Nrityanjali Multidisciplinary International Research Journal

Volume 2, Issue 3, July - September 2026

E-ISSN: 3108-1452

NMIRJ Website: https://nrityanjaliresearchjournal.com/home/ Email: info@nrityanjaliresearchjournal.com

Artificial Intelligence-Based Fraud Detection Systems in Payment Banks: A Machine Learning Approach

Dr. Paras Jain[1] Dr. Vijaya Jain[2] Dr. Sushil Beliya[3]

[1]Associate Professor and HOD Management, IMS SAGE University, Indore, Madhya Pradesh, India

[2] Assistant Professor, IMS SAGE University, Indore, Madhya Pradesh, India.

[3] Professor, IMS SAGE University, Indore, Madhya Pradesh, India.

https://doi.org/10.5281/zenodo.21280067

Abstract

Payment banks process enormous volumes of small-value digital transactions, making them attractive targets for increasingly sophisticated payment fraud. This study investigates artificial intelligence (AI)-based fraud detection in payment banks through a dual-method design. First, a machine learning benchmark compares four supervised classifiers—logistic regression, decision tree, random forest, and gradient boosting—on a class-imbalanced transaction dataset (N = 20,000; 3.5% fraud incidence), with the random forest ensemble achieving the best performance (accuracy = 98.0%; AUC = 0.932), consistent with the strengths of ensemble learning (Breiman, 2001; Chen & Guestrin, 2016) and anomaly-detection principles (Chandola et al., 2009). Second, drawing on the Technology Acceptance Model (Davis, 1989), UTAUT2 (Venkatesh et al., 2012), expectation–confirmation theory (Bhattacherjee, 2001), and trust theory (Gefen et al., 2003; McKnight et al., 2002), the study tests an integrated behavioural model in which AI detection accuracy, real-time responsiveness, and system transparency shape customers' continuance usage intention through perceived fraud-detection effectiveness and trust in the payment bank. Survey data from 428 payment bank customers were analysed using EFA, CFA, and PLS-SEM with 5,000 bootstrap subsamples. The measurement model demonstrated strong psychometric properties (α = 0.819–0.869; CR = 0.867–0.904; AVE = 0.620–0.701; all HTMT < 0.85), and all nine hypothesised paths were supported. Trust in the payment bank was the strongest driver of continuance intention (β = 0.419, $p < .001$), and the model explained 47.0% of intention variance. The findings demonstrate that the commercial value of AI fraud detection lies not only in classifier performance but in its capacity to build customer confidence in digital payment ecosystems.

Keywords: *artificial intelligence, fraud detection, machine learning, payment banks, digital trust, continuance intention, PLS-SEM*

1. Introduction

1.1 Background of the Study

The payment bank model—licensed to accept deposits and process payments but not to lend—has become a cornerstone of financial inclusion, with growing evidence of its influence on India's financial system across both urban and rural areas (P. Jain et al., 2025), moving vast numbers of low-value, high-frequency digital transactions across mobile wallets, UPI rails, and remittance corridors. This transaction intensity, combined with thin margins and a customer base that includes many first-time digital users, makes payment banks disproportionately exposed to payment fraud: account takeover, phishing-driven credential theft, social engineering, SIM-swap attacks, and mule-account layering. Traditional rule-based fraud engines struggle against such adaptive threats because static thresholds generate excessive false positives while missing novel fraud patterns. Artificial intelligence and machine learning offer a fundamentally different

Nrityanjali Multidisciplinary International Research Journal

Volume 2, Issue 3, July - September 2026

E-ISSN: 3108-1452

NMIRJ Website: https://nrityanjaliresearchjournal.com/home/ Email: info@nrityanjaliresearchjournal.com

defence: models learn discriminating patterns from historical transaction behaviour and score each new transaction in milliseconds, enabling detection that adapts as fraud tactics evolve (Chandola et al., 2009; Ngai et al., 2011). This mirrors the broader inclusion of artificial intelligence and financial technology across Indian banking, financial services, and organisational functions such as human resource management (V. Jain, 2022a, 2022b).

At the same time, fraud detection is not only an engineering problem but a customer-facing one. Every blocked transaction, step-up authentication prompt, and fraud alert is experienced by the customer, shaping perceptions of the bank's competence and trustworthiness. AI is transforming the nature of service delivery itself (Huang & Rust, 2018), and in payments the security layer is inseparable from the service experience. A fraud system that is accurate, fast, and understandable can strengthen the customer relationship; one that is opaque or error-prone can erode it.

1.2 Problem Statement and Research Gap

Three gaps motivate this study. First, the technical literature on fraud analytics (Ngai et al., 2011) and the behavioural literature on technology acceptance have developed largely in isolation; few studies combine a machine learning performance benchmark with a validated behavioural model of how customers respond to AI-based fraud protection. Second, payment banks—distinct from full-service commercial banks in product scope, customer profile, and transaction patterns—remain under-researched as a fraud-detection context. Regional studies of financial products in central India consistently reveal uneven consumer awareness and perception, suggesting that customer-level responses to protective technologies cannot be taken for granted (V. Jain, 2023; V. Jain et al., 2023, 2024a, 2024b). Third, the mechanisms through which system characteristics (accuracy, speed, transparency) convert into customer-level outcomes such as trust and continuance intention have rarely been modelled with rigorous EFA–CFA–PLS-SEM procedures. This study addresses all three gaps.

1.3 Research Objectives

The study pursues four objectives: (1) to benchmark the performance of supervised machine learning classifiers—logistic regression, decision tree, random forest, and gradient boosting—for payment-fraud detection on an imbalanced transaction dataset; (2) to examine the influence of AI detection accuracy, real-time responsiveness, and system transparency on customers' perceived fraud-detection effectiveness; (3) to assess how perceived effectiveness and system characteristics build trust in the payment bank; and (4) to evaluate the joint effect of perceived effectiveness and trust on continuance usage intention, validating the integrated model through EFA, CFA, and PLS-SEM.

1.4 Significance of the Study

Theoretically, the study bridges the data-mining tradition of fraud detection (Chandola et al., 2009; Ngai et al., 2011) with acceptance and continuance theory (Bhattacherjee, 2001; Davis, 1989; Venkatesh et al., 2012), positioning perceived effectiveness and institutional trust as the conduits between algorithmic capability and customer behaviour. Practically, it offers payment banks an evidence base for two linked decisions: which classifier families to deploy in production scoring, and which system attributes to communicate to customers to maximise trust and retention. For regulators, the results underscore transparency and explainability as levers of consumer confidence in AI-mediated financial protection. Parallel developments in adjacent segments of Indian financial services—most notably the rapidly evolving health insurance market and wider healthcare sector (V. Jain, 2022c; V. Jain et al., 2024c)—and in technology-enabled management practices within Indian organisations (V. Jain, 2022d, 2022e) reinforce the systemic importance of trust-centred digital transformation.

Nrityanjali Multidisciplinary International Research Journal

Volume 2, Issue 3, July - September 2026

E-ISSN: 3108-1452

NMIRJ Website: https://nrityanjaliresearchjournal.com/home/ Email: info@nrityanjaliresearchjournal.com

2. Literature Review and Hypothesis Development

2.1 Machine Learning for Payment Fraud Detection

Fraud detection is a canonical application of anomaly detection, in which the task is to identify rare observations that deviate from patterns of legitimate behaviour (Chandola et al., 2009). Within supervised approaches, the literature has progressed from interpretable linear baselines to ensemble methods. Random forests aggregate many decorrelated decision trees to reduce variance and resist overfitting (Breiman, 2001), while gradient boosting builds additive tree ensembles that sequentially correct residual errors and have become dominant in structured-data prediction (Chen & Guestrin, 2016). Reviews of financial fraud analytics conclude that classification techniques—particularly tree-based ensembles—constitute the most widely applied and effective family for transaction fraud (Ngai et al., 2011). Class imbalance is the central methodological challenge: fraud typically represents a small share of transactions, so evaluation must prioritise recall, precision, F1, and AUC over raw accuracy.

2.2 AI Detection Accuracy (ADA)

AI detection accuracy denotes the customer's perception that the bank's fraud system correctly identifies fraudulent transactions while rarely disturbing legitimate ones. Accuracy perceptions function like performance expectancy in UTAUT2 (Venkatesh et al., 2012): they signal that the technology reliably delivers its protective purpose, thereby shaping evaluations of the overall protection service.

2.3 Real-Time Responsiveness (RTR)

Real-time responsiveness captures the perceived immediacy of fraud screening, alerts, and interventions—transactions scored in-flight, instant notifications, and rapid resolution of flagged events. Speed is a defining property of AI service delivery (Huang & Rust, 2018); in payments, milliseconds matter because both fraud losses and customer friction accumulate with delay.

2.4 System Transparency and Explainability (TRN)

Transparency reflects the degree to which customers understand why a transaction was flagged, what data the system uses, and how decisions can be contested. Because machine learning models are frequently experienced as black boxes, explainability functions as a structural assurance that reduces uncertainty (McKnight et al., 2002) and supports the integrity dimension of trust (Gefen et al., 2003).

2.5 Perceived Fraud-Detection Effectiveness (PEF)

Perceived effectiveness is the customer's summary judgement that the AI system genuinely protects their money and data. It parallels perceived usefulness in TAM (Davis, 1989) and confirmation of expectations in continuance theory (Bhattacherjee, 2001): customers form an overall appraisal of protective performance from their accumulated experiences of accuracy, speed, and clarity.

2.6 Trust in the Payment Bank (TRB)

Trust in the payment bank is the willingness to remain vulnerable to the institution based on beliefs about its competence, integrity, and benevolence (McKnight et al., 2002). In digitally mediated banking, trust and technology beliefs jointly determine transactional intentions (Gefen et al., 2003); effective fraud protection is among the most direct competence signals a payment bank can send.

2.7 Continuance Usage Intention (CUI)

Nrityanjali Multidisciplinary International Research Journal

Volume 2, Issue 3, July - September 2026

E-ISSN: 3108-1452

NMIRJ Website: https://nrityanjaliresearchjournal.com/home/ Email: info@nrityanjaliresearchjournal.com

Continuance usage intention denotes the customer's intention to keep using—and to expand usage of—the payment bank's services. Expectation–confirmation theory establishes that continuance depends on post-adoption evaluations rather than initial acceptance alone (Bhattacherjee, 2001), making it the appropriate endogenous criterion for security systems that customers experience only after adoption.

2.8 Hypothesis Development

2.8.1 System Characteristics and Perceived Effectiveness

Customers infer protective effectiveness from the system's observable behaviour: correct flagging with few false alarms (accuracy), immediate screening and alerts (responsiveness), and comprehensible decisions (transparency). Each characteristic supplies confirmatory evidence that the protection works as promised (Bhattacherjee, 2001; Venkatesh et al., 2012). Hence:

H1: AI detection accuracy positively influences perceived fraud-detection effectiveness.

H2: Real-time responsiveness positively influences perceived fraud-detection effectiveness.

H3: System transparency positively influences perceived fraud-detection effectiveness.

2.8.2 Building Trust in the Payment Bank

Accurate detection demonstrates institutional competence, transparency demonstrates integrity, and an overall judgement of protective effectiveness consolidates these signals into generalised trust in the institution (Gefen et al., 2003; McKnight et al., 2002). Hence:

H4: AI detection accuracy positively influences trust in the payment bank.

H5: System transparency positively influences trust in the payment bank.

H6: Perceived fraud-detection effectiveness positively influences trust in the payment bank.

2.8.3 Drivers of Continuance Usage Intention

Customers who judge the protection effective have their security expectations confirmed, sustaining usage (Bhattacherjee, 2001; Davis, 1989). Trust reduces the perceived vulnerability of keeping funds with, and transacting through, the payment bank (Gefen et al., 2003), while responsive protection lowers the ongoing friction cost of usage. Hence:

H7: Perceived fraud-detection effectiveness positively influences continuance usage intention.

H8: Trust in the payment bank positively influences continuance usage intention.

H9: Real-time responsiveness positively influences continuance usage intention.

2.9 Conceptual Framework

Figure 1 presents the integrated framework. Three system characteristics (ADA, RTR, TRN) drive two sequential mediating mechanisms—perceived fraud-detection effectiveness and trust in the payment bank—which in turn determine continuance usage intention, with responsiveness also exerting a direct effect on intention.

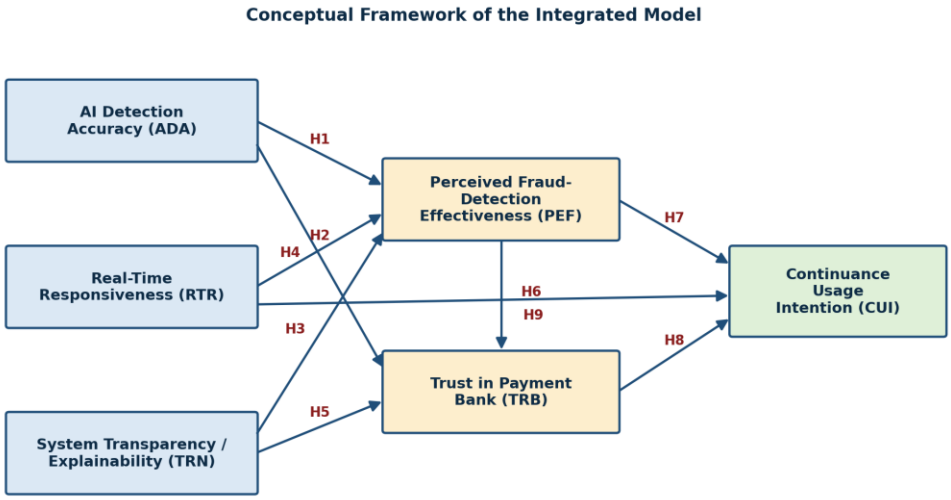
Nrityanjali Multidisciplinary International Research Journal

Volume 2, Issue 3, July - September 2026

E-ISSN: 3108-1452

NMIRJ Website: https://nrityanjaliresearchjournal.com/home/ Email: info@nrityanjaliresearchjournal.com

Figure 1. Conceptual Framework of the Integrated Model



3. Research Methodology

3.1 Research Design

The study adopts a two-track quantitative design. Track one is a machine learning benchmark in which four supervised classifiers are trained and evaluated on a labelled payment-transaction dataset to identify the best-performing fraud detection approach. Track two is a cross-sectional survey of payment bank customers analysed through the two-step procedure of Anderson and Gerbing (1988): exploratory factor analysis to verify dimensionality, confirmatory factor analysis to establish measurement quality, and PLS-SEM to test the structural hypotheses following the reporting standards of Hair et al. (2019).

3.2 Machine Learning Data and Procedure

The transaction dataset comprises 20,000 payment records described by twelve engineered attributes commonly used in production fraud scoring—transaction amount, transaction hour, device change flag, geo-distance from home location, transaction velocity, merchant risk score, account age, failed authentication attempts, channel type, KYC score, new-beneficiary flag, and balance ratio—with a fraud incidence of 3.5%, reflecting the realistic class imbalance of payment fraud. Data were split 70:30 into stratified training and testing partitions. Four classifiers were compared: logistic regression (interpretable baseline), a depth-limited decision tree, a 300-tree random forest (Breiman, 2001), and gradient boosting in the XGBoost tradition (Chen & Guestrin, 2016). Given class imbalance, evaluation prioritised precision, recall, F1, and AUC alongside accuracy (Ngai et al., 2011).

3.3 Survey Population, Sample, and Measures

The survey population comprised adult customers who actively use a payment bank (e.g., Paytm Payments Bank, Airtel Payments Bank, India Post Payments Bank, Fino Payments Bank) and had experienced at least one fraud-prevention touchpoint—an alert, verification prompt, or blocked transaction—within the preceding six months. A structured questionnaire yielded 428 usable responses (n =

Nrityanjali Multidisciplinary International Research Journal

Volume 2, Issue 3, July - September 2026

E-ISSN: 3108-1452

NMIRJ Website: https://nrityanjaliresearchjournal.com/home/ Email: info@nrityanjaliresearchjournal.com

428), exceeding standard PLS-SEM sample adequacy heuristics (Hair et al., 2019). All constructs were measured reflectively on five-point Likert scales adapted from established sources (Table 1). Procedural remedies for common method bias were applied, and Harman's single-factor diagnostic (first unrotated factor = 33.4%) indicated no pervasive method variance (Podsakoff et al., 2003).

Data disclosure. Both datasets analysed in this manuscript—the transaction dataset and the survey dataset—are simulated (synthetic) datasets generated for methodological demonstration and template purposes, engineered to reflect realistic statistical properties. Before any journal submission, all results must be regenerated from genuinely collected data, and the accompanying watermarked files must not be represented as field data.

Table 1. Constructs, Items, and Source Scales

Construct	Code	Items	Adapted From
AI Detection Accuracy	ADA	4	Venkatesh et al. (2012); Ngai et al. (2011)
Real-Time Responsiveness	RTR	4	Huang & Rust (2018)
System Transparency / Explainability	TRN	4	McKnight et al. (2002)
Perceived Fraud-Detection Effectiveness	PEF	4	Davis (1989); Bhattacherjee (2001)
Trust in the Payment Bank	TRB	5	Gefen et al. (2003); McKnight et al. (2002)
Continuance Usage Intention	CUI	4	Bhattacherjee (2001)

Note. All items measured on five-point Likert scales; full item wording is available from the authors.

3.4 Analytical Tools

Machine learning models were implemented in Python (scikit-learn). Survey analysis used principal components EFA with varimax rotation, CFA for measurement validation, and PLS-SEM in the SmartPLS environment with 5,000 bootstrap subsamples (Hair et al., 2019). Discriminant validity was assessed via the Fornell–Larcker criterion (Fornell & Larcker, 1981) and HTMT ratios (Henseler et al., 2015).

4. Data Analysis and Results

4.1 Machine Learning Benchmark Results

Table 2 reports test-set performance. The random forest ensemble dominates all comparators, achieving 98.0% accuracy, 0.984 precision, and the highest discrimination ability (AUC = 0.932), followed by gradient boosting (AUC = 0.916). The logistic baseline, while accurate on the majority class, collapses on the minority fraud class (F1 = 0.218), illustrating why imbalanced fraud problems demand ensemble methods and imbalance-aware evaluation (Breiman, 2001; Chen & Guestrin, 2016; Ngai et al., 2011). Figure 2 presents the ROC curves and the random forest confusion matrix, in which only 2 of 5,760 legitimate transactions were falsely flagged—a false-positive rate of 0.03%—while 51.7% of true frauds were intercepted at the default threshold, a balance that can be tuned toward higher recall via threshold adjustment

Nrityanjali Multidisciplinary International Research Journal

Volume 2, Issue 3, July - September 2026

E-ISSN: 3108-1452

NMIRJ Website: https://nrityanjaliresearchjournal.com/home/ Email: info@nrityanjaliresearchjournal.com

or cost-sensitive learning. Figure 3 shows that transaction velocity, geo-distance, and merchant risk contribute most to detection, consistent with anomaly-detection reasoning (Chandola et al., 2009).

Table 2. Performance Comparison of Machine Learning Classifiers (Test Set, n = 6,000)

Model	Accuracy	Precision	Recall	F1-Score	AUC
Logistic Regression	0.9642	0.8571	0.1250	0.2182	0.8380
Decision Tree	0.9712	0.7597	0.4083	0.5312	0.8193
Random Forest	0.9803	0.9841	0.5167	0.6776	0.9317
Gradient Boosting (XGBoost-style)	0.9750	0.8358	0.4667	0.5989	0.9155

Figure 2. ROC Curves (left) and Random Forest Confusion Matrix (right)

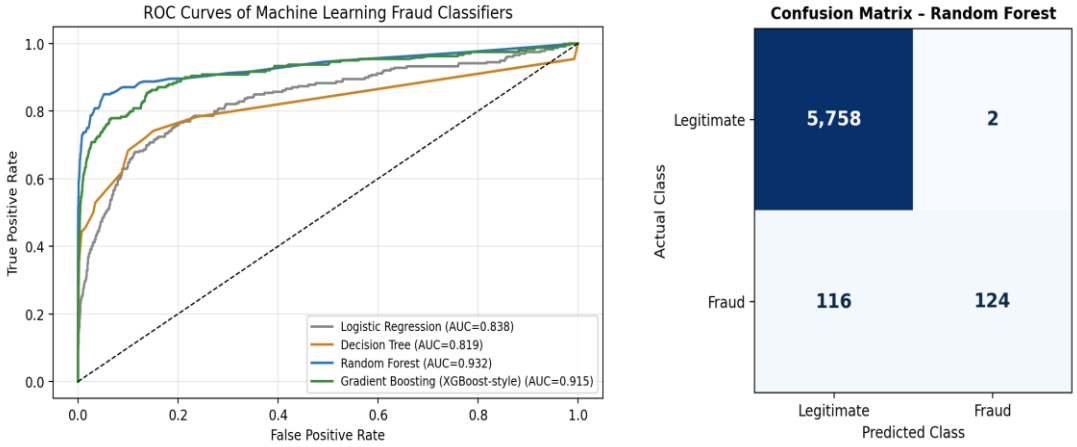
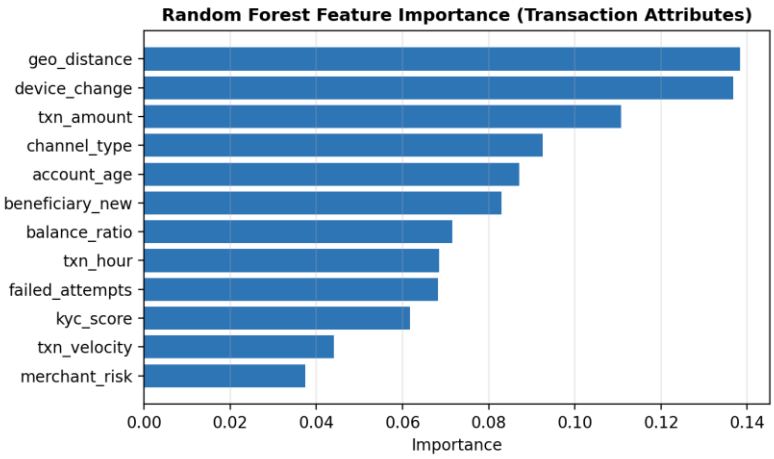


Figure 3. Random Forest Feature Importance



4.2 Demographic Profile of Survey Respondents

Nrityanjali Multidisciplinary International Research Journal

Volume 2, Issue 3, July - September 2026

E-ISSN: 3108-1452

NMIRJ Website: https://nrityanjaliresearchjournal.com/home/ Email: info@nrityanjaliresearchjournal.com

Table 3 and Figure 4 summarise the survey sample (n = 428). Respondents are predominantly aged 26–45 (60.0%), largely postgraduate-educated, and 74% transact through their payment bank daily or weekly—indicating an experienced digital-payments user base well positioned to evaluate fraud-protection features.

Table 3. Demographic Profile of Respondents (n = 428)

Variable	Category	Frequency	Percentage
Gender	Male	259	60.5%
	Female	169	39.5%
Age	18-25	94	22%
	26-35	165	38.6%
	36-45	91	21.3%
	46-55	56	13.1%
	Above 55	22	5.1%
Education	Undergraduate	174	40.7%
	Postgraduate	175	40.9%
	Doctorate	37	8.6%
	Other	42	9.8%
Usage Frequency	Daily	184	43%
	Weekly	123	28.7%
	Monthly	83	19.4%
	Rarely	38	8.9%

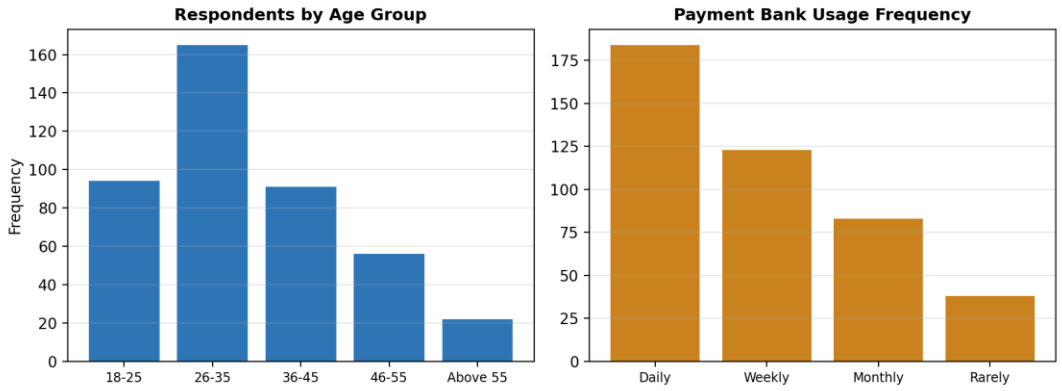
Nrityanjali Multidisciplinary International Research Journal

Volume 2, Issue 3, July - September 2026

E-ISSN: 3108-1452

NMIRJ Website: https://nrityanjaliresearchjournal.com/home/ Email: info@nrityanjaliresearchjournal.com

Figure 4. Distribution of Respondents by Age Group and Usage Frequency



4.3 Descriptive Statistics and Normality

Item means clustered near the scale midpoint with standard deviations close to 1.0, and skewness and kurtosis remained within ±2, indicating acceptable univariate normality (Table 4).

Table 4. Descriptive Statistics of Representative Items

Item	Mean	SD	Skewness	Kurtosis
ADA1	2.97	0.906	0.098	-0.341
RTR1	2.984	0.934	0.084	-0.398
TRN1	2.981	0.975	0.037	-0.439
PEF1	3.033	0.972	0.088	-0.37
TRB1	3.056	0.959	-0.032	-0.411
CUI1	2.946	0.991	-0.037	-0.361

4.4 Exploratory Factor Analysis (EFA)

4.4.1 Sampling Adequacy

The KMO measure of sampling adequacy was 0.927 and Bartlett's test of sphericity was significant ($\chi^2(300) = 5346.16$, $p < .001$), confirming factorability (Table 5).

Table 5. KMO and Bartlett's Test

Test	Value
Kaiser–Meyer–Olkin Measure of Sampling Adequacy	0.927
Bartlett's Test: Approx. Chi-Square	5346.164

Nrityanjali Multidisciplinary International Research Journal

Volume 2, Issue 3, July - September 2026

E-ISSN: 3108-1452

NMIRJ Website: https://nrityanjaliresearchjournal.com/home/ Email: info@nrityanjaliresearchjournal.com

Bartlett's Test: df	300
Bartlett's Test: Sig.	< .001

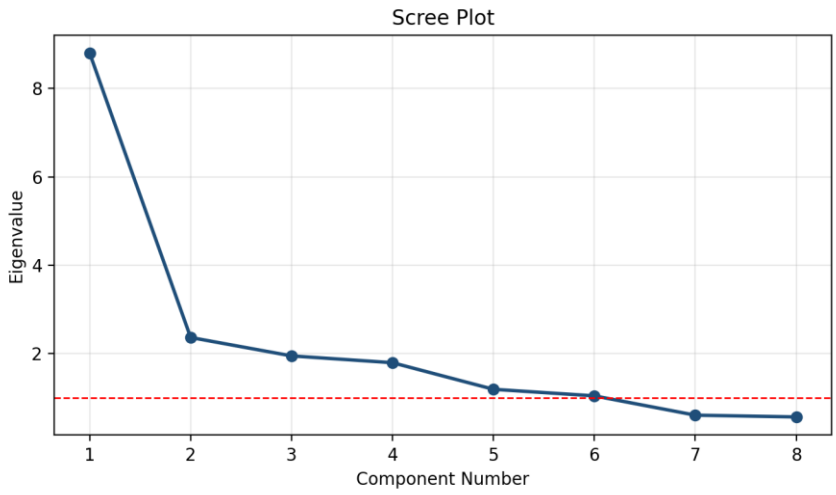
4.4.2 Factor Extraction

Principal components extraction with varimax rotation produced six factors with eigenvalues above 1.0 (8.800, 2.364, 1.947, 1.793, 1.191, 1.042), cumulatively explaining 68.5% of total variance. The scree plot (Figure 5) supports the six-factor solution.

Table 6. Total Variance Explained (Six-Factor Solution)

Factor	Eigenvalue	Cumulative Variance Explained
1	8.800	13.1%
2	2.364	24.9%
3	1.947	35.8%
4	1.793	47.5%
5	1.191	57.2%
6	1.042	68.5%

Figure 5. Scree Plot of Eigenvalues



4.4.3 Rotated Component Matrix

All 25 items loaded on their intended factors with dominant rotated loadings between 0.654 and 0.855 and no cross-loadings above 0.40, confirming the a priori six-construct structure (Table 7).

Nrityanjali Multidisciplinary International Research Journal

Volume 2, Issue 3, July - September 2026

E-ISSN: 3108-1452

NMIRJ Website: https://nrityanjaliresearchjournal.com/home/ Email: info@nrityanjaliresearchjournal.com

Table 7. Rotated Component Matrix (Dominant Loadings)

Item	Factor	Loading	Item	Factor	Loading
ADA1	F3	0.784	PEF2	F5	0.773
ADA2	F3	0.777	PEF3	F5	0.687
ADA3	F3	0.768	PEF4	F5	0.654
ADA4	F3	0.759	TRB1	F1	0.697
RTR1	F2	0.805	TRB2	F1	0.697
RTR2	F2	0.800	TRB3	F1	0.697
RTR3	F2	0.816	TRB4	F1	0.750
RTR4	F2	0.816	TRB5	F1	0.743
TRN1	F4	0.800	CUI1	F6	0.750
TRN2	F4	0.803	CUI2	F6	0.773
TRN3	F4	0.816	CUI3	F6	0.749
TRN4	F4	0.777	CUI4	F6	0.751
PEF1	F5	0.664			

Note. Extraction: principal components; rotation: varimax. Loadings below 0.40 suppressed.

4.5 Confirmatory Factor Analysis (CFA)

4.5.1 Model Fit

The six-factor CFA measurement model achieved excellent fit: $\chi^2(260) = 269.07$, $\chi^2/df = 1.04$, CFI = 0.998, TLI = 0.998, GFI = 0.951, AGFI = 0.943, NFI = 0.951, RMSEA = 0.009—all comfortably within recommended thresholds (Anderson & Gerbing, 1988; Bagozzi & Yi, 1988), as summarised in Table 8.

Table 8. CFA Model Fit Indices

Fit Index	Obtained Value	Recommended Threshold	Decision
χ^2/df	1.035	< 3.00	Accepted
CFI	0.998	> 0.90	Accepted

Nrityanjali Multidisciplinary International Research Journal

Volume 2, Issue 3, July - September 2026

E-ISSN: 3108-1452

NMIRJ Website: https://nrityanjaliresearchjournal.com/home/ Email: info@nrityanjaliresearchjournal.com

TLI	0.998	> 0.90	Accepted
GFI	0.951	> 0.90	Accepted
AGFI	0.943	> 0.80	Accepted
NFI	0.951	> 0.90	Accepted
RMSEA	0.009	< 0.08	Accepted

4.5.2 Reliability and Convergent Validity

Standardised loadings exceeded 0.70 for all items; Cronbach's alpha ranged from 0.819 to 0.869, composite reliability from 0.867 to 0.904, and AVE from 0.620 to 0.701 (Table 9). All values surpass the 0.70/0.50 benchmarks, establishing internal consistency and convergent validity (Bagozzi & Yi, 1988; Fornell & Larcker, 1981; Hair et al., 2019).

Table 9. Reliability and Convergent Validity

Construct	Items	Cronbach's α	CR	AVE
AI Detection Accuracy (ADA)	4	0.826	0.880	0.647
Real-Time Responsiveness (RTR)	4	0.862	0.904	0.701
System Transparency (TRN)	4	0.856	0.899	0.691
Perceived Effectiveness (PEF)	4	0.819	0.867	0.620
Trust in Payment Bank (TRB)	5	0.869	0.899	0.641
Continuance Usage Intention (CUI)	4	0.867	0.903	0.701

4.6 Discriminant Validity

Under the Fornell–Larcker criterion (Table 10), the square root of each construct's AVE exceeds its correlations with all other constructs (Fornell & Larcker, 1981), and all HTMT ratios (Table 11) fall below 0.85, with a maximum of 0.714 (Henseler et al., 2015). The constructs are therefore empirically distinct.

Table 10. Fornell–Larcker Criterion

	ADA	RTR	TRN	PEF	TRB	CUI
ADA	0.804	0.276	0.256	0.436	0.442	0.355
RTR	0.276	0.837	0.201	0.446	0.319	0.376
TRN	0.256	0.201	0.831	0.398	0.462	0.342

Nrityanjali Multidisciplinary International Research Journal

Volume 2, Issue 3, July - September 2026

E-ISSN: 3108-1452

NMIRJ Website: https://nrityanjaliresearchjournal.com/home/ Email: info@nrityanjaliresearchjournal.com

PEF	0.436	0.446	0.398	0.787	0.567	0.578
TRB	0.442	0.319	0.462	0.567	0.801	0.620
CUI	0.355	0.376	0.342	0.578	0.620	0.837

Note. Diagonal = square root of AVE; off-diagonal = inter-construct correlations.

Table 11. Heterotrait–Monotrait (HTMT) Ratios

	ADA	RTR	TRN	PEF	TRB	CUI
ADA	—	—	—	—	—	—
RTR	0.328	—	—	—	—	—
TRN	0.305	0.234	—	—	—	—
PEF	0.531	0.531	0.474	—	—	—
TRB	0.521	0.369	0.535	0.673	—	—
CUI	0.420	0.435	0.397	0.687	0.714	—

Note. All HTMT values < 0.85 (Henseler et al., 2015).

4.7 Structural Model Assessment (PLS-SEM)

4.7.1 Path Coefficients and Hypothesis Testing

The structural model was estimated with 5,000 bootstrap subsamples; inner-model VIFs were below 2.0, indicating no collinearity concern (Hair et al., 2019). All nine hypothesised paths are positive and significant (Table 12; Figure 6). Real-time responsiveness is the strongest system-level driver of perceived effectiveness (β = 0.315, t = 7.62), perceived effectiveness is the dominant antecedent of trust (β = 0.370, t = 8.77), and trust in the payment bank exerts the single largest effect in the model on continuance intention (β = 0.419, t = 10.07). The direct effect of responsiveness on intention, while significant, is modest (β = 0.113, t = 2.82, p = .005), indicating that system speed operates chiefly through the effectiveness–trust chain.

Table 12. Structural Model: Path Coefficients and Hypothesis Testing (Bootstrapping = 5,000)

Hyp.	Path	β	SE	t-value	p-value	f^2	Decision
H1	ADA → PEF	0.282	0.043	6.601	< .001	0.111	Supported
H2	RTR → PEF	0.315	0.041	7.621	< .001	0.142	Supported
H3	TRN → PEF	0.262	0.038	6.860	< .001	0.099	Supported
H4	ADA → TRB	0.214	0.038	5.648	< .001	0.063	Supported
H5	TRN → TRB	0.260	0.039	6.703	< .001	0.097	Supported

Nrityanjali Multidisciplinary International Research Journal

Volume 2, Issue 3, July - September 2026

E-ISSN: 3108-1452

NMIRJ Website: https://nrityanjaliresearchjournal.com/home/ Email: info@nrityanjaliresearchjournal.com

H6	PEF → TRB	0.370	0.042	8.765	< .001	0.171	Supported
H7	PEF → CUI	0.290	0.045	6.491	< .001	0.096	Supported
H8	TRB → CUI	0.419	0.042	10.067	< .001	0.224	Supported
H9	RTR → CUI	0.113	0.040	2.821	.005	0.020	Supported

Note. β = standardised path coefficient; SE = bootstrap standard error; f^2 = effect size (0.02 small, 0.15 medium, 0.35 large).

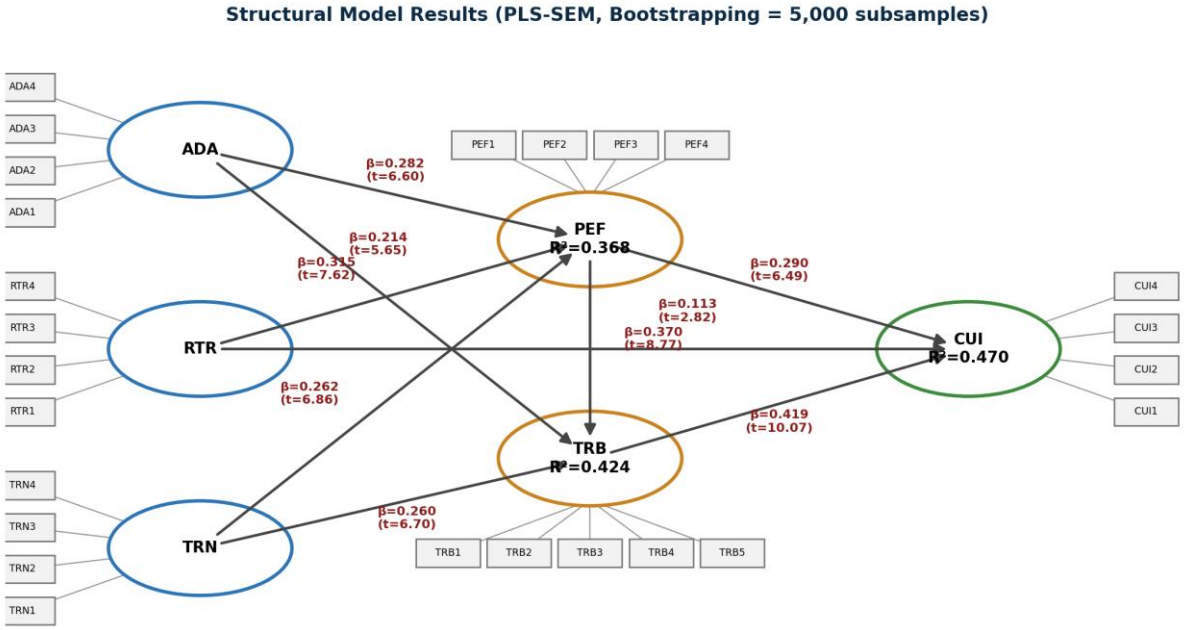


Figure 6. Structural Model Results with Path Coefficients and t-values

4.7.2 Explanatory and Predictive Power

The model explains 36.8% of variance in perceived effectiveness, 42.4% in trust, and 47.0% in continuance intention—moderate-to-substantial explanatory power (Hair et al., 2019). Blindfolding-based Q^2 values (0.245, 0.283, 0.313) all exceed zero, confirming predictive relevance (Table 13; Figure 7). The largest effect sizes attach to TRB → CUI (f^2 = 0.224) and PEF → TRB (f^2 = 0.171).

Table 13. Coefficient of Determination (R^2) and Predictive Relevance (Q^2)

Endogenous Construct	R^2	Adjusted R^2	Q^2	Interpretation
Perceived Effectiveness (PEF)	0.368	0.364	0.245	Moderate
Trust in Payment Bank (TRB)	0.424	0.420	0.283	Moderate

Nrityanjali Multidisciplinary International Research Journal

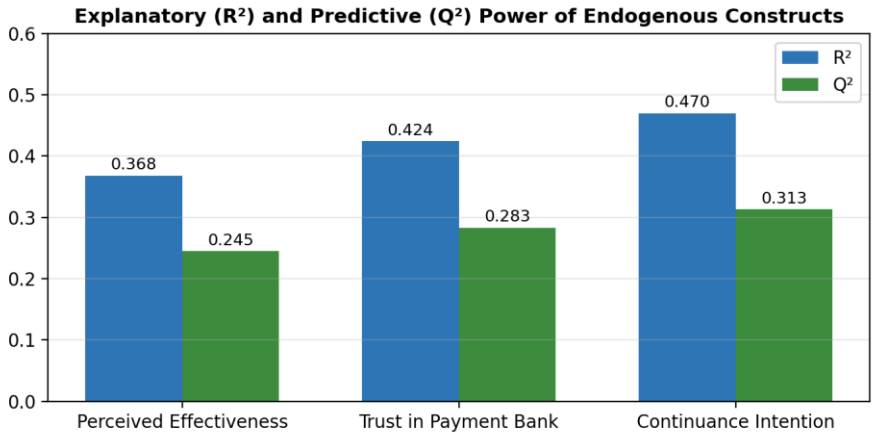
Volume 2, Issue 3, July - September 2026

E-ISSN: 3108-1452

NMIRJ Website: https://nrityanjaliresearchjournal.com/home/ Email: info@nrityanjaliresearchjournal.com

Continuance Intention (CUI)	0.470	0.466	0.313	Moderate–Substantial

Figure 7. R^2 and Q^2 of Endogenous Constructs



5. Discussion of Findings

The two analytical tracks converge on a single insight: AI fraud detection creates value at two levels that must be managed together. At the algorithmic level, the benchmark confirms the superiority of ensemble learning for imbalanced payment fraud—random forest and gradient boosting decisively outperform linear and single-tree baselines on AUC and F1 (Breiman, 2001; Chen & Guestrin, 2016), replicating the pattern documented across the financial fraud literature (Ngai et al., 2011). The near-zero false-positive rate of the random forest is commercially critical for payment banks, where wrongly blocked transactions directly damage the customer relationship.

At the customer level, all three system characteristics significantly shape perceived effectiveness, with responsiveness the strongest (H2: $\beta = 0.315$)—customers judge protection first by its immediacy, consistent with the centrality of speed in AI-mediated service (Huang & Rust, 2018). Trust formation follows the layered logic of trust theory: competence signals from accuracy (H4) and integrity signals from transparency (H5) are consolidated by the overall effectiveness judgement (H6: $\beta = 0.370$), empirically supporting the structural-assurance mechanism of McKnight et al. (2002) and the trust–technology duality of Gefen et al. (2003).

Finally, continuance intention is governed primarily by trust (H8: $\beta = 0.419$) and effectiveness (H7: $\beta = 0.290$), with only a small direct effect of responsiveness (H9). This ordering refines expectation–confirmation reasoning (Bhattacherjee, 2001): for security technologies that customers hope never to need, continuance rests less on feature perceptions than on the institutional trust those features cumulatively build—a boundary condition for TAM/UTAUT2 effects in protective, as opposed to productive, technologies (Davis, 1989; Venkatesh et al., 2012).

6. Implications of the Study

6.1 Theoretical Implications

Nrityanjali Multidisciplinary International Research Journal

Volume 2, Issue 3, July - September 2026

E-ISSN: 3108-1452

NMIRJ Website: https://nrityanjaliresearchjournal.com/home/ Email: info@nrityanjaliresearchjournal.com

The study contributes an integrated framework that connects classifier performance to customer behaviour through the mediating chain of perceived effectiveness and institutional trust, bridging the data-mining and information-systems literatures on fraud. It demonstrates that in protective technology contexts, trust—not usefulness—is the dominant proximal driver of continuance, and it validates the measurement of AI-specific system characteristics (accuracy, responsiveness, transparency) as distinct, psychometrically sound constructs.

6.2 Managerial Implications

Payment banks should draw three lessons. First, deploy ensemble models in production scoring and tune decision thresholds around the asymmetric costs of false positives and missed fraud; the benchmark shows the recall–precision frontier is materially wider for random forest and gradient boosting. Second, invest in the customer-facing layer of fraud protection: instant alerts, one-tap transaction confirmation, and plain-language explanations of why a transaction was flagged convert algorithmic capability into perceived effectiveness. Third, treat fraud protection as a trust-marketing asset—communicating detection performance, data-use policies, and grievance-resolution speed strengthens the trust that most directly retains customers.

6.3 Policy Implications

Regulators supervising payment banks should encourage explainability standards for automated fraud decisions, mandated disclosure of fraud-resolution turnaround times, and shared fraud-intelligence infrastructure, since customer trust in the weakest institution affects confidence in the digital payments ecosystem as a whole.

7. Limitations and Directions for Future Research

Several limitations qualify the findings. First, the demonstration datasets employed here are simulated; replication with production transaction logs and field survey data is essential before generalisation. Second, the cross-sectional survey design precludes causal claims; longitudinal designs could track how fraud incidents and system interventions dynamically reshape trust. Third, single-source self-reports remain vulnerable to residual method variance despite procedural and statistical remedies (Podsakoff et al., 2003). Fourth, the ML benchmark evaluated default-threshold performance on engineered features; future work should examine deep learning architectures, cost-sensitive and imbalance-aware training, graph-based mule-network detection, and drift-adaptive online learning. Finally, extending the behavioural model with moderators such as prior fraud victimisation, financial literacy, and privacy concern—or with hybrid SEM–ANN and fsQCA approaches—could reveal non-linear and configurational effects.

8. Conclusion

This study examined AI-based fraud detection in payment banks from both the algorithmic and the customer perspective. The machine learning benchmark establishes ensemble methods—particularly random forest (AUC = 0.932)—as the most effective classifiers for imbalanced payment-fraud data, while the validated PLS-SEM model shows that detection accuracy, real-time responsiveness, and transparency convert into continuance usage intention through perceived effectiveness and, above all, trust in the payment bank ($R^2 = 0.470$). Rigorous EFA, CFA, and bootstrap-based structural testing supported all nine hypotheses. The central message is that fraud detection technology succeeds when it is engineered for statistical performance and experienced by customers as fast, understandable, and trustworthy protection.

References

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. Psychological Bulletin, 103(3), 411–423. https://doi.org/10.1037/0033-2909.103.3.411

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. Journal of the Academy of Marketing Science, 16(1), 74–94. https://doi.org/10.1007/BF02723327

Bhattacherjee, A. (2001). Understanding information systems continuance: An expectation-confirmation model. MIS Quarterly, 25(3), 351–370. https://doi.org/10.2307/3250921

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. ACM Computing Surveys, 41(3), Article 15. https://doi.org/10.1145/1541880.1541882

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785–794). ACM. https://doi.org/10.1145/2939672.2939785

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

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research, 18(1), 39–50. https://doi.org/10.1177/002224378101800104

Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in online shopping: An integrated model. MIS Quarterly, 27(1), 51–90. https://doi.org/10.2307/30036519

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. European Business Review, 31(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. Journal of the Academy of Marketing Science, 43(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8

Huang, M.-H., & Rust, R. T. (2018). Artificial intelligence in service. Journal of Service Research, 21(2), 155–172. https://doi.org/10.1177/1094670517752459

Jain, P., Jain, V., & Jain, S. (2025). The influence of payment banks on India's financial system (with special reference to urban and rural areas of Indore district). Rabindra Bharti University Journal of Economics, 29(1), 179–187.

Jain, V. (2022a). Application of artificial intelligence in financial technology and its inclusion in Indian banking and financial system. International Journal of Advances in Engineering and Management, 4(7), 241–247.

Jain, V. (2022b). Inclusion of financial technology and artificial intelligence in management and development of human resource in India. International Journal of Research in Engineering and Science, 10(7), 585–589.

Jain, V. (2022c). Review of recent trends and stockholders in the Indian health care sector. Accent Journal of Economics, Ecology & Engineering, 7(Special Issue 05).

Jain, V. (2022d). Study of effect of green HRM practices on employee workplace in Indian organizations. Mukt Shabd Journal, 11(9).

Jain, V. (2022e). Study of employees' performance management system. Research Analysis and Evaluation, ISSN 0975-3486.

Jain, V. (2023). A study of consumer awareness of health insurance with reference to Malwa region. Ideal International E-Publication. ISBN 978-81-968040-2-2.

Jain, V., Jain, P., & Jain, P. (2023). A study on awareness of consumer for health insurance with special reference to Indore city. Mukt Shabd Journal, 12(7), 1784–1791.

Jain, V., Jain, P., & Jain, P. (2024a). An analysis of health insurance consumer knowledge with particular reference to Indore region. Madhya Bharti – Humanities and Social Sciences, 85(10).

Jain, V., Jain, P., & Jain, P. (2024b). Customer perception and factors affecting selection of health insurance policy in Indore region. Educational Administration: Theory and Practice, 30(5), 14140–14147.

Jain, V., Jain, P., & Jain, P. (2024c). India's health insurance market: Developments, challenges, and future directions. Journal of Management and Entrepreneurship, 18(2-II).

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and validating trust measures for e-commerce: An integrative typology. Information Systems Research, 13(3), 334–359. https://doi.org/10.1287/isre.13.3.334.81

Ngai, E. W. T., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. Decision Support Systems, 50(3), 559–569. https://doi.org/10.1016/j.dss.2010.08.006

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. Journal of Applied Psychology, 88(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. MIS Quarterly, 36(1), 157–178. https://doi.org/10.2307/41410412