

Original Research Article

Impact of AI-Based Fraud Prevention on Customer Trust in Digital Payment Systems: An Empirical Investigation of Continuance Intention among Indian Users

Dr. Devanjali Dutta*Assistant Professor, K.M. Agrawal College of Arts, Commerce and Science, India*

Received: 10 September, 2026 | Accepted: 24 September, 2026 | Published: 6 October, 2026

Abstract

The rapid expansion of digital payment ecosystems in India, anchored by the Unified Payments Interface (UPI), has been accompanied by a sharp rise in payment fraud, eroding consumer confidence and threatening long-term adoption. Artificial intelligence (AI)-driven fraud prevention mechanisms—real-time anomaly detection, behavioural biometrics, and adaptive authentication—are increasingly deployed by payment service providers to mitigate this risk. Drawing on the Expectation-Confirmation Model of Information Systems Continuance (Bhattacherjee, 2001) and the trust literature (Mayer, Davis, & Schoorman, 1995), this study empirically examines the influence of perceived AI security features on customers' continuance intention toward digital payment applications. A cross-sectional survey of 350 active digital payment users in India was conducted using a structured, five-point Likert-scale instrument. Pearson correlation analysis revealed a very strong, positive, and statistically significant association between AI security features and continuance intention ($r = .900$, $p < .001$, 95% CI [.878, .918]). Simple linear regression confirmed that AI security features significantly predict continuance intention, explaining 76.7% of the variance ($R^2 = .767$, $F(1, 348) = 1147.00$, $p < .001$; $\beta = .876$, $t = 33.87$, $p < .001$). The findings substantiate the proposition that visible, intelligent fraud-prevention capabilities function as a salient trust-building cue, thereby reinforcing users' decision to continue using digital payment platforms. The paper discusses theoretical and managerial implications, acknowledges limitations, and outlines directions for future research.

Keywords: artificial intelligence; fraud prevention; digital payments; customer trust; continuance intention; expectation-confirmation model; UPI; India

1. Introduction

India's digital payments landscape has undergone a structural transformation since the launch of the Unified Payments Interface (UPI) in 2016, with transaction value on the network surpassing ₹200 trillion in the two years preceding 2024 (Bloomberg, 2024). This expansion has democratised access to financial services and accelerated the country's transition toward a less-cash economy. However, the same architectural features that make digital payments convenient—real-time settlement, interoperability, and mobile-first delivery—have also widened the attack surface for malicious actors. According to the Reserve Bank of India's Annual Report 2023-24, the value of frauds reported in the card- and internet-based transaction segment rose more than fivefold to approximately ₹14.57 billion in fiscal year 2023-24, up from ₹2.77 billion in the previous year (Reserve Bank of India, 2024). Such figures underscore an emerging paradox: the very systems designed to extend financial inclusion are simultaneously generating novel vectors of consumer harm.

Artificial intelligence (AI) and machine learning (ML) have emerged as the principal technological response to this challenge. Payment networks, issuing banks, and third-party application providers now deploy AI-driven solutions that analyse transaction streams in milliseconds, flag anomalous behavioural patterns, and adapt authentication thresholds to perceived risk (Mastercard, 2025; Adewumi & Akinyelu, 2017). Industry evidence suggests that AI-based systems not only improve detection accuracy but also reduce false positives, thereby lowering customer friction and churn (Mastercard, 2025). Yet, from a consumer-behaviour standpoint, a critical question remains: do users' perceptions of these AI-driven safeguards translate into higher trust and a stronger intention to continue using the underlying payment platform? Answering this question is consequential because the long-term viability of digital payment ecosystems depends not on first-time adoption alone but on sustained, voluntary continuance (Bhattacharjee, 2001).

Anchored in the Expectation-Confirmation Model (ECM) of Information Systems Continuance (Bhattacharjee, 2001) and complemented by the integrative trust framework of Mayer, Davis, and Schoorman (1995), this study empirically investigates the relationship between perceived AI security features and continuance intention in the context of Indian digital payment applications. Using a cross-sectional survey of 350 users, the research tests the central hypothesis that perceived AI-based fraud prevention exerts a significant positive effect on continuance intention. The remainder of the paper is organised as follows: Section 2 reviews the relevant literature and develops the research hypothesis; Section 3 details the methodology; Section 4 presents the empirical results; Section 5 discusses the findings; Section 6 addresses limitations and future research directions; and Section 7 concludes.

2. Literature Review and Hypothesis Development

2.1 Digital Payments, Fraud, and the Indian Context

Digital payment systems encompass any instrument that enables the electronic transfer of value between transacting parties, including mobile wallets, UPI applications, and internet-banking platforms. India occupies a distinctive position in this domain: UPI processed the majority of the country's retail non-cash payments in 2024-25, with digital transactions comprising 99.9% of retail non-cash volumes (Reserve Bank of India, 2025). This

concentration creates systemic importance but also systemic risk. PwC India (2024) observes that the growth of digital payments has been paralleled by increasingly sophisticated fraud typologies, including mule-account networks, social engineering, and phishing via spoofed domains. The Reserve Bank of India has responded with a risk-based authentication framework and AI-driven initiatives such as MuleHunter.ai and the Digital Payments Intelligence Platform (Reserve Bank of India, 2025). These institutional responses reflect a growing consensus that technological countermeasures are indispensable for sustaining consumer confidence.

2.2 AI-Based Fraud Prevention in Payment Systems

AI-based fraud prevention refers to the application of machine-learning algorithms, behavioural analytics, and real-time decisioning engines to identify, score, and intercept potentially fraudulent transactions. Adewumi and Akinyelu (2017), in their survey of machine-learning and nature-inspired credit card fraud detection techniques, document the gradual shift from rule-based systems to supervised, unsupervised, and hybrid models that can adapt to evolving fraud tactics. More recent industry evidence indicates that AI not only detects fraud with greater precision but also reduces false-positive rates, which are a leading source of customer dissatisfaction and declined legitimate transactions (Mastercard, 2025; Federal Reserve Bank of Philadelphia, 2024). According to Mastercard's 2025 research, 83% of industry leaders report that AI has reduced both false positives and customer churn, marking what the report characterises as a new era in fraud prevention (Mastercard, 2025). The Philadelphia Federal Reserve (2024) similarly reports that, after recalibrating AI fraud-detection systems, one large bank reduced low-value false-positive escalations by 25-30% without compromising true-fraud detection.

2.3 Trust and Continuance Intention

Trust has long been theorised as a multidimensional construct comprising ability, benevolence, and integrity (Mayer, Davis, & Schoorman, 1995). In technology-mediated contexts, trust operates as a psychological mechanism that reduces perceived risk and enables commitment in the absence of complete information. Continuance intention, conceptualised by Bhattacharjee (2001) in the Expectation-Confirmation Model, refers to a user's intention to continue using an information system after initial acceptance. The ECM posits that continuance intention is determined by satisfaction and perceived usefulness, which are in turn shaped by the confirmation of prior expectations. Subsequent research has extended the ECM by integrating trust as an antecedent of both perceived usefulness and continuance intention (Shao et al., 2019; Yan et al., 2024). In the mobile payment literature specifically, Shao et al. (2019) demonstrate that platform characteristics such as security and reputation positively influence trust and continuance intention, while Yan et al. (2024), in a study of 674 mobile payment users, identify perceived usefulness, self-efficacy, and subjective norms as significant drivers of continuance intention.

2.4 Hypothesis Development

Integrating the preceding streams, AI-based fraud prevention can be theorised as a salient trust cue that signals the platform's technical ability to protect users' financial assets. According to Mayer et al.'s (1995) ability dimension of trust, demonstrated competence in fraud detection strengthens the user's perception that the service provider possesses the requisite expertise to safeguard transactions. This perception, in turn, satisfies the perceived-usefulness criterion of the ECM (Bhattacharjee, 2001), as users come to view continued engagement with the platform as both safe and beneficial. Chawla and Joshi (2019), in an empirical study of Indian consumers' mobile-wallet adoption, similarly find that perceived security is among the strongest predictors of adoption intention. Building on this convergent evidence, the present study tests the following hypothesis:

H1: Perceived AI-based fraud prevention features have a significant positive effect on users' continuance intention toward digital payment applications.

3. Methodology

3.1 Research Design and Approach

The study adopts a quantitative, cross-sectional, survey-based research design consistent with the dominant methodological tradition in technology-adoption and continuance-intention research (Venkatesh, Morris, Davis, & Davis, 2003; Bhattacharjee, 2001). A positivist epistemology underpins the investigation, treating perceived AI security features and continuance intention as latent constructs that can be measured through reflective indicators and modelled using parametric statistics. The unit of analysis is the individual digital payment user.

3.2 Sampling and Data Collection

Data were collected from 350 active users of Indian digital payment applications through a structured online questionnaire distributed via professional and academic networks. A combination of purposive and convenience sampling was employed, with the screening criterion that respondents had used at least one digital payment application in the preceding 30 days. The sample size exceeds the threshold of 200 recommended by Hair, Black, Babin, and Anderson (2010) for multivariate techniques and provides adequate statistical power (>0.99) to detect medium-to-large effects at $\alpha = .05$. Table 1 summarises the demographic profile of the sample.

Table 1. Demographic Profile of Respondents (N = 350)

Variable	Category	n	%
Gender	Male	191	54.6
	Female	159	45.4
Age	18-25	132	37.7
	26-35	113	32.3
	36-45	74	21.1
	46 and above	31	8.9
Education	Undergraduate	159	45.4
	Postgraduate	123	35.1
	Professional	56	16.0
	Doctorate	12	3.4
Occupation	Private employee	139	39.7
	Student	90	25.7
	Business	52	14.9
	Government employee	44	12.6

	Self-employed	25	7.1
Monthly income (INR)	Less than 25,000	107	30.6
	25,001-50,000	140	40.0
	50,001-75,000	69	19.7
	Above 75,000	34	9.7
Primary payment app	Google Pay	115	32.9
	PhonePe	96	27.4
	Paytm	67	19.1
	Bank app	44	12.6
	BHIM	28	8.0

Note. Percentages may not sum to 100 due to rounding.

As Table 1 indicates, the sample is gender-balanced (54.6% male; 45.4% female) and skewed toward younger, educated users, with 70% of respondents aged between 18 and 35 and over 80% holding at least an undergraduate degree. Private-sector employees (39.7%) and students (25.7%) form the largest occupational groups, and the income distribution captures the full range from low- to high-earners. Google Pay (32.9%), PhonePe (27.4%), and Paytm (19.1%) dominate the primary-payment-app category, broadly mirroring national market-share patterns.

3.3 Measurement Instrument

Two reflective constructs were operationalised using established, multi-item scales adapted from prior literature. Perceived AI security features (five items, AI1-AI5) were adapted from the security-perception items used in mobile-payment trust research (Shao et al., 2019) and reframed to emphasise AI-specific cues such as real-time monitoring, behavioural analytics, and adaptive authentication. Continuance intention (five items, CI1-CI5) was measured using the scale proposed by Bhattacherjee (2001). All items were rated on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Internal-consistency reliability was assessed using Cronbach's alpha, with values of .70 or above considered acceptable (Nunnally, 1978).

3.4 Data Analysis Strategy

Descriptive statistics were computed for all items and constructs. The bivariate relationship between perceived AI security features and continuance intention was examined using Pearson's product-moment correlation. To test H1, a simple linear regression was estimated with continuance intention as the dependent variable and AI security features as the sole predictor. Model fit was evaluated using R, R^2 , the omnibus F-test, and the standardised and unstandardised coefficients with their 95% confidence intervals. All analyses were conducted on the complete sample of $N = 350$.

4. Results

4.1 Descriptive Statistics

Table 2 reports the descriptive statistics for the ten measured items. Item means ranged from 3.60 to 3.73 on the five-point scale, with medians of 4.00 throughout, indicating an overall positive perception of AI security features and a favourable continuance intention. Standard deviations hovered around 0.90, suggesting moderate dispersion and limited ceiling effects. No item exhibited skewness or kurtosis values exceeding the conventional thresholds of |2| and |7| respectively, supporting the assumption of univariate normality required for the parametric analyses that follow.

Table 2. Descriptive Statistics for Study Items (N = 350)

Item	N	Mean	Median	SD	Min	Max
AI1	350	3.66	4.00	0.891	2	5
AI2	350	3.61	4.00	0.913	2	5
AI3	350	3.60	4.00	0.970	1	5
AI4	350	3.61	4.00	0.925	2	5
AI5	350	3.65	4.00	0.913	1	5
CI1	350	3.73	4.00	0.926	2	5
CI2	350	3.67	4.00	0.923	1	5
CI3	350	3.67	4.00	0.910	1	5
CI4	350	3.71	4.00	0.884	1	5
CI5	350	3.71	4.00	0.918	1	5

Note. AI1-AI5 = perceived AI security feature items; CI1-CI5 = continuance intention items. SD = standard deviation.

4.2 Correlation Analysis

Pearson's product-moment correlation was computed to assess the bivariate association between the two composite constructs. As reported in Table 3, the analysis revealed a very strong, positive, and statistically significant relationship between perceived AI security features and continuance intention ($r = .900$, $p < .001$). The 95% confidence interval for the correlation coefficient ranged from .878 to .918, indicating a tightly bounded and substantively large effect. By Cohen's (1988) conventions, this value substantially exceeds the threshold for a large effect ($r \geq .50$), suggesting that the two constructs share approximately 81% of their variance.

Table 3. Pearson Correlation Matrix for Study Constructs (N = 350)

Variable	AI Security Features	Continuance Intention
AI Security Features	—	.900***
Continuance Intention	.900***	—

Note. * $p < .05$, ** $p < .01$, *** $p < .001$ (two-tailed).

4.3 Regression Analysis

A simple linear regression was estimated to test H1, with continuance intention modelled as a function of perceived AI security features. The overall model was statistically significant, $F(1, 348) = 1147.00$, $p < .001$, and explained 76.7% of the variance in continuance intention ($R^2 = .767$), as reported in Table 4. The ANOVA decomposition confirmed that the model sum of squares (4103) substantially exceeded the residual sum of squares (1245),

indicating that perceived AI security features account for the lion's share of explained variance in the dependent variable.

Table 4. Regression Model Fit Summary

Model	R	R ²	F	df1	df2	p
1	.876	.767	1147.00	1	348	<.001

Note. Model estimated using $N = 350$. Predictor: AI security features; dependent variable: continuance intention.

Table 5. Regression Coefficients

Predictor	B	SE	β	t	p	95% CI for B
(Intercept)	1.776	0.5036	—	3.53	<.001	[0.786, 2.767]
AI Security Features	0.936	0.0276	.876	33.87	<.001	[0.881, 0.990]

Note. B = unstandardised coefficient; SE = standard error; β = standardised coefficient; CI = confidence interval. Dependent variable: continuance intention.

As Table 5 shows, the coefficient for AI security features was positive, large, and highly significant ($B = 0.936$, $SE = 0.0276$, $\beta = .876$, $t = 33.87$, $p < .001$). The 95% confidence interval for the unstandardised coefficient ([0.881, 0.990]) excluded zero by a wide margin, reinforcing the robustness of the effect. Collectively, these results provide strong empirical support for H1: perceived AI-based fraud prevention features exert a significant positive effect on users' continuance intention toward digital payment applications.

5. Discussion

The findings of this study provide robust empirical support for the proposition that perceived AI-based fraud prevention is a powerful determinant of continuance intention in digital payment systems. The correlation of $r = .900$ and the regression R^2 of .767 indicate that AI security perceptions account for nearly four-fifths of the explained variance in continuance intention—a magnitude that is unusually high compared with typical technology-adoption studies, where R^2 values of .30 to .50 are more common (Venkatesh et al., 2003). Several theoretical and contextual factors help explain this pronounced effect.

First, the result aligns with the ability dimension of the integrative trust model (Mayer et al., 1995). AI-driven fraud detection is highly visible to users through real-time alerts, biometric prompts, and post-transaction summaries, and these visible cues function as direct evidence of the platform's competence. Such competence signals reduce perceived transaction risk and, in doing so, satisfy the perceived-usefulness criterion specified in Bhattacharjee's (2001) Expectation-Confirmation Model. Second, the findings resonate with the broader mobile-payment literature. Shao et al. (2019) demonstrate that platform-level security characteristics positively influence both trust and continuance intention, and the present study extends this evidence by isolating AI-specific security features as a particularly potent trust cue. Third, the result is consistent with industry evidence that AI reduces false positives and customer friction (Mastercard, 2025; Federal Reserve Bank of Philadelphia, 2024); by minimising declined legitimate transactions, AI not only protects users but also preserves the seamless experience that drives continued use.

From a managerial perspective, the implications are twofold. For payment service providers, the findings suggest that investments in AI-based fraud prevention should be conceptualised not merely as cost-side risk controls but as strategic assets that directly reinforce user retention. Communicating these capabilities—through transparent in-app messaging, security dashboards, and post-incident explanations—can amplify the trust effect. For regulators, the evidence supports continued investment in shared AI infrastructure such as the Reserve Bank of India's MuleHunter.ai and Digital Payments Intelligence Platform (Reserve Bank of India, 2025), as institution-level AI capability appears to translate into user-level continuance behaviour. The result also carries implications for financial inclusion: if AI-driven security sustains trust among existing users, it may also lower the psychological barrier to adoption among prospective users who currently perceive digital payments as risky.

It is important to interpret these results in light of the broader theoretical literature on continuance behaviour. While Bhattacharjee (2001) emphasises satisfaction and perceived usefulness as proximal drivers of continuance intention, the present study suggests that, in security-sensitive contexts such as digital payments, perceived security may function as an upstream antecedent that shapes both satisfaction and usefulness. This repositioning is consistent with Yan et al. (2024), who identify perceived usefulness as a key mediator between antecedent beliefs and continuance intention, and with Chawla and Joshi (2019), who document the primacy of perceived security in the Indian mobile-wallet context.

6. Limitations and Future Research

This study has several limitations that should inform the interpretation of its findings and guide subsequent research. First, the cross-sectional design precludes causal inference; while the regression results are consistent with a causal effect of AI security perceptions on continuance intention, reverse causation and omitted-variable bias cannot be ruled out. Longitudinal or experimental designs would strengthen causal claims. Second, the reliance on self-report measures introduces the risk of common-method bias and social-desirability effects, particularly for a construct as value-laden as security perception. Future research could complement survey data with behavioural measures such as actual transaction frequency or retention rates.

Third, the sample is concentrated among younger, educated, urban users and over-represents private-sector employees and students, which may limit the generalisability of the findings to older, rural, or less digitally literate populations. Stratified sampling across demographic and geographic segments would improve external validity. Fourth, the model is intentionally parsimonious, focusing on a single predictor; future studies could integrate additional constructs such as perceived ease of use, social influence, facilitating conditions (Venkatesh et al., 2003), and perceived risk to assess the incremental explanatory contribution of AI security features. Finally, comparative research across countries with varying fraud-prevention infrastructures would help disentangle the role of institutional context in shaping the AI-security-trust relationship.

7. Conclusion

This study set out to examine the impact of AI-based fraud prevention on customer trust and continuance intention in digital payment systems. Drawing on the Expectation-Confirmation Model (Bhattacharjee, 2001) and the integrative trust framework (Mayer et al., 1995), and using survey data from 350 Indian digital payment users, the analysis documents a very strong, positive, and statistically significant relationship between perceived AI security features and continuance intention ($r = .900$, $p < .001$). Linear regression further demonstrates that AI security features explain 76.7% of the variance in continuance intention ($\beta = .876$, $p < .001$), providing strong support for the central hypothesis.

The findings carry a clear message for payment service providers, regulators, and platform designers: AI-based fraud prevention is not merely a defensive technology but a strategic trust-building asset that sustains user engagement. As digital payment ecosystems continue to expand, the visible demonstration of intelligent, adaptive security will be increasingly central to maintaining the consumer confidence on which long-term adoption depends. Future research that combines longitudinal designs, behavioural measures, and cross-cultural comparisons will further enrich understanding of how AI reshapes the architecture of trust in digital finance.

References

- Adewumi, A. O., & Akinyelu, A. A. (2017). A survey of machine-learning and nature-inspired based credit card fraud detection techniques. *International Journal of System Assurance Engineering and Management*, 8(2), 937-953. <https://doi.org/10.1007/s13198-016-0551-y>
- Bhattacharjee, A. (2001). Understanding information systems continuance: An expectation-confirmation model. *MIS Quarterly*, 25(3), 351-370. <https://doi.org/10.2307/3250921>
- Bloomberg. (2024, May 30). Online payment frauds jump over 400% in India, RBI data shows. Bloomberg. <https://www.bloomberg.com/news/articles/2024-05-30/online-payment-frauds-jump-over-400-in-india-rbi-data-shows>
- Shao, Z., Zhang, L., Li, X., & Guo, Y. (2019). Antecedents of trust and continuance intention in mobile payment platforms: The moderating effect of gender. *Electronic Commerce Research and Applications*, 33, 100823. <https://doi.org/10.1016/j.elerap.2018.100823>
- Chawla, D., & Joshi, H. (2019). Consumer attitude and intention to adopt mobile wallet in India: An empirical study. *International Journal of Bank Marketing*, 37(7), 1590-1618. <https://doi.org/10.1108/IJBM-09-2018-0276>
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340. <https://doi.org/10.2307/249008>
- Federal Reserve Bank of Philadelphia. (2024). AI-enabled fraud is on the rise. Federal Reserve Bank of Philadelphia. <https://www.philadelphiafed.org/-/media/FRBP/Assets/Economy/Articles/ai-enabled-fraud-is-on-the-rise-article.pdf>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Pearson Prentice Hall.
- Mastercard. (2025). On the right side of AI: Shaping the future of payment fraud prevention. Mastercard. <https://www.mastercard.com/content/dam/mccom/shared/news-and-trends/insights/2025/one-the-right-side-of-ai/the-right-side-of-ai-shaping-the-future-of-payment-fraud-prevention.pdf>
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organisational trust. *Academy of Management Review*, 20(3), 709-734. <https://doi.org/10.2307/258792>
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill.

- PwC India. (2024). Combating payments fraud in India's digital payments landscape. PricewaterhouseCoopers India. <https://www.pwc.in/ghost-templates/combating-payments-fraud-in-Indias-digital-payments-landscape.html>
- Reserve Bank of India. (2024). Annual report 2023-24. Reserve Bank of India. https://rbidocs.rbi.org.in/rdocs/AnnualReport/PDFs/0ANNUALREPORT202324_FULLLDF549205FA214F62A2441C5320D64A29.PDF
- Reserve Bank of India. (2025). Annual report 2024-25: Payment and settlement systems. Reserve Bank of India. <https://www.rbi.org.in/Scripts/AnnualReportMainDisplay.aspx>
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425-478. <https://doi.org/10.2307/30036540>
- Yan, Y., Chen, K.-H., Zhao, Y., & Liu, F. (2024). Factors influencing continuance intention to use mobile payment: A perspective of user adaptation. *Humanities and Social Sciences Communications*, 11, Article 540. <https://doi.org/10.1057/s41599-024-02778-z>