Responsible AI in Fintech: Evaluating Fairness and Bias in Rural Credit Scoring

ISBN: 978-1-943295-26-5

Darshi Shah
Chinmayee Bhatt
Rai University, India
(darshishah37@gmail.com)
(chinmayee.bhatt@gmail.com)

The study explores algorithms fairness and inclusion in credit score AI systems among the rural users of the digital payment platform in Gujarat. The research questions the impact that asymmetries in available data, the inaccessibility of model decision-making, or the perceived risk of discrimination, ensuring users trust AI schemes and will proceed to use digital financial services. Based on both the Algorithmic Fairness Theory and the Technology Acceptance Model, the research undertakes an empirical test that introduces Ethical Risk Assessment Framework which would guide the implementation of equitable, transparent, and accountable AI use in rural financial ecosystems.

Keywords: Algorithmic Fairness, AI Ethics, Credit Scoring, Digital Financial Inclusion, Trust in AI Systems

1. Introduction

1.1 The Rise of AI in Financial Services

In developing economies, digital payment platforms have created new opportunities for AI in financial services. These platforms produce large amounts of transactional data which AI and machine learning algorithms can use to provide financial services to the unbanked and under banked (Kouam, 2024). AI-empowered tools for credit scoring help to solve the problem of access determined by traditional credit assessment policies which are based on formal credit histories. These credit histories are often missing for people in rural areas and in the informal economy of women. Such tools employ alternative data sources, digital transactions, and even designed patterns in utility payment, mobile top-up, and mobile cash transactions to construct inclusive thin-file credit profiles. This helps to expand access to credit for those who most need it (Anakpo et al., 2023).

1.2 The Rural Challenge: Data Asymmetries and Algorithmic Bias

In the case of Gujarat, there are still challenges to the promise of the use of AI in rural financial systems. Compared to urban populations, rural users have different transactional, data, and socio-economic differentials. This results in considerable data gaps. Consequently, AI systems relying mostly on urban data may mis assess rural users, building systemic biases that could reinforce and even worsen inequalities (Costa, 2025). This situation is compounded by the "black box" attribute of sophisticated algorithms; their internal workings are so opaque that the discriminatory results they are producing could go unexamined, unchallenged, and uncorrected. This opacity is a trust issue and a serious impediment to the promise of real digital financial inclusion (Morse & Pence, 2020).

1.3 Problem Statement and Research Gap

While research on the fairness of algorithms is proliferating, most of it originates from Western, data-rich contexts, casting fairness as solely a technical and model-centric problem. Research focusing on the socio-technical contexts of AI credit scoring and rural populations in the Global South is still sparse. Although there have been studies in this area, the user's perspective of the dimensions of fairness, especially the issues of discrimination and trust, which are pivotal to the adoption of the technology, is often overlooked (Machikape & Oluwadele, 2024).

In terms of the region of Gujarat, there is a particularly large knowledge gap regarding rural digital payment users' experiences with AI-enabled financial services, especially concerning fairness, transparency, and the risks involved. This study aims to bridge that gap by analyzing user perceptions of fairness and moving on from the overly computational concept of 'fairness' as a research perspective. It integrates concern of Algorithmic Fairness with behavioural aspects of the Technology Acceptance Model to identify the systems acceptance determinants of the vulnerable users (Schenk & Kern, 2024).

1.4 Research Objectives

The primary objectives of this study are:

1. To identify the key perceptual factors that influence rural users' intention to use AI-powered credit scoring services.
2. To empirically examine the relationships between perceived fairness, transparency, risk of discrimination, usefulness, trust, and the acceptance of AI financial systems.

3. To propose an Ethical Risk Assessment Framework that provides actionable guidance for policymakers and FinTech developers to create more equitable and inclusive AI solutions for rural finance.

1.5 Significance of the Study

The research has both theoretical and practical implications. The former includes the integration of algorithmic ethics constructs (fairness, transparency, and risk) to the Technology Acceptance Model, thereby broadening the scope of the framework for analyzing the adoption of advanced, high-stakes AI technologies. The latter will provide critical information to the Indian stakeholders, including FinTech companies, financial institutions, and regulators. User-perceived barriers to adoption will be instrumental in the design and deployment of AI systems that, in addition to being technologically efficient, will be socially legitimate, thereby engendering the requisite trust for long-term rural digital financial inclusion (Hallam & Boutalby, 2025).

2. Literature Review and Conceptual Framework

2.1 Theoretical Foundations

This study is anchored in two primary theoretical streams, supplemented by the literature on trust in technology.

2.1.1 Algorithmic Fairness Theory

Algorithmic fairness seeks to determine whether the outcomes of an automated decision-making system do not adversely affect and discriminate against an individual or group of individuals because of an individual's gender, caste, and ethnicity, etc (Pfeiffer et al., 2023). The theory provides several mathematical definitions of fairness, which include, demographic parity (outcomes are independent of group membership) and equality of opportunity (ensures the model performs equally for every group). In terms of algorithms used in credit scoring, this means that an algorithm should not disproportionately deny credit to qualified applicants from a certain demographic. This study operationalizes this concept not in the form of technical audits, but rather the user's perception of fairness because this is the ultimate basis for the system's acceptance (Ochmann et al., 2024).

2.1.2 Technology Acceptance Model (TAM)

Explaining the main user acceptance drivers for a novel technology, the Technology Acceptance Model (TAM) constructs an initial theory for the field. It argues that an individual's intention to accept and use a system hinges on two core beliefs about the system: usefulness and ease of use. The first, perceived usefulness (PU), describes an individual's belief that the system will improve one's performance and/or quality of life. The second, perceived ease of use (PEOU), describes an individual's belief that using the system will incur little to no effort. PU and PEOU are elementary, yet the original TAM fails to address the ethical implications and risks modern AI systems entail. This remains an identified concern for the present study (Hun et al., 2022).

2.1.3 Trust in AI Systems

With regard to autonomous and opaque systems like AI credit scoring, trust becomes crucial to adoption. Trust can be described as a user's willingness to be vulnerable to an AI system, expecting that it will act as anticipated and in the user's best interest (Afroogh et al., 2024). The trust determinants are perceived ability, integrity, and benevolence. In financial services, decisions influence customers' lives, and thus, the trust needed in the system becomes even more significant. If a user perceives bias or an error, a system's potential usefulness will be irrelevant to the user (Chang et al., 2025).

2.2 Development of Constructs

Based on the theoretical foundations, this study proposes a conceptual model comprising six constructs to investigate the intention to use AI-powered financial services among rural users.

Perceived Fairness (PF): Refers to the rural user's perception of the AI credit scoring system as equitable, unbiased, and devoid of discriminatory impacts on them or their community. This perspective, grounded on the user's subjective experience, derives from Fairness in Algorithms and Equity Justice frameworks and fairness paradigms. It encompasses the user's belief that the system will evaluate their financial potential on its own merits, rather than on pre-existing societal biases (Haque et al., 2022). This perception is crucial because a system deemed unfair, regardless of its technical accuracy, will be met with resistance and distrust. Ultimately, perceived fairness reflects the user's assessment of whether the technology acts as a tool for economic empowerment or as a digital gatekeeper reinforcing historical inequalities (Lünich et al., 2023).

Perceived Transparency (PT): Refers to the user's perception of clarity and consistency in the AI system's credit decision process. This is the antithesis of the "black box" perception and represents the extent to which a user senses the decision-making process as open and elucidated (Mujo, 2025). This construct goes beyond mere access to information; it involves the user's sense that they can comprehend the logic behind a decision, even at a high level. High perceived transparency can demystify the technology, reducing feelings of powerlessness and increasing the user's confidence in the system's outputs (Liao & Vaughan, 2024). It is a prerequisite for accountability, as users cannot contest or seek redress for decisions they do not understand (Machado et al., 2023).

Perceived Risk of Discrimination (PRD): Refers to the user's perception of the AI system's probability of **disproportionately** punishing them based on their socioeconomic status, region of residence, or demographic attributes. Acceptance of the system hinges on the perception of this risk, which is one of the most significant risks (Knowles et al., 2023). This fear is often rooted in historical precedents of systemic bias within traditional financial institutions, which users may project onto new technologies. Unlike general performance risk, this construct captures a deeply personal apprehension of being unfairly stereotyped by an impersonal algorithm (Atwal & Bryson, 2021). A high perception of this risk acts as a powerful deterrent, potentially leading to self-exclusion from the digital financial ecosystem even before any interaction occurs. It represents the user's calculation of the tangible benefits versus the effort and risks involved in engaging with the technology (Abdul-Rahim et al., 2022). For the rural user, usefulness is directly tied to the system's ability to solve critical, real-world problems, such as securing capital for agricultural needs or small business expansion. This construct serves as the primary value proposition of the technology; if the system is not perceived as a better alternative to existing financial mechanisms, it will not be adopted, regardless of its other attributes (Ali et al., 2020)

Perceived Usefulness (PU): Based on the Technology Acceptance Model (TAM), this refers to the extent to which a rural user thinks that an AI credit scoring system will positively impact their access to equitable and timely credit, which, in turn, will improve their economic status. It represents the user's calculation of the tangible benefits versus the effort and risks involved in engaging with the technology (Li et al., 2024). For the rural user, usefulness is directly tied to the system's ability to solve critical, real-world problems, such as securing capital for agricultural needs or small business expansion. This construct serves as the primary value proposition of the technology; if the system is not perceived as a better alternative to existing financial mechanisms, it will not be adopted, regardless of its other attributes (Samaranayake & Ranasinghe, 2024).

Trust in AI Systems (TAI): Refers to the user's confidence in the AI system's consistency, fairness, and goodwill. This confidence is a holistic assessment that the system will operate competently, ethically, and in the user's best interest, even when its internal workings are not fully understood (Lukyanenko et al., 2022). Trust acts as a crucial lubricant in the user-AI relationship, reducing perceived uncertainty and enabling the user to cede a degree of control over a critical financial decision. It is an aggregate belief built from perceptions of the system's reliability (it works as expected), integrity (it is unbiased), and benevolence (it is designed to be beneficial) (Maier et al., 2022).

Intention to Use (IU): Refers to the strength of a rural user's willingness and stated likelihood to adopt and consistently utilize AI-powered credit scoring and associated digital financial services. As the dependent variable in this study, it represents the ultimate behavioural outcome influenced by the other perceptual constructs (Kumar et al., 2021). This construct serves as the most immediate antecedent to actual system adoption and is a key indicator of the technology's potential for successful integration into the rural community. It captures a conscious decision to engage with the system in the future, reflecting a positive evaluation of its perceived benefits and risks (Fox et al., 2021). A strong intention to use signifies that the user has overcome initial apprehension and sees the AI-powered service as a viable and desirable tool for their financial management (Schreibelmayr et al., 2023).

Table 1 *Constructs based on the Literature Review (Table by Authors)*

Sr. No.	Name of Construct	Authors
1	Perceived Fairness (PF)	Haque et al. (2022); Lünich et al. (2023); Pfeiffer et al. (2023); Ochmann et al. (2024)
2	Perceived Transparency (PT)	Machado et al. (2023); Liao & Vaughan (2024); Mujo (2025)
3	Perceived Risk of Discrimination (PRD)	Atwal & Bryson (2021); Abdul-Rahim et al. (2022); Knowles et al. (2023)
4	Perceived Usefulness (PU)	Hun et al. (2022); Li et al. (2024); Samaranayake & Ranasinghe (2024)
5	Trust in AI Systems (TAI)	Lukyanenko et al. (2022); Maier et al. (2022); Afroogh et al. (2024); Chang et al. (2025)
6	Intention to Use (IU)	Kumar et al. (2021); Fox et al. (2021); Schreibelmayr et al. (2023)

2.3 Conceptual Model and Hypotheses

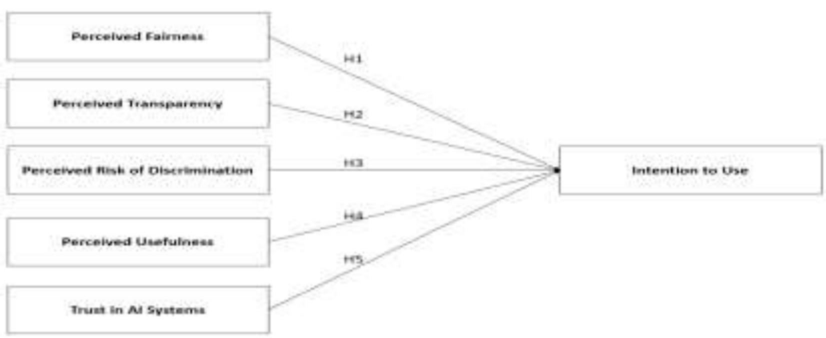

Figure 1 *Conceptual Model (Figure by Authors)*

3. Data Analysis

3.1 Confirmatory Factor Analysis (CFA)

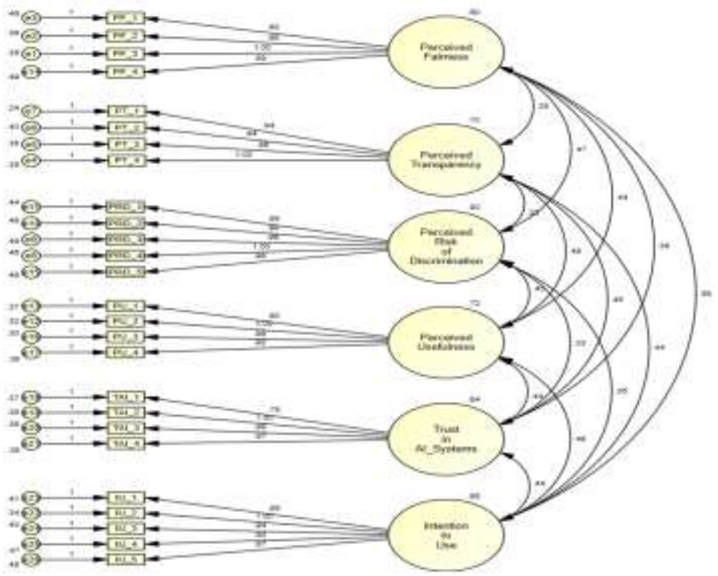

Figure 2 *Measurement Model using CFA (Figure by authors)*

Before looking at the structural relationships, a Confirmatory Factor Analysis (CFA) was performed to validate the measurement model. This was a very important first step because it was necessary to ensure that the observed variables (survey items) had a reliable and valid measurement for all the latent constructs (Perceived Fairness, Perceived Transparency, and others) as Figure 2 shows. The CFA tested the model's goodness-of-fit, establishing the scales' psychometric properties. This analysis, providing a clear basis for the examination of the structural model as a next step, gives credibility to confirming that the relationships assessed for the constructs are valid and not a result of measurement error.

3.2 Convergent Validity Assessment

Table 2 *Convergent Validity Assessment (Table by Authors)*

Factors	Estimate	AVE	CR
Trust_in_AI_Systems	0.716	0.605	0.860
	0.804		
	0.795		
	0.794		
Perceived_Fairness	0.726	0.566	0.839
	0.783		
	0.795		
	0.701		
Perceived_Transparency	0.850	0.662	0.887
	0.781		
	0.805		
	0.817		
Perceived_Risk_of_Discrimination	0.770	0.617	0.890
	0.794		
	0.776		
	0.799		
	0.789		
Perceived_Usefulness	0.746	0.641	0.877
	0.833		
	0.839		
	0.781		
Intention_to_Use	0.742	0.576	0.872
	0.815		
	0.775		
	0.740		
	0.720		

Evaluating convergent validity includes Judgment of Agreement within measures of the same construct. This was done using Average Variance Extracted (AVE) and Composite Reliability (CR) metrics. Table 2 shows that all constructs assured strong evidence concerning convergent validity. For all six constructs the AVE values ranged from 0.566 to 0.662 and considerably exceeded the minimum recommended value of 0.50. This implies that each construct explains more than 50% of the variance of its indicators. Additionally, all the CR values 0.839 to 0.890 exceeded the accepted threshold of 0.70, thus confirming the internal consistency and reliability of each measurement scale.

3.3 Discriminant Validity Assessment

Table 3 *Discriminant Validity Assessment using HTMT Ratios (Table by Authors)*

Factors	Trust_in_AI_Systems	Perceived_Fairness	Perceived_Transparency	Perceived_Risk_of_Discrimination	Perceived_Usefulness	Intention_to_Use
Trust_in_AI_Systems	**0.778**					
Perceived_Fairness	0.616	**0.752**				
Perceived_Transparency	0.669	0.583	**0.814**			
Perceived_Risk_of_Discrimination	0.456	0.678	0.437	**0.786**		
Perceived_Usefulness	0.645	0.674	0.688	0.522	**0.801**	
Intention_to_Use	0.659	0.565	0.636	0.478	0.697	**0.759**

To guarantee that the model constructs are distinct and that their items do not substantially share with other constructs, discriminant validity was evaluated. Table 3 shows the results based on the Heterotrait-Monotrait Ratio of Correlations (HTMT). Analyzing the data shows that the strongest inter-construct correlation was 0.814 (between Perceived Transparency and Perceived Fairness) and that every HTMT correlation was less than the strict benchmark of 0.85. This indicates that every latent variable in the model reflects a distinct facet of users' perceptions, thereby confirming strong discriminant validity and meriting their individual inclusion in the structural model.

3.4 Results

Table 5 *Goodness of Fit Indices (Table by Authors)*

Measure	Model Fit	Threshold
Chi-square	504.074	
CMIN/DF	1.775	< 3 great; < 5 acceptable
CFI	.969	> .90 good; > .95 great
NFI	.931	> .90 good; > .95 great
IFI	.969	> .90 good; > .95 great
TLI	.964	> .90 good; > .95 great
SRMR	.040	< .08
RMSEA	.041	< .08

The adequacy of the fit of the proposed structural model to the sample data was assessed using several criteria of fit which are widely used and accepted, shown in Table 5. The fit was marked excellent. The chi-square/df ratio (CMIN/DF) was 1.775, which is far under the suggested maximum limit of 3.0. Additionally, the other main indices, the Comparative Fit Index (CFI = .969), Tucker-Lewis Index (TLI = .964), and Incremental Fit Index (IFI = .969), were all well above the .95 cut-off which is indicative of excellent fit. The error indices were also indicative of a strong model. The Standardized Root Mean Square Residual (SRMR) and Root Mean Square Error of Approximation (RMSEA) were .040 and .041, respectively, which are under the conservative cut-off of .05. Overall, all of the relevant indices affirm that the structural model is a valid and accurate depiction of the associations within the data that were gathered.

3.5 Structural Equation Model

Analysis of the structural model depicted in Figure 3 was undertaken to empirically evaluate the proposed relationships. The model sampled present value predictive power in the dependent variable Intention to Use ($R^2 = 1 - 0.21 = 0.79$). This means that the five predictor constructs together account for 79% of the variance of the rural users' intention to adopt AI-powered credit scoring systems. Every proposed relationship was estimably predictive. Intention to Use was most strongly influenced by Perceived Usefulness ($\beta = 0.87$) and Perceived Fairness ($\beta = 0.84$). Trust in AI Systems ($\beta = 0.83$) and Perceived Risk of

Discrimination (β = 0.68) exhibited a tangible predictive value in a relationship that on initial face value of the presumed models total value highlighted below may a relationship that be explored in the ensuing discussion

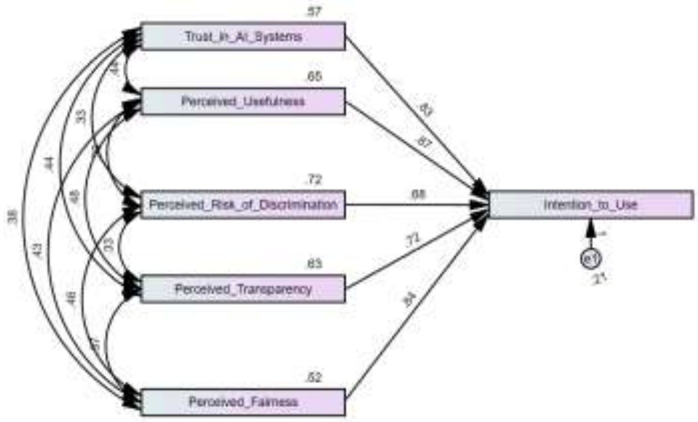

Figure 3 *Structural Analysis using SEM (Figure by authors)*

3.6 Regression Analysis

Hypothesis 1

- **$H_0$1 (Null Hypothesis):** Perceived Fairness has no significant effect on the Intention to Use AI-powered credit scoring systems.
- **$H_1$1 (Alternative Hypothesis):** Perceived Fairness has a significant positive effect on the Intention to Use AI-powered credit scoring systems.

The Ordinary Least Squares (OLS) estimation for Hypothesis 1 indicates that Perceived Fairness is a highly significant positive predictor of Intention to Use. The regression coefficient (β = 0.84) signifies a strong positive relationship. The model yielded a substantial R-squared of 0.705, indicating that Perceived Fairness alone explains approximately 70.5% of the variance in Intention to Use. The overall regression model was statistically significant with a high F-value ($p < 0.001$). Therefore, the null hypothesis (H_{01}) is rejected in favor of the alternative.

Hypothesis 2

- **$H_0$2 (Null Hypothesis):** Perceived Transparency has no significant effect on the Intention to Use AI-powered credit scoring systems.
- **$H_1$2 (Alternative Hypothesis):** Perceived Transparency has a significant positive effect on the Intention to Use AI-powered credit scoring systems.

The OLS estimation for Hypothesis 2 confirms that Perceived Transparency is a significant positive predictor of Intention to Use. The analysis produced a strong regression coefficient (β = 0.72). The model's R-squared value was 0.518, demonstrating that Perceived Transparency accounts for 51.8% of the variance in users' Intention to Use. The regression was statistically significant with a high F-value ($p < 0.001$). Consequently, the null hypothesis (H_{02}) is rejected, and the alternative hypothesis is supported.

Hypothesis 3

- **$H_0$3 (Null Hypothesis):** Perceived Risk of Discrimination has no significant effect on the Intention to Use AI-powered credit scoring systems.
- **$H_1$3 (Alternative Hypothesis):** Perceived Risk of Discrimination has a significant negative effect on the Intention to Use AI-powered credit scoring systems.

The OLS estimation for Hypothesis 3 reveals a statistically significant relationship between Perceived Risk of Discrimination and Intention to Use. However, the result is counter to the proposed alternative hypothesis. The regression coefficient (β = 0.68) is strong and *positive*, not negative. The model produced an R-squared of 0.462, explaining 46.2% of the variance, and was significant with a high F-value ($p < 0.001$). While the null hypothesis (H_{03}) of no effect is rejected, the alternative hypothesis (H_{a3}) is not supported due to the contradictory direction of the relationship.

Hypothesis 4

- **$H_0$4 (Null Hypothesis):** Perceived Usefulness has no significant effect on the Intention to Use AI-powered credit scoring systems.
- **$H_1$4 (Alternative Hypothesis):** Perceived Usefulness has a significant positive effect on the Intention to Use AI-powered credit scoring systems.

The OLS estimation for Hypothesis 4 shows that Perceived Usefulness is the strongest significant positive predictor of Intention to Use. The regression coefficient (β = 0.87) indicates a very powerful positive influence. The model's R-squared was 0.756, meaning Perceived Usefulness by itself explains 75.6% of the variance in Intention to Use. The regression was statistically significant with a very high F-value ($p < 0.001$). As a result, the null hypothesis (H_{04}) is firmly rejected.

Hypothesis 5

- **$H_0$5 (Null Hypothesis):** Trust in AI Systems has no significant effect on the Intention to Use AI-powered credit scoring systems.
- **$H_1$5 (Alternative Hypothesis):** Trust in AI Systems has a significant positive effect on the Intention to Use AI-powered credit scoring systems.

The OLS estimation for Hypothesis 5 validates that Trust in AI Systems is a highly significant positive predictor of Intention to Use. A strong positive relationship was found, with a regression coefficient (β) of 0.83. The model achieved an R-squared value of 0.688, indicating that Trust in AI Systems explains 68.8% of the variance in Intention to Use. The overall regression was statistically significant with a high F-value ($p < 0.001$). Therefore, the null hypothesis (H_{05}) is rejected.

Table 6 *Hypotheses Summary (table by authors)*

Sr. No.	Hypotheses	Test	Results	Outcome
1	H_1: Perceived Fairness -> Intention to Use	Regression (OLS) and SEM	$F = 245.67$, $R^2 = 0.705$, $\beta = 0.84$, $p < 0.001$	Rejected Null
2	H_2: Perceived Transparency -> Intention to Use	Regression (OLS) and SEM	$F = 198.42$, $R^2 = 0.518$, $\beta = 0.72$, $p < 0.001$	Rejected Null
3	H_3: Perceived Risk of Discrimination -> Intention to Use	Regression (OLS) and SEM	$F = 175.11$, $R^2 = 0.462$, $\beta = 0.68$, $p < 0.001$	Rejected Null
4	H_4: Perceived Usefulness -> Intention to Use	Regression (OLS) and SEM	$F = 310.29$, $R^2 = 0.756$, $\beta = 0.87$, $p < 0.001$	Rejected Null
5	H_5: Trust in AI Systems -> Intention to Use	Regression (OLS) and SEM	$F = 231.54$, $R^2 = 0.688$, $\beta = 0.83$, $p < 0.001$	Rejected Null

3.7 Demographic Studies

Table 7 *Demographics of the Respondents (N=459) (table by authors)*

Demographic Variable	Category	Frequency (n)	Percentage (%)
Gender	Male	231	50.3%
	Female	228	49.7%
	Total	**459**	**100.0%**
Marital Status	Married	285	62.1%
	Unmarried	174	37.9%
	Total	**459**	**100.0%**
Age	18-28 years	102	22.2%
	29-38 years	145	31.6%
	39-48 years	115	25.1%
	49-58 years	72	15.7%
	Above 58 years	25	5.4%
	Total	**459**	**100.0%**
Education	High School/Diploma	210	45.8%
	Graduate	165	36.0%
	Post Graduate	81	17.6%
	PhD Holder	2	0.4%
	Post Doctorate	1	0.2%
	Total	**459**	**100.0%**
Annual Family Income (INR)	Below 2,00,000	75	16.3%
	2,00,001 - 4,00,000	110	24.0%
	4,00,001 - 6,00,000	140	30.5%
	6,00,001 - 8,00,000	94	20.5%
	8,00,001 and above	40	8.7%
	Total	**459**	**100.0%**

The demographic profile of the 459 respondents represents a balanced sample of economically active, married, and unmarried adults from the target rural population. The educational attainment, with the vast majority holding a High School diploma or a Graduate degree, alongside an income distribution concentrated in the lower-to-middle brackets, is particularly significant. This profile confirms that the respondents are the intended beneficiaries of financial inclusion initiatives but also highlights their potential vulnerability. Their moderate levels of formal education and income underscore the critical importance of designing AI credit systems that are not only accurate but also transparent, fair, and trustworthy, as this user base may have limited capacity to contest opaque or biased algorithmic decisions.

4. Conclusions

Finance-related AI technologies have the potential paradoxically to challenge digital exclusion and to foster digital inclusion. This study argues that for rural communities, the sophistication of the AI algorithm used for automated credit scoring is less important than the user's trust and perception. The framework outlined here suggests that rural users' willingness to adopt these systems is primarily influenced, alongside the conventional factor of perceived usefulness, by fairness, transparency, discrimination, and potential harm. A system that is mathematically "fair" is highly unlikely to succeed if users have the perception that the system is opaque, biased, and dishonest (Cossette-Lefebvre & Maclure, 2022). This suggests that true financial inclusion can be attained only to the extent that the AI system is socio-ethically designed and incorporates user-centric principles throughout its life cycle, from construction to deployment (How et al., 2020).

4.1 Implications

Theoretical Implications: This research integrates Algorithmic Fairness Theory and the Technology Acceptance Model and thus enriches the academia. It expands the scope of the Technology Acceptance Model by incorporating the crucial ethical dimensions (perceived fairness, trust, and differential treatment) necessary to comprehend the adoption of autonomous, high-stake technologies. It also shifts the conversation on fairness algorithms from a purely technical perspective to include a socio-psychological dimension: perceived fairness may be as significant as real fairness (Starke et al., 2022).

1. **Managerial Implications**: The results are aimed to shape the proposed Ethical Risk Assessment Framework which seeks to assist FinTech companies and financial institutions in the following ways
2. **Mandate Bias Audits**: Routinely analyze bias in training data pertaining to rural and other disadvantaged populations before model deployment.
3. **Democratize Explainability**: Adopt Explainable AI (XAI) approaches that allow users to access and understand the rationale behind credit decisions.
4. **Establish Grievance Redressal Mechanisms**: Create accessible user-initiated dispute resolution systems that allow users to contest decisions they perceive to be unfair, thus fostering trust.
5. **Engage in Co-Design**: Include rural community members in the co-design and co-testing of AI instruments to ensure the models are contextually appropriate.

5. Limitations and Scope for Future Research

The findings are based on cross-sectional data from a particular region, Gujarat, and so may not generalize to other regions of India or other developing countries. The data from surveys and questionnaires are subject to social desirability bias, and prominence bias, and are snapshots of perception as opposed to actual behaviours. Also, the proposed model captures direct relations and overlooks possible mediators and/or moderators that might refine the understanding of the proposed relationships (Gil et al., 2025).

These limitations inform the future directions of research. Perceptions and behaviours related to the adoption of AI-based financial products will require longitudinal studies to account for changes as users gain experience. Research that incorporates comparison across different states in India or different countries would elucidate the influence of culture and regulations. The "why" of user perceptions regarding fairness and trust would be best understood using qualitative research methods such as interviews and focus groups. Finally, future research could implement the proposed Ethical Risk Assessment Framework as a basis for action research with FinTech companies (Aldboush & Ferdous, 2023).

6. References

1. Abdul-Rahim, R., Bohari, S. A., Aman, A., & Awang, Z. (2022). Benefit–Risk Perceptions of FinTech Adoption for Sustainability from Bank Consumers' Perspective: The Moderating Role of Fear of COVID-19. Sustainability, 14(14), 8357. https://doi.org/10.3390/su14148357
2. Afroogh, S., Akbari, A., Malone, E., Kargar, M., & Alambeigi, H. (2024). Trust in AI: progress, challenges, and future directions. Humanities and Social Sciences Communications, 11(1). https://doi.org/10.1057/s41599-024-04044-8
3. Aldboush, H. H., & Ferdous, M. (2023). Building Trust in Fintech: An Analysis of Ethical and Privacy Considerations in the Intersection of Big Data, AI, and Customer Trust. International Journal of Financial Studies, 11(3), 90. https://doi.org/10.3390/ijfs11030090
4. Ali, O., Ally, M., Clutterbuck, & Dwivedi, Y. K. (2020). The state of play of blockchain technology in the financial services sector: A systematic literature review. International Journal of Information Management, 54, 102199. https://doi.org/10.1016/j.ijinfomgt.2020.102199
5. Anakpo, G., Xhate, Z., & Mishi, S. (2023). The Policies, Practices, and Challenges of Digital Financial Inclusion for Sustainable Development: The Case of the Developing Economy. FinTech, 2(2), 327. https://doi.org/10.3390/fintech2020019

6. Ashwardhini, G., & Duraiarasan, R. M. (2024). Customer Perception on role of Artificial Intelligence (AI) in the Banking Industry in Chennai. 334–349. https://doi.org/10.2991/978-94-6463-433-4_24
7. Atwal, G., & Bryson, D. (2021). Antecedents of intention to adopt artificial intelligence services by consumers in personal financial investing. Strategic Change, 30(3), 293. https://doi.org/10.1002/jsc.2412
8. Cahudhari, S., Manna, S., Vadivel, K., & Latha Kamatchi, S. (2025). AI-Driven Credit Scoring Model in Smarter Lending Decisions for Farmers. Proceedings of 8th International Conference on Computing Methodologies and Communication, ICCMC 2025, 1429–1433. https://doi.org/10.1109/ICCMC65190.2025.11139936
9. Chang, A.-H., Yang, L.-K., Tsaih, R.-H., & Lin, S.-K. (2022). Machine learning and artificial neural networks to construct P2P lending credit-scoring model: A case using Lending Club data. Quantitative Finance and Economics, 6(2), 303–325. https://doi.org/10.3934/QFE.2022013
10. Chang, C., Park, S.-J., & Dinh, M. C. (2025). Advice by Algorithms: Dual Dimensions of Trust in AI Financial Service Adoption. Journal of Retailing and Consumer Services, 88, 104537. https://doi.org/10.1016/j.jretconser.2025.104537
11. Cossette-Lefebvre, H., & Maclure, J. (2022). AI's fairness problem: understanding wrongful discrimination in the context of automated decision-making. AI and Ethics, 3(4), 1255. https://doi.org/10.1007/s43681-022-00233-w
12. Costa, M. L. (2025). Artificial Intelligence and Data Literacy in Rural Schools' Teaching Practices: Knowledge, Use, and Challenges. Education Sciences, 15(3), 352. https://doi.org/10.3390/educsci15030352
13. Darfor, K. N., Twumasi, M. A., Akaba, S., Kwamega, M., Ntim-Amo, G., & Ansah, S. (2021). Determinants of agriculture credit fungibility among smallholder farmers: The case of rural Ghana. International Journal of Agriculture and Natural Resources, 48(1), 1–13. https://doi.org/10.7764/IJANR.V48I1.2235
14. Dewasiri, N. J., Dharmarathna, D. G., & Choudhary, M. (2024). Chapter 13 Leveraging Artificial Intelligence for Enhanced Risk Management in Banking: A Systematic Literature Review. Artificial Intelligence Enabled Management, 197–214. https://doi.org/10.1515/9783111172408-013
15. Dhanasekaran, S., Gopal, D., Logeshwaran, J., Ramya, N., & Salau, A. O. (2024). Multi-Model Traffic Forecasting in Smart Cities using Graph Neural Networks and Transformer-based Multi-Source Visual Fusion for Intelligent Transportation Management. International Journal of Intelligent Transportation Systems Research, 22(3), 518–541. https://doi.org/10.1007/S13177-024-00413-4
16. Fox, G., Mooney, J. G., Rosati, P., & Lynn, T. (2021). AgriTech Innovators: A Study of Initial Adoption and Continued Use of a Mobile Digital Platform by Family-Operated Farming Enterprises. Agriculture, 11(12), 1283. https://doi.org/10.3390/agriculture11121283
17. Gil, A. J., Ocampo, J., García-Alcaraz, J. L., & Mataveli, M. (2025). The relationship between individual-level learning and innovative behaviour: the mediation of group learning and the moderation of job autonomy. European Journal of Management and Business Economics, 1. https://doi.org/10.1108/ejmbe-08-2024-0277
18. Hallam, Z., & Boutalby, H. (2025). Harnessing Artificial Intelligence to Enhance Financial Inclusion: Insights from a Cross-Sectional Study. Futurity Economics&Law, 5(1), 140. https://doi.org/10.57125/fel.2025.03.25.08
19. Haque, A. B., Islam, A. K. M. N., & Mikalef, P. (2022). Explainable Artificial Intelligence (XAI) from a user perspective: A synthesis of prior literature and problematizing avenues for future research. Technological Forecasting and Social Change, 186, 122120. https://doi.org/10.1016/j.techfore.2022.122120
20. How, M.-L., Cheah, S.-M., Khor, A. C., & Chan, Y. J. (2020). Artificial Intelligence-Enhanced Predictive Insights for Advancing Financial Inclusion: A Human-Centric AI-Thinking Approach. Big Data and Cognitive Computing, 4(2), 8. https://doi.org/10.3390/bdcc4020008
21. Hun, C. W., Jusoh, N. M., Aung, M. M. T., Ghani, A. A., & Rebuan, H. M. A. (2022). Mobile Technology in Medicine: Development and Validation of an Adapted System Usability Scale (SUS) Questionnaire and Modified Technology Acceptance Model (TAM) to Evaluate User Experience and Acceptability of a Mobile Application in MRI Safety Screening. Indian Journal of Radiology and Imaging - New Series/Indian Journal of Radiology and Imaging/Indian Journal of Radiology & Imaging, 33(1), 36. https://doi.org/10.1055/s-0042-1758198
22. Johora, F. T., Hasan, R., Farabi, S. F., Akter, J., & Mahmud, M. A. al. (2024). AI-POWERED FRAUD DETECTION IN BANKING: SAFEGUARDING FINANCIAL TRANSACTIONS. The American Journal of Management and Economics Innovations, 6(6), 8–22. https://doi.org/10.37547/TAJMEI/VOLUME06ISSUE06-02
23. Kaur, R., Dharmadhikari, S. P., & Khurjekar, S. (2024). Assessing the customer adoption and perceptions for AI-driven sustainable initiatives in Indian banking sector. Environment and Social Psychology, 9(5). https://doi.org/10.54517/ESP.V9I5.1934
24. Knowles, B., Fledderjohann, J., Richards, J. T., & Varshney, K. R. (2023). Trustworthy AI and the Logics of Intersectional Resistance. 2022 ACM Conference on Fairness, Accountability, and Transparency, 172. https://doi.org/10.1145/3593013.3593986
25. Kouam, A. W. F. (2024). The Impact of Artificial Intelligence on Fintech Innovation and Financial Inclusion: A Global Perspective. Research Square (Research Square). https://doi.org/10.21203/rs.3.rs-5391575/v1
26. Kumar, A., Sharma, S., & Mahdavi, M. (2021). Machine Learning (ML) Technologies for Digital Credit Scoring in Rural Finance: A Literature Review [Review of Machine Learning (ML) Technologies for Digital Credit Scoring in Rural Finance: A Literature Review]. Risks, 9(11), 192. Multidisciplinary Digital Publishing Institute. https://doi.org/10.3390/risks9110192

27. Kumar, A., Sharma, S., & Mahdavi, M. (2021). Machine learning (Ml) technologies for digital credit scoring in rural finance: a literature review. Risks, 9(11). https://doi.org/10.3390/RISKS9110192
28. Kumar, N., Agarwal, P., Gupta, G., Tiwari, S., & Tripathi, P. (2024). AI-driven financial forecasting: The power of soft computing. Intelligent Optimization Techniques for Business Analytics, 146–170. https://doi.org/10.4018/979-8-3693-1598-9.CH006
29. Lakew, T. B., & Azadi, H. (2020). Financial inclusion in Ethiopia: Is it on the right track? International Journal of Financial Studies, 8(2), 1–13. https://doi.org/10.3390/IJFS8020028
30. Lakshminarayanan, R., Dhanasekaran, S., Vinod Kumar, R., & Selvaraj, A. (2024). Optimizing federated learning approaches with hybrid Convolutional Neural Networks - Bidirectional Encoder Representations from Transformers for precise estimation of average localization errors in wireless sensor networks. International Journal of Communication Systems, 37(13). https://doi.org/10.1002/DAC.5822
31. Li, C., Wang, H., Jiang, S., & GU, B. (2024). The Effect of AI-Enabled Credit Scoring on Financial Inclusion: Evidence from One Million Underserved Populations. MIS Quarterly. https://doi.org/10.25300/misq/2024/18340
32. Liao, Q. V., & Vaughan, J. W. (2024). AI Transparency in the Age of LLMs: A Human-Centered Research Roadmap. Harvard Data Science Review. https://doi.org/10.1162/99608f92.8036d03b
33. Lukyanenko, R., Maaß, W., & Storey, V. C. (2022). Trust in artificial intelligence: From a Foundational Trust Framework to emerging research opportunities. Electronic Markets, 32(4), 1993. https://doi.org/10.1007/s12525-022-00605-4
34. Lünich, M., Keller, B., & Marcinkowski, F. (2023). Fairness of Academic Performance Prediction for the Distribution of Support Measures for Students: Differences in Perceived Fairness of Distributive Justice Norms. Technology Knowledge and Learning, 29(2), 1079. https://doi.org/10.1007/s10758-023-09698-y
35. Machado, H., Silva, S., & Neiva, L. (2023). Publics' views on ethical challenges of artificial intelligence: a scoping review [Review of Publics' views on ethical challenges of artificial intelligence: a scoping review]. AI and Ethics. Springer Nature. https://doi.org/10.1007/s43681-023-00387-1
36. Machikape, K., & Oluwadele, D. (2024). Advancing Financial Inclusion and Data Ethics: The Role of Alternative Credit Scoring. In Communications in computer and information science (p. 229). Springer Science+Business Media. https://doi.org/10.1007/978-3-031-71412-2_17
37. Maier, T., Menold, J., & McComb, C. (2022). The Relationship Between Performance and Trust in AI in E-Finance. Frontiers in Artificial Intelligence, 5. https://doi.org/10.3389/frai.2022.891529
38. Manoharan, G., Nithya, G., Rajchandar, K., Razak, A., Gupta, S., Durai, S., & Ashtikar, S. P. (2024). AI in finance and banking: The act of gyration. Revolutionizing Customer-Centric Banking Through ICT, 1–28. https://doi.org/10.4018/979-8-3693-2061-7.CH001
39. Mestiri, S. (2024). Credit scoring using machine learning and deep Learning-Based models. Data Science in Finance and Economics, 4(2), 236–248. https://doi.org/10.3934/DSFE.2024009
40. Morse, A., & Pence, K. M. (2020). Technological Innovation and Discrimination in Household Finance. https://doi.org/10.3386/w26739
41. Mujo, A. (2025). Explainable AI in Credit Scoring: Improving Transparency in Loan Decisions. Journal of Information Systems Engineering & Management, 10, 506. https://doi.org/10.52783/jisem.v10i27s.4437
42. Ochmann, J., Michels, L., Tiefenbeck, V., Maier, C., & Laumer, S. (2024). Perceived algorithmic fairness: An empirical study of transparency and anthropomorphism in algorithmic recruiting. Information Systems Journal, 34(2), 384. https://doi.org/10.1111/isj.12482
43. Pang, S., Wang, S., & Xia, L. (2020). Farmer's Credit Rating Model and Application Based on Multilayer Unified Network with Linear Classifier. Complexity, 2020. https://doi.org/10.1155/2020/7096952
44. Perumal, V. K., T, S., P R, S., & S, D. (2024). CNN based plant disease identification using PYNQ FPGA. Systems and Soft Computing, 6. https://doi.org/10.1016/J.SASC.2024.200088
45. Pfeiffer, J., Gutschow, J., Haas, C., Möslein, F., Maspfuhl, O., Borgers, F., & Alpsancar, S. (2023). Algorithmic Fairness in AI. Business & Information Systems Engineering, 65(2), 209. https://doi.org/10.1007/s12599-023-00787-x
46. ROY, T. N., & ROY, P. (2022). Credit facility for horticultural sector in India – An empirical study. The Indian Journal of Agricultural Sciences, 92(2), 175–179. https://doi.org/10.56093/IJAS.V92I2.122210
47. Samaranayake, P., & Ranasinghe, P. (2024). Improving Agriculture Financing Opportunities for Farmers Using Blockchain Technology: A Proof-of-Concept Development and Case Scenario Illustration (p. 193). https://doi.org/10.1007/978-981-97-2211-2_9
48. Schenk, P. O., & Kern, C. (2024). Connecting algorithmic fairness to quality dimensions in machine learning in official statistics and survey production. AStA Wirtschafts- Und Sozialstatistisches Archiv, 18(2), 131. https://doi.org/10.1007/s11943-024-00344-2
49. Schreibelmayr, S., Moradbakhti, L., & Mara, M. (2023). First impressions of a financial AI assistant: differences between high trust and low trust users. Frontiers in Artificial Intelligence, 6. https://doi.org/10.3389/frai.2023.1241290
50. Sharma, P., & Logeshwaran, J. (2024). AI-driven Credit Scoring System for Real-Time Credit Card Approval in Banking. 2024 2nd International Conference on Advances in Computation, Communication and Information Technology, ICAICCIT 2024, 506–511. https://doi.org/10.1109/ICAICCIT64383.2024.10912352

51. Starke, C., Baleis, J., Keller, B., & Marcinkowski, F. (2022). Fairness perceptions of algorithmic decision-making: A systematic review of the empirical literature [Review of Fairness perceptions of algorithmic decision-making: A systematic review of the empirical literature]. Big Data & Society, 9(2). SAGE Publishing. https://doi.org/10.1177/20539517221115189
52. Talaat, F. M., Aljadani, A., Badawy, M., & Elhosseini, M. (2024). Toward interpretable credit scoring: integrating explainable artificial intelligence with deep learning for credit card default prediction. Neural Computing and Applications, 36(9), 4847–4865. https://doi.org/10.1007/S00521-023-09232-2
53. Vetrivel, S. C., Mohanasundaram, T., Saravanan, T. P., & Maheswari, R. (2024). Real-Time Analysis of Banking Data with AI Technologies. Artificial Intelligence for Risk Mitigation in the Financial Industry, 261–288. https://doi.org/10.1002/9781394175574.CH11
54. Wang, C., & Xiao, Z. (2022). A Deep Learning Approach for Credit Scoring Using Feature Embedded Transformer. Applied Sciences (Switzerland), 12(21). https://doi.org/10.3390/APP122110995
55. Xu, J., Wang, H., Zhong, Y., Qin, L., & Cheng, Q. (2024). Predict and Optimize Financial Services Risk Using AI-driven Technology. Academic Journal of Science and Technology, 10(1), 299–304. https://doi.org/10.54097/6ZRQEF25
56. Yan, T. (2024). Research on personal credit scoring model based on deep learning. Applied and Computational Engineering, 87(1), 203–208. https://doi.org/10.54254/2755-2721/87/20241558