\ \{T_1(x), T_2(x), \dots, T_n(x)\} \tag{1}$$

This algorithm was chosen for the study because of its robustness against overfitting, especially in datasets with high dimensionality and mixed data types such as demographic and consumer behavior variables. Random Forest effectively handles both numerical and categorical data (after encoding) and provides a measure of feature importance, allowing identification of the most influential factors affecting consumer satisfaction with AI in digital retail. Table 2 summarizes the main hyperparameters used in the Random Forest model applied in this study along with their descriptions.

Table 2. Key Hyperparameters of the Random Forest Algorithm

Hyperparameter	Description
n_estimators	The number of trees in the forest.
max_depth	The maximum depth of each tree, controlling model complexity.
min_samples_split	The minimum number of samples required to split an internal node.
min_samples_leaf	The minimum number of samples required to be at a leaf node.

Tuning these parameters helps balance bias and variance, optimizing model performance. Given these strengths, Random Forest is well-suited for analyzing and predicting satisfaction based on complex consumer data.

3.4. Data Analysis

The dataset was first divided into training and testing subsets using an 80:20 split to allow model training and unbiased evaluation on unseen data. Hyperparameter tuning was performed on the training set using Grid Search with 5-fold cross-validation to identify the optimal values for critical parameters such as the number of trees (n_estimators), tree

ISSN 2776-8007

depth (max_depth), minimum samples required to split a node (min_samples_split), and minimum samples required at a leaf node (min_samples_leaf). This process helps improve model performance by systematically searching across predefined parameter combinations. After tuning, the best model configuration was trained on the entire training data and then evaluated on the test set. Model performance was assessed using metrics such as accuracy, precision, recall, and F1-score, complemented by visual tools like confusion matrices and ROC curves to provide comprehensive insight into classification effectiveness.

After training the model, its performance was evaluated using several key metrics to comprehensively measure classification quality. Accuracy represents the proportion of total correct predictions among all samples and is calculated as:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{2}$$

TP and TN are true positives and true negatives, while FP and FN are false positives and false negatives.

Precision measures the accuracy of positive predictions and is defined as:

$$Precision = \frac{TP}{TP+FP} \tag{3}$$

Indicating the proportion of correctly predicted positive samples out of all predicted positives.

Recall (also known as sensitivity) evaluates the model's ability to identify all relevant positive cases:

$$Recall = \frac{TP}{TP+FN} \tag{4}$$

Representing the proportion of true positives detected out of actual positives.

The F1-Score is the harmonic mean of precision and recall, balancing the two metrics as:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{5}$$

This metric is particularly useful when classes are imbalanced, as it accounts for both false positives and false negatives. These metrics collectively offer a thorough evaluation of the model's classification performance on the test set.

4. Results and Discussion

4.1. Result

Table 1 presents the evaluation metrics of the Random Forest classification model predicting consumer satisfaction with AI technology usage in digital retail. The model achieves a precision of 0.84 and a recall of 0.97 for the "Satisfied" class, indicating it correctly identifies most satisfied consumers with high accuracy. However, for the "Unsatisfied" class, precision drops to 0.70 and recall is notably lower at 0.27, suggesting the model struggles to detect dissatisfied consumers effectively. The F1-score, which balances precision and recall, is 0.90 for the "Satisfied" class but only 0.39 for the "Unsatisfied" class. Overall accuracy stands at 83%, reflecting good general classification performance. The macro average scores, which treat all classes equally, highlight moderate performance with values of 0.77 (precision), 0.62 (recall), and 0.65 (F1-score). Meanwhile, the weighted average, accounting for class imbalance, shows slightly higher scores, emphasizing the model's stronger performance on the majority class. These results indicate the model is reliable at identifying satisfied consumers but has limited ability in detecting dissatisfaction, likely due to class imbalance.

Table 1. Performance Metrics of the Random Forest Model

Class	Precision	Recall	F1-Score	Support
Satisfied	0.84	0.97	0.90	106
Unsatisfied	0.70	0.27	0.39	26
Accuracy			0.83	132

Figure 2 displays a heatmap representing the Grid Search cross-validation (CV) accuracy scores for different combinations of the hyperparameters n_estimators and max_depth in a machine learning model. The color intensity varies according to the accuracy values, with darker shades indicating higher accuracy. The results reveal that the highest CV accuracy, approximately 0.8563, is achieved with a lower max_depth of 5 and either 50 or 200 n_estimators. As max_depth increases to 10 and 20, the accuracy slightly decreases across all values of n_estimators. This suggests that shallower trees with a moderate number of estimators tend to perform better in this model configuration. Figure 2 thus helps in identifying the optimal hyperparameter settings to maximize model performance.

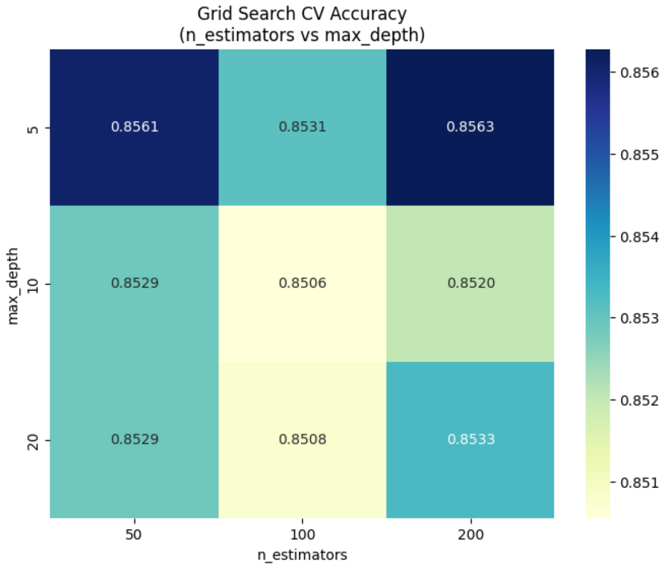

Figure 2. Grid Search CV Accuracy

Figure 3 illustrates the feature importance of various demographic and consumer behavior factors after model tuning. The chart highlights that AI_Enhance_Experience is the most influential feature, significantly contributing to the model's predictions, followed by Online_Service_Preference. Other features such as Country, Education, and Product_Category_Clothing also show moderate importance but are considerably less impactful compared to the top two factors. The remaining features, including age, gender, and payment methods, contribute minimally to the model's overall performance. This distribution suggests that user experience and online service preferences play a critical role in shaping the outcomes, while demographic details have a relatively smaller influence.

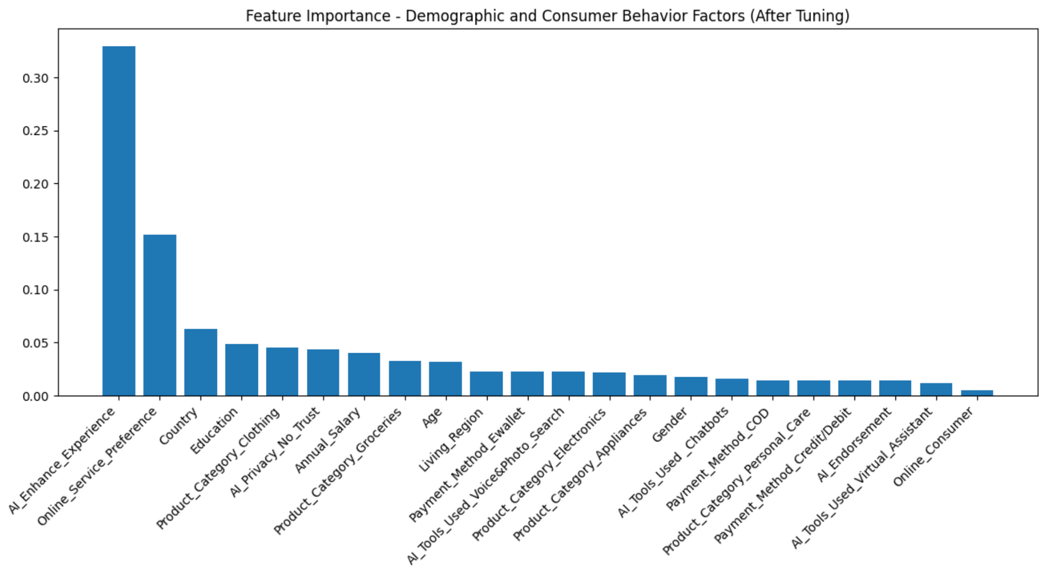

Figure 3. Feature Importance

ISSN 2776-8007

Figure 4 presents a confusion matrix illustrating the performance of a classification model predicting customer satisfaction. The matrix compares the actual customer satisfaction labels against the predicted ones. The model correctly classified 103 satisfied customers and 7 unsatisfied customers, indicating true positives for both classes. However, it misclassified 3 satisfied customers as unsatisfied and 19 unsatisfied customers as satisfied. This suggests that while the model performs well in identifying satisfied customers, it struggles more with accurately predicting unsatisfied customers, resulting in a higher number of false positives in that category.

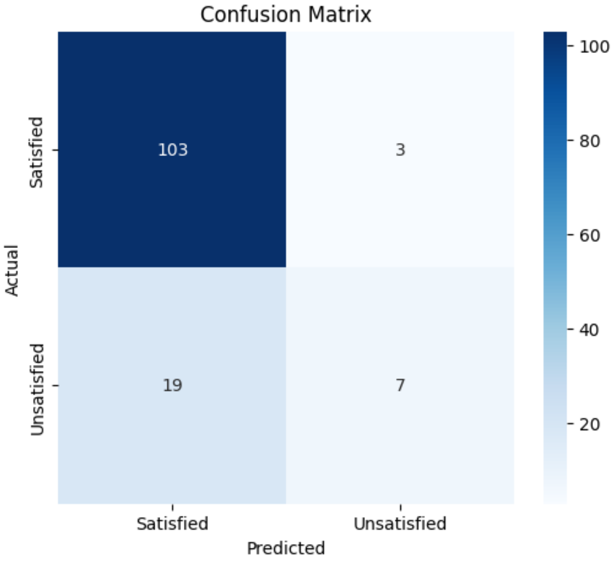

Figure 4. Confusion Matrix

Figure 5 shows the multi-class ROC curve for the “Satisfied” class in the classification model. The ROC curve plots the true positive rate against the false positive rate at various threshold settings. The area under the curve (AUC) for the “Satisfied” class is 0.18, indicating poor model performance in distinguishing satisfied customers from unsatisfied ones. The curve lies mostly close to the diagonal line, which represents random guessing, suggesting that the model has limited discriminatory power for this class.

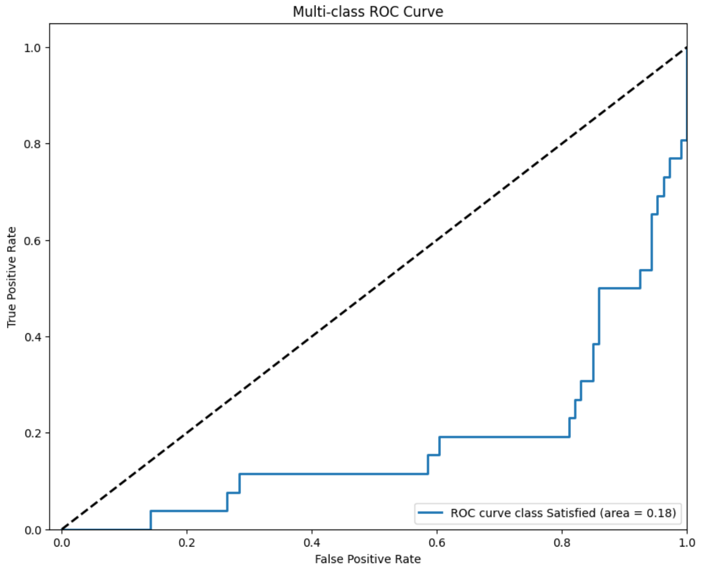

Figure 5. Multi-class ROC Curve

Figure 6 presents a correlation matrix illustrating the relationships among various demographic, behavioral, and AI-related features with AI satisfaction. The matrix highlights both positive and negative correlations, with the strength of correlations visually represented through a color gradient ranging from blue (negative correlation) to red (positive correlation). Notably, features such as "AI_Enhance_Experience" and "AI_Tools_Used_Virtual_Assistant" exhibit a moderate positive correlation with "AI_Satisfaction," suggesting that users who have enhanced experiences or frequently use virtual assistants tend to report higher satisfaction levels. Conversely, variables like "Living_Region" and "Country" show a negative correlation with AI satisfaction, indicating potential regional differences in user perceptions. Other factors, including age, gender, and payment methods, demonstrate relatively weak correlations with AI satisfaction, reflecting more subtle or complex influences. Overall, this figure provides valuable insights into which factors most significantly relate to AI satisfaction, guiding future improvements and targeted user engagement strategies.

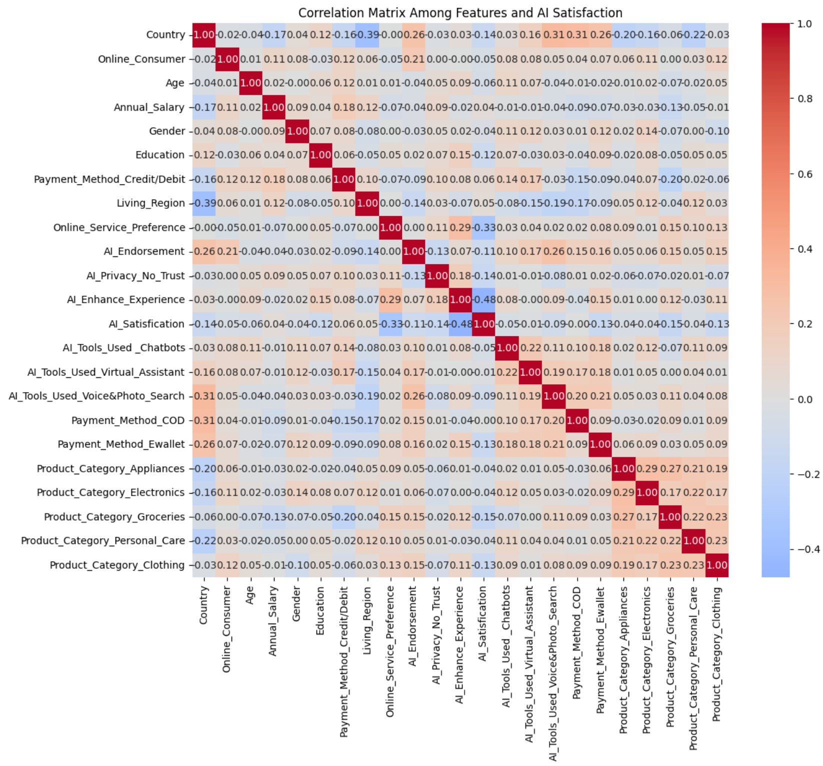

Figure 6. Correlations Matrix

4.2. Discussion

The findings of this study confirm and extend previous literature on user satisfaction with AI technologies in digital retail. The Random Forest model demonstrated strong predictive performance, particularly in identifying satisfied consumers, achieving a high precision (0.84), recall (0.97), and F1-score (0.90). This supports earlier work by Luo [15] and Huang [38], who showed that Random Forest is well-suited for modeling complex consumer satisfaction data due to its capacity to handle multi-dimensional features and imbalanced datasets. However, the model's poor performance in detecting unsatisfied consumers (recall = 0.27, F1-score = 0.39) highlights a limitation frequently observed in prior studies namely, the challenge of class imbalance in real-world data [40]. This was evident in the confusion matrix, where most unsatisfied consumers were misclassified, indicating a need for advanced balancing techniques or hybrid models to enhance minority class detection.

Feature importance analysis underscores the centrality of experience-related variables such as AI_Enhance_Experience and Online_Service_Preference, affirming the claims by Raji et al. [13], Lee & Park [4], and Srivastava & Pal [21] that personalization and service quality significantly shape consumer satisfaction. In contrast, demographic variables such as age, gender, and education showed relatively low importance, which aligns with findings from Sudirjo & Tjahyadi [35] and Prakash et al. [25] that such factors are contextual and less predictive when experiential factors dominate. The correlation matrix further validates these findings, revealing strong associations between AI-specific interactions and satisfaction, consistent with the Technology Acceptance Model (TAM) which emphasizes perceived usefulness and ease of use [24], as well as Customer Satisfaction Theory, where performance-expectation alignment determines satisfaction [42].

The grid search optimization illustrates that shallower trees (max_depth = 5) with moderate estimators yield the best cross-validation accuracy (0.8563), demonstrating that model simplicity can enhance generalization. This supports Kumar and Kumar [39] conclusion that hyperparameter tuning plays a critical role in achieving reliable classification outcomes.

The practical implications of this research are significant for digital retailers aiming to enhance customer satisfaction using AI tools. First, businesses should focus on improving the experiential quality of AI applications particularly personalization and service enhancement features as these are the strongest predictors of satisfaction. The findings suggest that instead of over-investing in demographic-based segmentation, firms should invest in real-time, data-driven personalization strategies. Moreover, the model's limitation in identifying unsatisfied users implies a need for more nuanced approaches, such as integrating sentiment analysis or feedback loops to better capture dissatisfaction signals, thereby improving retention strategies.

This study contributes to existing knowledge by integrating advanced machine learning techniques with established theoretical models to provide a nuanced understanding of AI satisfaction in digital retail. While prior research has addressed AI effectiveness and consumer trust [32], [22], this study is among the first to empirically evaluate satisfaction predictors using Random Forest and highlight the trade-off between model precision and class balance in real-world satisfaction data. Additionally, it offers insight into hyperparameter effects on model accuracy, which has practical relevance for applied data science in retail.

Despite its contributions, the study has several limitations. First, the imbalance in class distribution constrained the model’s ability to detect dissatisfied consumers. Future research should consider techniques like SMOTE, ensemble methods, or cost-sensitive learning to address this issue. Second, the dataset focused largely on behavioral and demographic features, potentially overlooking affective or contextual variables such as emotional engagement or situational constraints, which have been emphasized by Miah et al. [37]. Third, the ROC-AUC value for the “Satisfied” class (0.18) suggests limited discriminatory power, warranting exploration of alternative models (e.g., XGBoost or neural networks) and deeper feature engineering.

Lastly, this research is based on a specific digital retail context, which may limit generalizability across industries. Future studies should replicate the analysis in sectors such as healthcare, finance, or education, where AI satisfaction drivers may differ. Qualitative studies could also complement this work by exploring the psychological and emotional dimensions of satisfaction not captured in quantitative models.

5. Conclusion

This study provides an in-depth analysis of the demographic and behavioral factors influencing consumer satisfaction with AI technology usage in digital retail, using the Random Forest algorithm. The findings reveal that experiential attributes, particularly enhanced AI experience and online service preferences, are the most influential predictors of satisfaction, while demographic variables play a relatively minor role. The Random Forest model demonstrated high accuracy in identifying satisfied consumers but struggled to detect dissatisfied ones due to class imbalance, a common challenge in consumer behavior datasets.

This research contributes theoretically by integrating the Technology Acceptance Model (TAM) and Customer Satisfaction Theory to interpret the results and practically by offering insights for digital retailers on how to personalize AI-driven experiences to foster satisfaction and loyalty. The application of machine learning for predictive satisfaction modeling, especially using hyperparameter-tuned Random Forest, presents a methodological advancement with real-world relevance.

However, limitations such as class imbalance, limited emotional and contextual variables, and industry-specific focus should be addressed in future research. Expanding the model with qualitative insights and applying it across different sectors could enhance its generalizability and explanatory power. Overall, this study underscores the importance of user-centric AI development and data-driven personalization strategies to succeed in the evolving digital retail landscape.

ISSN 2776-8007

6. Declarations

6.1. Author Contributions

Conceptualization: E.P., A.S., P.A.P.; Methodology: E.P., A.S.; Software: P.A.P.; Validation: E.P., A.S.; Formal Analysis: P.A.P.; Investigation: E.P.; Resources: A.S.; Data Curation: P.A.P.; Writing – Original Draft Preparation: P.A.P.; Writing – Review and Editing: E.P., A.S.; Visualization: P.A.P.; All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] J. Chen, “Analysis of Walmarts Marketing Strategy in the Context of New Retail,” *Adv. Econ. Manag. Polit. Sci.*, vol. 44, no. 1, hal. 21–27, 2023, doi: 10.54254/2754-1169/44/20232176.

[2] R. Bhagat, V. S. Chauhan, dan P. Bhagat, “Investigating the Impact of Artificial Intelligence on Consumer’s Purchase Intention in E-Retailing,” *Foresight*, vol. 25, no. 2, hal. 249–263, 2022, doi: 10.1108/fs-10-2021-0218.

[3] M. Tiutiu dan D. Dabija, “Improving Customer Experience Using Artificial Intelligence in Online Retail,” *Proc. Int. Conf. Bus. Excell.*, vol. 17, no. 1, hal. 1139–1147, 2023, doi: 10.2478/picbe-2023-0102.

[4] M. Lee dan J. Park, “Do Parasocial Relationships and the Quality of Communication With AI Shopping Chatbots Determine Middle-aged Women Consumers’ Continuance Usage Intentions?,” *J. Consum. Behav.*, vol. 21, no. 4, hal. 842–854, 2022, doi: 10.1002/cb.2043.

[5] H. E. Ai-zhong dan Y. Zhang, “AI-powered Touch Points in the Customer Journey: A Systematic Literature Review and Research Agenda,” *J. Res. Interact. Mark.*, vol. 17, no. 4, hal. 620–639, 2022, doi: 10.1108/jrim-03-2022-0082.

[6] S. Josimovski, L. P. Ivanovska, dan D. Dodevski, “Understanding the Consumer Dynamics of AI in North Macedonian E-Business,” *Econ. Cult.*, vol. 20, no. 2, hal. 64–75, 2023, doi: 10.2478/jec-2023-0016.

[7] S. J. Sardesai, E. D’Souza, dan S. Govekar, “Analysing the Impacts of Artificial Intelligence Service Quality and Human Service Quality on Customer Satisfaction and Customer Loyalty in the Hospitality Sector,” *Turizam*, vol. 28, no. 1, hal. 37–48, 2024, doi: 10.5937/turizam28-45450.

[8] N. Pandey, N. Gupta, S. Rastogi, R. R. Singh, dan M. Mishra, “Digital Coupons and Gen Z: An Application of Technology Acceptance Model With Coupon Proneness as a Moderator,” *Innov. Mark.*, vol. 20, no. 2, hal. 65–77, 2024, doi: 10.21511/im.20(2).2024.06.

[9] D. Girollet, “Digital Divides Among Microenterprises: Evidence From Sub-Saharan Africa,” *J. Int. Dev.*, vol. 36, no. 2, hal. 1350–1380, 2023, doi: 10.1002/jid.3860.

[10] A. Devisakti, M. Muftahu, dan X. Hu, “Digital Divide Among B40 Students in Malaysian Higher Education Institutions,” *Educ. Inf. Technol.*, vol. 29, no. 2, hal. 1857–1883, 2023, doi: 10.1007/s10639-023-11847-w.

[11] M. Rahevar dan S. Darji, “The Adoption of Ai-Driven Chatbots Into a Recommendation for E-Commerce Systems to Targeted Customer in the Selection of Product,” *Ijmec*, vol. 1, no. 2, hal. 128–137, 2024, doi: 10.62737/m1vpdq75.

[12] X. Cheng, X. Zhang, B. Yang, dan Y. Fu, “An Investigation on Trust in AI-enabled Collaboration: Application of AI-Driven Chatbot in Accommodation-Based Sharing Economy,” *Electron. Commer. Res. Appl.*, vol. 54, no. July, hal. 101164, 2022, doi: 10.1016/j.elerap.2022.101164.

[13] M. A. Raji, H. B. Olodo, T. T. Oke, W. A. Addy, O. C. Ofodile, dan A. T. Oyewole, “E-Commerce and Consumer Behavior: A Review of AI-powered Personalization and Market Trends,” *GSC Adv. Res. Rev.*, vol. 18, no. 3, hal. 66–77, 2024, doi: 10.30574/gscarr.2024.18.3.0090.

[14] W. Wang *et al.*, “A User Purchase Behavior Prediction Method Based on XGBoost,” *Electronics*, vol. 12, no. 9, hal. 1–17, 2023, doi: 10.3390/electronics12092047.

[15] R. Luo, “Improved Random Forest Based on Grid Search for Customer Satisfaction Prediction,” *Adv. Econ. Manag. Polit. Sci.*, vol. 38, no. 1, hal. 198–207, 2023, doi: 10.54254/2754-1169/38/20231913.

[16] A. Ridwan, U. Muzakir, dan S. Nurhidayati, “Optimizing E-Commerce Inventory to Prevent Stock Outs Using the Random Forest Algorithm Approach,” *Int. J. Softw. Eng. Comput. Sci.*, vol. 4, no. 1, hal. 107–120, 2024, doi: 10.35870/ijsecs.v4i1.2326.

[17] S. McKee, S. Sands, J. Pallant, dan E. Cohen, “The Evolving Direct-to-consumer Retail Model: A Review and Research Agenda,” *Int. J. Consum. Stud.*, vol. 47, no. 6, hal. 2816–2842, 2023, doi: 10.1111/ijcs.12972.

[18] P. N. Nair dan G. Trehan, “Examining the Use of Ai-Powered Social Media Analytics for Target Customer Segmentation: A Systematic Review in Retail Industry,” *Educ. Adm. Theory Pract.*, vol. 30, no. 4, hal. 1798–1805, 2024, doi: 10.53555/kuey.v30i4.1753.

[19] J. M. Lopes, L. Silva, dan I. Massano-Cardoso, “AI Meets the Shopper: Psychosocial Factors in Ease of Use and Their Effect on E-Commerce Purchase Intention,” *Behav. Sci. (Basel).*, vol. 14, no. 7, hal. 616, 2024, doi: 10.3390/bs14070616.

[20] R. Ciuchita, J. Gummerus, M. Holmlund, dan E. L. Linhart, “Programmatic Advertising in Online Retailing: Consumer Perceptions and Future Avenues,” *J. Serv. Manag.*, vol. 34, no. 2, hal. 231–255, 2022, doi: 10.1108/josm-06-2021-0238.

[21] K. Srivastava dan D. Pal, “Importance of AI Attributes in Indian Retail Stores: A Conjoint Analysis Approach,” *Int. J. Retail Distrib. Manag.*, vol. 52, no. 3, hal. 355–371, 2024, doi: 10.1108/ijrdm-11-2022-0456.

[22] C. R. Guerra-Tamez, K. K. Flores, G. M. Serna-Mendiburu, D. C. Robles, dan J. I. Cortés, “Decoding Gen Z: AI’s Influence on Brand Trust and Purchasing Behavior,” *Front. Artif. Intell.*, vol. 7, no. March, hal. 1–14, 2024, doi: 10.3389/frai.2024.1323512.

[23] H. A. D. M. Arachchi dan D. Samarasinghe, “Impact of Embedded AI Mobile Smart Speech Recognition On consumer Attitudes Towards AI and Purchase Intention Across Generations X and Y,” *Eur. J. Manag. Stud.*, vol. 29, no. 1, hal. 3–29, 2023, doi: 10.1108/ejms-03-2023-0019.

[24] S. Dutta, “Revolutionizing Retail: An Empirical Study on the Impact of Generative AI in Omnichannel Strategies,” *Int. J. Multidiscip. Res.*, vol. 6, no. 5, hal. 1–12, 2024, doi: 10.36948/ijfmr.2024.v06i05.28799.

[25] P. Prakash, P. Kumar, P. Kishore, D. Jaganathan, S. Immanuel, dan S. V Raj, “Determinant of Access to Credit and Availing Subsidies for Protected Cultivation in Maharashtra,” *Indian J. Ext. Educ.*, vol. 58, no. 2, hal. 167–172, 2022, doi: 10.48165/ijee.2022.58240.

[26] U. Bashir, S. M. Mohideen, dan M. A. Sheikh, “Evaluating the Influence of Frontline Demonstration and Training on Vermicomposting Adoption Among Farmers,” *Skuast J. Res.*, vol. 26, no. 1, hal. 126–130, 2024, doi: 10.5958/2349-297x.2024.00015.1.

[27] D. Pandita, V. Bhatt, V. V. R. Kumar, A. Fatma, dan F. Vapiwala, “Electrifying the Future: Analysing the Determinants of Electric Vehicle Adoption,” *Int. J. Energy Sect. Manag.*, vol. 18, no. 6, hal. 1767–1786, 2024, doi: 10.1108/ijesm-06-2023-0004.

[28] M. M. H. Emon, “Unveiling the Progression Towards Solar Power Adoption: A Comprehensive Analysis of Understanding, Awareness, and Acceptance of Solar Technology in Bangladesh,” *Econ. Growth Environ. Sustain.*, vol. 2, no. 2, hal. 105–111, 2023, doi: 10.26480/egnes.02.2023.105.111.

[29] N. Wexler dan Y. Fan, “Gauging Public Attitudes and Preferences Toward a Hypothetical Future Public Shared Automated Vehicle System: Examining the Roles of Gender, Race, Income, and Health,” *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2676, no. 10, hal. 588–600, 2022, doi: 10.1177/03611981221090512.

[30] M. Sallam *et al.*, “Assessing Health Students’ Attitudes and Usage of ChatGPT in Jordan: Validation Study,” *JMIR Med. Educ.*, vol. 9, no. e48254, hal. 1–15, 2023, doi: 10.2196/48254.

[31] G. O. Ikoyo-Eweto, I. F. Adedokun, J. Archibong, dan G. F. Okwuokenye, “Rural Farmers Access to Extension Services:

Implications for Increased Adoption of Improved Farm Technologies in Delta State, Nigeria," *J. Agric. Food Sci.*, vol. 21, no. 2, hal. 135–152, 2024, doi: 10.4314/jafs.v21i2.11.

[32] G. ElSayad dan H. Mamdouh, "Are Young Adult Consumers Ready to Be Intelligent Shoppers? The Importance of Perceived Trust and the Usefulness of AI-powered Retail Platforms in Shaping Purchase Intention," *Young Consum. Insight Ideas Responsible Mark.*, vol. 25, no. 6, hal. 969–989, 2024, doi: 10.1108/yc-02-2024-1991.

[33] R. S. Sumi dan M. Ahmed, "Investigating Young Consumers' Online Buying Behavior In COVID-19 Pandemic: Perspective of Bangladesh," *Iim Ranchi J. Manag. Stud.*, vol. 1, no. 2, hal. 108–123, 2022, doi: 10.1108/irjms-09-2021-0127.

[34] N. A. Murshed dan E. Uğurlu, "Navigating the Digital Marketplace: A Holistic Model Integrating Social Media Engagement and Consumer Behavior Factors to Enhance Online Shopping Adoption," *J. Theory Appl. Manag.*, vol. 16, no. 3, hal. 542–559, 2023, doi: 10.20473/jmtt.v16i3.52059.

[35] F. Sudirjo dan I. Tjahyadi, "Assessing Factors Influencing Online Shopping Adoption Among Indonesian Consumers: A Quantitative Study," *West Sci. Interdiscip. Stud.*, vol. 1, no. 05, hal. 51–57, 2023, doi: 10.58812/wsis.v1i05.73.

[36] A. Rakate dan H. V Gaikwad, "The Impact of Social Media in Islamic Studies on Consumer Behavior Towards Smart Watches With Special Reference for the City of Islampur," *Relig. Educ.*, vol. 2, no. 2, hal. 80–91, 2022, doi: 10.17509/re.v2i2.51151.

[37] M. R. Miah, A. Hossain, R. Shikder, T. Saha, dan M. Neger, "Evaluating the Impact of Social Media on Online Shopping Behavior During COVID-19 Pandemic: A Bangladeshi Consumers' Perspectives," *Heliyon*, vol. 8, no. 9, hal. e10600, 2022,

[38] W. Huang, "Analysis of Promotional Online Shopping Behavior Based on Machine Learning," *Highlights Sci. Eng. Technol.*, vol. 56, no. July, hal. 65–72, 2023, doi: 10.54097/hset.v56i.9817.

[39] B. Kumar dan S. M. Kumar, "Recognition of Idleness Among Patients Based on Activity Using Random Forest Over Decision Tree Algorithm.," *J. Pharm. Negat. Results*, vol. 13, no. 04, hal. 1803–1808, 2022, doi: 10.47750/pnr.2022.13.s04.217.

[40] J. Ma, Q. Chen, H. Li, Y. Chen, Y. Lu, dan H. Yang, "A Model for Identifying Glass Artifacts Based on SOM Clustering Analysis and Random Forest Algorithm," *Highlights Sci. Eng. Technol.*, vol. 62, no. July, hal. 164–173, 2023, doi: 10.54097/hset.v62i.10439.

[41] Y. Liu, "Discussion on the Enterprise Financial Risk Management Framework Based on AI Fintech," *Decis. Mak. Appl. Manag. Eng.*, vol. 7, no. 1, hal. 254–269, 2023, doi: 10.31181/dmame712024942.

[42] O. Irshad, S. Ahmad, dan S. Mahmood, "Fostering Purchase Intentions Through CSR and Service Quality: The Role of Customer Satisfaction, Brand Loyalty, and Admiration," *Sustainability*, vol. 16, no. 23, hal. 10584, 2024, doi: 10.3390/su162310584.

[43] Q. Zhang dan S. Ahmad, "Linking Corporate Social Responsibility, Consumer Identification and Purchasing Intention," *Sustainability*, vol. 14, no. 19, hal. 12552, 2022, doi: 10.3390/su141912552.