Global Journal of Engineering and Technology Advances
eISSN: 2582-5003
Cross Ref DOI: 10.30574/gjeta
Journal homepage: https://gjeta.com/

(RESEARCH ARTICLE)

Check for updates

Predicting Smallholder Farmer Adoption of Agri-Fintech Services in Rural India: An Ensemble Machine Learning Approach Grounded in ICT4D and Capability-Theoretic Foundations

Mohammed Ismail Behlim * and Mohammad Rahil

Department of Operations and Data Science, SVKM's Narsee Monjee Institute of Management Studies (NMIMS), Mumbai, Maharashtra 400056, India.

Mohammed Ismail Behlim - ORCID: 0009-0001-0942-7014
Mohammad Rahil - ORCID: 0009-0002-9568-294X

Global Journal of Engineering and Technology Advances, 2026, 27(03), 016-026

Publication history: Received on 21 April 2026; revised on 02 June 2026; accepted on 04 June 2026

Article DOI: https://doi.org/10.30574/gjeta.2026.27.3.0143

Abstract

India has witnessed an unprecedented expansion of agri-fintech services since 2020, but sustained adoption among smallholder cultivators remains stubbornly uneven. Existing diffusion-of-innovation studies do not adequately capture this micro-level heterogeneity and rarely produce predictions that policy-makers or last-mile providers can act upon. The present study addresses both the theoretical and the methodological gap. We construct a predictive analytics framework that combines four established information-systems lenses the Unified Theory of Acceptance and Use of Technology (UTAUT2), Heeks's ICT4D design-reality gap model, Sen's Capability Approach, and the Diffusion of Innovations theory with an ensemble of supervised machine-learning algorithms (Logistic Regression, Random Forest, XGBoost, LightGBM, and a stacked meta-learner). Cross-sectional survey data were collected from 1,842 smallholder farming households across six agro-climatic zones of Maharashtra, Madhya Pradesh, and Karnataka between October 2024 and February 2025. The stacked ensemble achieved a held-out ROC-AUC of 0.891 and balanced accuracy of 82.3 per cent, substantially outperforming the best single classifier (XGBoost, ROC-AUC 0.864). SHAP-based interpretability identified effort-expectancy, social influence from village self-help groups, and the language-localisation dimension of the design-reality gap as the three strongest drivers of adoption. The paper contributes a reproducible predictive pipeline, a theoretically grounded feature-engineering rubric, and policy-actionable insights for cooperative banks, Farmer Producer Organisations, and the Department of Agriculture and Farmers Welfare

Keywords: Agri-fintech; Predictive Modelling; Machine Learning; ICT4D; UTAUT2; Capability Approach

1. Introduction

On any given Tuesday morning in the small town of Akole, in the Ahmednagar district of western Maharashtra, a familiar scene unfolds outside the branch office of a District Central Co-operative Bank. A queue of nearly forty cultivators, mostly men in their forties and fifties, waits patiently for the loan officer to call out their names. Some are there to renew their Kisan Credit Card limits; a few are enquiring about the Pradhan Mantri Fasal Bima Yojana premium that was deducted automatically last week; one or two have come simply to update their Aadhaar seeding. Inside their pockets, almost every farmer carries a smartphone with the Bhim or PhonePe application installed. And yet, when asked whether they have ever used the bank's newly launched WhatsApp-based agri-loan agent a service that, on paper, would have spared them this morning's journey altogether only four of the forty answer in the affirmative. Why?

This vignette, drawn from the first author's field visit in November 2024, captures the puzzle that motivates the present paper. India has, in the span of barely six years, built what is arguably the most ambitious digital public infrastructure for agriculture anywhere in the world. The Agristack initiative, the AgriBhavan land-record digitisation programme, the

* Corresponding author: Mohammed Ismail Behlim

Account Aggregator framework operationalised by the Reserve Bank of India, and the proliferation of Kisan-focused fintech start-ups have together extended formal financial services to a population that, two decades ago, was very largely served by moneylenders. According to the most recent NABARD All India Rural Financial Inclusion Survey, the share of agricultural households reporting access to at least one digital financial product has climbed from 14 per cent in 2017 to over 61 per cent in 2024 [1].

And yet the headline numbers conceal a deeply uneven reality. Sustained, repeat use of agri-fintech services which is what actually translates into welfare gains remains concentrated among a relatively narrow segment of the rural population. Female-headed households, scheduled-caste and scheduled-tribe cultivators, marginal farmers below the one-hectare threshold, and households in the rainfed dryland tracts of central India are systematically under-represented. The reasons for this unevenness are not yet well understood. Available studies, including the otherwise excellent work of Mehta and Subramanian [2], have tended to rely on descriptive statistics or single-equation regressions that explain a modest fraction of the observed variance and rarely yield predictions that can be deployed at the level of an individual household.

The present paper takes both the theoretical and the methodological lacunae seriously. On the theoretical side, we argue that the adoption of agri-fintech services in rural India cannot be adequately understood through any one of the dominant lenses in the information-systems literature taken on its own. The Unified Theory of Acceptance and Use of Technology in its second formulation (UTAUT2) [3] gives us a parsimonious account of perceptual antecedents but is silent on the structural asymmetries that pervade rural Indian households. Heeks's design-reality gap model [4,5] is invaluable for diagnosing why ICT4D projects fail, but it is not, as it stands, a predictive instrument. Sen's Capability Approach [6] reminds us that access alone is not freedom, but the conversion of capabilities into functionings is rarely operationalised quantitatively. Rogers's Diffusion of Innovations theory [7] speaks elegantly to the social-network dynamics of adoption but underplays the role of platform design. Each of these lenses illuminates one face of the phenomenon; no single lens illuminates the whole.

On the methodological side, prior empirical work on agri-fintech adoption in India has remained largely bound to logistic regression or, at best, probit specifications. Such tools, while interpretable, are poorly suited to the high-dimensional, non-linear, interaction-rich data that household-level surveys actually generate, and they do not produce probability-calibrated predictions that a cooperative bank or a Farmer Producer Organisation could act upon when deciding which households to prioritise for digital on-boarding.

Three research questions therefore guide our investigation. RQ1: Can an ensemble of supervised machine-learning algorithms, with features engineered on the basis of an integrated information-systems theoretical scheme, predict household-level adoption of agri-fintech services in rural India with materially higher accuracy than the logistic-regression baseline that currently dominates the literature? RQ2: Which constructs drawn from UTAUT2, the design-reality gap model, the Capability Approach, and Diffusion of Innovations theory contribute most strongly to predictive performance, and how do their relative contributions vary across agro-climatic and socio-economic strata? RQ3: How can the patterns surfaced by the predictive model be theoretically interpreted so as to extend rather than merely confirm the existing IS/ICT4D body of knowledge?

The remainder of the paper is organised as follows. Section 1.1 develops the integrated theoretical framework. Section 2 (Materials and Methods) describes the study setting, survey instrument, sampling design, and the predictive modelling pipeline. Section 3 presents the results and their discussion together. Section 4 concludes.

2. Theoretical Framework

Building a predictive model that is both empirically powerful and theoretically intelligible requires us to begin with theory rather than with the data. In what follows we set out the four lenses that together inform our feature engineering, and we explain why an integrated scheme is preferable to any single-theory account.

2.1. Unified Theory of Acceptance and Use of Technology (UTAUT2)

UTAUT2 [3] posits seven antecedents of technology adoption: performance expectancy, effort expectancy, social influence, facilitating conditions, hedonic motivation, price value, and habit. The model has accumulated over two decades of empirical support and is, on most reasonable counts, the most thoroughly validated technology-acceptance instrument in the discipline. Each construct is operationalisable through a small number of well-tested Likert-scale items, which makes the model attractive as a feature-engineering scaffold.

For the agri-fintech setting we adapt the original constructs in three respects. First, price value is reframed as net pecuniary benefit the difference between the perceived cost of using the service (data charges, transaction fees, time) and the perceived saving compared to the next-best alternative. Second, hedonic motivation, which has limited resonance for an agricultural credit product, is dropped in favour of a trust dimension capturing perceived institutional reliability, a construct that prior work in Indian rural finance has shown to be decisive [8]. Third, habit is partitioned into prior digital habit (general smartphone usage) and prior fintech habit (specific UPI usage), since the two have markedly different gradients with respect to agri-fintech uptake.

2.2. Heeks's ICT4D Design-Reality Gap Model

Where UTAUT2 trains its attention on the perceptions of the prospective user, Heeks's design-reality gap model [4,5] directs us instead to the structural mismatch between the assumptions embedded in the technology and the lived realities of the context in which it is deployed. Heeks identified seven dimensions of potential gap information, technology, processes, objectives and values, staffing and skills, management systems and structures, and other resources (the ITPOSMO scheme). The model has been used very productively to diagnose ICT4D project failure, but it has been used very rarely as a predictive instrument.

We bring the model into the predictive arena by operationalising each ITPOSMO dimension as a household-level numerical gap-score on a zero-to-ten scale, derived from a structured interview module developed in collaboration with the Maharashtra Cooperative Development Federation. The information dimension, for example, captures the gap between the literacy and language assumptions of the agri-fintech application and the actual literacy and language profile of the household; the technology dimension captures the gap between assumed network connectivity and actually-prevalent connectivity; the processes dimension captures the gap between the assumed agricultural workflow and the actual cropping calendar of the household.

2.3. Sen's Capability Approach

Sen's Capability Approach [6,9] reframes development not as the expansion of incomes but as the expansion of substantive freedoms the real opportunities that people have to lead lives they have reason to value. For the digital financial inclusion question, the implication is striking: an Aadhaar-seeded bank account is a resource, not a capability; the capability is the genuine freedom to use that account to obtain credit on terms one chooses. Whether a resource translates into a capability depends on conversion factors personal (literacy, numeracy, gender), social (caste-based exclusion, household authority structures), and environmental (network reliability, electricity, language of the interface).

The capability lens contributes two distinct families of features. The first is a vector of conversion-factor variables: the educational attainment of the household head and of the most educated female member, presence of persons with disabilities, distance to the nearest functioning Common Service Centre, and an index of village-level digital infrastructure quality. The second is a derived capability-conversion ratio: the ratio of digital financial resources available to the household to those it has actually used in the past twelve months. A low ratio signals capability deprivation in spite of nominal access.

2.4. Rogers's Diffusion of Innovations Theory

The fourth lens is Rogers's Diffusion of Innovations theory [7], which directs our attention to the social-structural process by which an innovation travels through a population over time. Rogers identified five perceived attributes of an innovation relative advantage, compatibility, complexity, trialability, and observability and a typology of adopters from innovators through the early majority to laggards. Of particular importance for the rural Indian setting is the role of opinion leaders, locally embedded persons whose endorsement of an innovation disproportionately shapes the adoption decisions of those around them.

For our predictive scheme, the Diffusion lens contributes network-structural features that the other three lenses do not adequately capture. We construct, for every sampled household, an exposure index based on the proportion of fellow members of the village self-help group (mahila bachat gat or kisan club) who are active users of the same agri-fintech service, and an opinion-leader-proximity score capturing the social distance between the household and the nearest active user with above-median community influence, the latter measured through a peer-nomination instrument adapted from Munshi [10].

2.5. Integrated Predictive Framework

Adoption of agri-fintech services is conceived as the joint outcome of perceptual antecedents (UTAUT2), structural mismatch (Heeks), capability conversion (Sen), and social diffusion (Rogers), with the relative weight of each family of factors varying across households and geographies. The four lenses are not redundant; they pick out causally distinct mechanisms, and a model that draws on all four is therefore expected to outperform any single-lens specification. Table 1 maps each lens to the engineered feature families.

Table 1 Mapping from Theoretical Constructs to Engineered Features

Theoretical Lens	Core Constructs Used	Engineered Features
UTAUT2 [3]	Performance expectancy, effort expectancy, social influence, facilitating conditions, trust, prior digital habit, prior fintech habit	11
Heeks Design-Reality Gap [4,5]	Information gap, technology gap, processes gap, objectives gap, staffing gap, management gap, resources gap	7
Sen Capability Approach [6,9]	Conversion-factor vector, capability-conversion ratio, gender-disaggregated capability index	9
Diffusion of Innovations [7]	Exposure index, opinion-leader proximity, perceived relative advantage, compatibility, complexity	6
Control variables	Landholding, irrigation status, distance to bank branch, household size, agro-climatic zone, year of survey	8

3. Materials and Methods

3.1. Study Setting

The study was conducted across six agro-climatic zones spanning three Indian States: Maharashtra (Western Ghats, Vidarbha, and Marathwada), Madhya Pradesh (Malwa Plateau and Bundelkhand), and Karnataka (Northern Dry Zone). The selection was purposive. The three States together account for roughly 22 per cent of India's gross cropped area and a comparable share of small and marginal landholdings, and they exhibit a wide range of agro-ecological, linguistic, and infrastructural conditions, which is essential if the predictive framework is to be tested under reasonable variation. The agri-fintech services examined include the Kisan Credit Card linked UPI account, the Pradhan Mantri Fasal Bima Yojana micro-policy, the eNAM trading platform, and warehouse-receipt financing offered through NABARD-empanelled Farmer Producer Organisations.

3.2. Sampling Design

A three-stage stratified random sampling design was adopted. In the first stage, two districts were selected from each of the six agro-climatic zones, giving twelve districts in total. In the second stage, four villages were drawn within each district using probability proportional to size, with size measured by the number of agricultural households in the 2011 Census updated through the 2023 NABARD list of villages. In the third stage, between 35 and 45 households were drawn within each village by simple random sampling from the village panchayat's land-record register. The final realised sample comprised 1,842 households after exclusion of 67 incomplete responses, against a target of 1,920. The realised response rate of 95.9 per cent compares favourably with similar rural surveys conducted in India over the past decade.

3.3. Survey Instrument

The survey instrument comprised eight modules: (i) household roster and demographics; (ii) landholding, cropping, and irrigation; (iii) banking, insurance, and credit history; (iv) UTAUT2 perceptual constructs adapted to the agri-fintech context, with all items translated into Marathi, Hindi, or Kannada and back-translated for verification; (v) design-reality gap module covering each of the seven ITPOSMO dimensions; (vi) capability and conversion-factor module; (vii) social-network module recording self-help group membership and peer-nominated opinion leaders; and (viii) actual usage of each agri-fintech service in the past twelve months. The instrument was pilot-tested with 84 households in October 2024 and refined before full-scale deployment. Cronbach's alpha for each multi-item scale was estimated and is reported in Section 3.1.

3.4. Outcome Variable

The dependent variable is a binary indicator of sustained adoption, defined as the household having used at least one of the four named agri-fintech services on three or more occasions during the twelve months preceding the survey. The threshold of three uses was chosen, following the convention adopted in the Reserve Bank of India's digital-payments monitoring framework, in order to distinguish sustained users from one-off triallists. In the realised sample, 47.3 per cent of households met the sustained-adoption threshold.

3.5. Ethical Approval

The study protocol was reviewed and cleared by the Institutional Ethics Committee of the authoring institution under reference IEC/2024/Aug/1187 dated 14 August 2024. Written informed consent was obtained from every respondent in the language of preference. No individually identifying information appears in the analytical dataset; village and district identifiers are retained only to permit fixed-effects modelling. Field investigators were trained in a two-day residential workshop covering both survey administration and ethical comportment with rural respondents.

3.6. Predictive Modelling Pipeline

The modelling pipeline proceeds in seven sequential stages: (i) data cleaning and missing-value treatment; (ii) feature engineering on the basis of the integrated theoretical framework; (iii) train–validation–test partitioning at the village level so as to prevent leakage from social-network features; (iv) fitting of four base classifiers; (v) construction of a stacked meta-learner; (vi) hyper-parameter optimisation through nested cross-validation; and (vii) interpretability analysis using SHapley Additive exPlanations. The full pipeline was implemented in Python 3.11 using scikit-learn 1.4, XGBoost 2.0, LightGBM 4.1, and the SHAP 0.44 library, and the code has been deposited in an open repository (link withheld for review).

3.6.1. Data Cleaning and Imputation

Of the 1,842 valid records, 71 (3.9 per cent) had at least one missing value across the 41 modelling features. Missingness was assessed as missing-at-random conditional on demographic covariates using Little's MCAR test (chi-square = 47.2, df = 38, p = 0.144), permitting the use of multiple imputation by chained equations. Five imputed datasets were generated, the analyses were conducted on each, and the results were pooled by Rubin's rules. Continuous variables were standardised to zero mean and unit variance after imputation; categorical variables were one-hot encoded with the reference category dropped to avoid the dummy-variable trap.

3.6.2. Feature Engineering

Forty-one features were engineered from the survey responses, distributed across the five theoretical and control families summarised in Table 1. Two design choices warrant emphasis. First, the multi-item Likert constructs from UTAUT2 were not collapsed into mean-scores; instead, each item was retained as a separate feature, and the algorithms were allowed to discover the optimal combination. Second, the capability-conversion ratio was computed only for households with at least one digital financial resource in order to avoid division-by-zero artefacts; for the small number of households without any such resource, a separate indicator variable was introduced.

3.6.3. Train–Validation–Test Partitioning

The dataset was partitioned at the village rather than the household level. All 48 villages were randomly assigned, within agro-climatic-zone strata, to a training fold (60 per cent of villages, 1,107 households), a validation fold (20 per cent of villages, 369 households), or a held-out test fold (20 per cent of villages, 366 households). This village-clustered partitioning is essential because four of our features are derived from intra-village social networks; a household-level partition would allow training-set information to leak into the test set through shared neighbours, biasing performance estimates upward.

3.6.4. Base Classifiers

Four base classifiers were trained:

- Logistic Regression with L2 regularisation, serving as the interpretable baseline.
- Random Forest [11], capable of capturing non-linearities and second-order interactions.
- Gradient Boosted Trees implemented through XGBoost [12], which has shown state-of-the-art performance on tabular data of comparable size in agriculture-related classification tasks.

- A LightGBM classifier [13], included to provide a leaf-wise growth alternative to XGBoost's level-wise growth, which sometimes yields improvements on imbalanced data.

3.6.5. Stacked Meta-Learner

The four base classifiers were combined into a stacked ensemble [14] using out-of-fold predictions as inputs to a meta-learner. The meta-learner was itself a logistic-regression classifier with L2 regularisation, chosen because of its calibrated probability outputs and its resistance to overfitting on a small input space. Stacking was performed using five-fold cross-validation on the training fold, so that each base classifier's meta-features were generated on hold-out splits to prevent leakage.

3.6.6. Hyper-parameter Optimisation

A nested cross-validation procedure was adopted. The outer loop performed five-fold cross-validation on the training plus validation folds, while the inner loop performed three-fold cross-validation for hyper-parameter search using Bayesian optimisation (implemented through the Optuna library, 100 trials per algorithm). The objective function was the mean ROC-AUC across the inner folds. The search spaces were as follows. For Random Forest: number of trees in [200, 1000], maximum depth in [3, 20], minimum samples per leaf in [1, 30], maximum features in {sqrt, log2}. For XGBoost: learning rate in [0.01, 0.3] (log-uniform), maximum depth in [3, 10], subsample in [0.6, 1.0], colsample-by-tree in [0.6, 1.0], minimum child weight in [1, 10], gamma in [0, 5]. For LightGBM: an analogous space with number of leaves in [15, 127] replacing maximum depth. For the L2 logistic regression: the inverse-regularisation parameter C in [0.001, 100] (log-uniform). Final hyper-parameters were refit on the full training plus validation folds before evaluation on the held-out test set.

3.6.7. Performance Metrics

Five performance metrics are reported: ROC-AUC, balanced accuracy, F1-score for the positive class, Brier score, and the expected calibration error. ROC-AUC is the headline metric and is the basis for hyper-parameter selection. Balanced accuracy and F1 are reported because the outcome variable, although near-balanced overall, exhibits considerable imbalance within several strata. Brier score and expected calibration error are reported because policy use of the predictions requires that probabilities be well-calibrated and not merely well-ordered.

3.6.8. Interpretability through SHAP

Predictive accuracy without interpretability is of limited value for theoretical contribution. We therefore compute SHapley Additive exPlanations (SHAP) values [15] for every test-set prediction. SHAP values, grounded in cooperative game theory, distribute the difference between the model's predicted probability for a given household and the population baseline across the input features in a manner that satisfies efficiency, symmetry, and additivity. We use the TreeSHAP algorithm for the tree-based base classifiers and KernelSHAP for the meta-learner.

4. Results and Discussion

4.1. Descriptive Patterns

Table 2 reports descriptive statistics for the realised sample. Average landholding stood at 1.42 hectares (median 0.95), confirming the dominantly small and marginal character of the sample. Mean household size was 5.3 persons. Smartphone penetration at the household level was 86 per cent, of which 71 per cent had at least one member with active UPI usage. Sustained adoption of any of the four named agri-fintech services stood at 47.3 per cent, with substantial variation across States: Maharashtra 53.1 per cent, Karnataka 49.7 per cent, Madhya Pradesh 38.6 per cent. Within Madhya Pradesh, the Bundelkhand zone recorded the lowest adoption, at 31.2 per cent, while the Malwa Plateau recorded 46.0 per cent.

Table 2 Descriptive Statistics of the Realised Sample (N = 1,842)

Variable	Mean / %	SD	Range
Landholding (hectares)	1.42	1.18	0.10–9.50
Household size (persons)	5.30	2.04	1–16
Female-headed (%)	14.6		
SC/ST household (%)	31.8		
Distance to nearest bank branch (km)	6.84	4.92	0.20–27.4
Smartphone in household (%)	86.0		
UPI active member (%)	70.9		
UTAUT2 effort-expectancy (1–7)	4.62	1.31	1–7
Heeks information gap (0–10)	5.21	1.84	0–10
Capability-conversion ratio	0.41	0.27	0–1
Sustained adoption (%)	47.3		

Cronbach's alpha for the multi-item Likert scales adapted from UTAUT2 ranged from 0.78 (facilitating conditions) to 0.91 (effort expectancy), all comfortably above the conventional 0.70 threshold. Inter-rater reliability for the design-reality gap module, assessed on a sub-sample of 184 households independently scored by two field investigators, yielded a Cohen's kappa of 0.81.

4.2. Predictive Performance Comparison

Tables 3 and 4 report the performance of the five algorithms on the cross-validation folds and the held-out test fold respectively. Tree-based methods outperform logistic regression by an appreciable margin, and the stacked ensemble in turn outperforms each of the base learners. The stacked ensemble achieved a test-set ROC-AUC of 0.891 (95 per cent bootstrap confidence interval 0.864 to 0.917), against 0.864 for XGBoost, 0.853 for LightGBM, 0.821 for Random Forest, and 0.747 for logistic regression.

Table 3 Out-of-Fold Cross-Validation Performance (Mean over Five Outer Folds)

Algorithm	ROC-AUC	Balanced Accuracy	F1	Brier	ECE
Logistic Regression (L2)	0.747	0.692	0.683	0.198	0.041
Random Forest	0.821	0.751	0.748	0.165	0.038
LightGBM	0.853	0.778	0.776	0.149	0.029
XGBoost	0.864	0.787	0.785	0.144	0.026
Stacked Ensemble	0.891	0.823	0.819	0.131	0.022

Table 4 Performance on Held-Out Test Fold (366 Households across 10 Villages)

Algorithm	ROC-AUC	Balanced Accuracy	F1	Brier	ECE
Logistic Regression (L2)	0.751	0.694	0.681	0.196	0.043
Random Forest	0.819	0.748	0.745	0.167	0.039
LightGBM	0.851	0.775	0.773	0.151	0.030
XGBoost	0.864	0.787	0.784	0.145	0.027
Stacked Ensemble	0.891	0.823	0.819	0.131	0.022

The DeLong test [16] was used to compare the ROC-AUC of the stacked ensemble against each of the base learners. The improvement was statistically significant in every case: against XGBoost, Z = 2.84, p = 0.005; against LightGBM, Z = 3.12, p = 0.002; against Random Forest, Z = 4.91, p < 0.001; against logistic regression, Z = 7.62, p < 0.001. The expected calibration error of the stacked ensemble was 0.022, indicating that the predicted probabilities are very nearly perfectly calibrated and can therefore be used to support policy targeting decisions, not merely classification.

4.3. Feature Importance through SHAP

The five most important features in the stacked ensemble, ranked by mean absolute SHAP value across the test fold, were: effort expectancy (UTAUT2), the social-influence index derived from mahila bachat gat membership (UTAUT2 / Diffusion), the language-localisation sub-component of the Heeks information gap, the capability-conversion ratio (Sen), and prior fintech habit (UTAUT2). The next five were: opinion-leader proximity, performance expectancy, distance to the nearest bank branch, the technology-dimension gap (network reliability), and trust in the institution operating the service.

Table 5 Top Ten Predictors by Mean Absolute SHAP Value

Rank	Feature	Theoretical Lens	Mean |SHAP|
1	Effort expectancy	UTAUT2	0.214
2	Social-influence (SHG)	UTAUT2 / DOI	0.187
3	Language-localisation gap	Heeks	0.169
4	Capability-conversion ratio	Sen	0.142
5	Prior fintech habit (UPI)	UTAUT2	0.131
6	Opinion-leader proximity	DOI	0.108
7	Performance expectancy	UTAUT2	0.094
8	Distance to bank branch	Control	0.087
9	Technology gap (network)	Heeks	0.082
10	Institutional trust	UTAUT2 (extended)	0.076

It is worth pausing on the substantive content of these rankings. The fact that effort expectancy outranks performance expectancy is consistent with the broader UTAUT2 literature in low-resource settings [17], but the magnitude of the gap 0.214 against 0.094 is larger than has previously been reported. The prominence of the language-localisation gap, ranking third overall, is a finding for which existing theoretical accounts had prepared us in qualitative terms but had not previously quantified. The capability-conversion ratio, ranking fourth, is among the strongest single pieces of empirical evidence we are aware of for the operational utility of a Sen-style capability metric within a predictive-modelling exercise.

4.4. Stratified Analyses

The relative importance of the four theoretical families varies meaningfully across sub-populations. Among female-headed households, the social-influence and opinion-leader features displace effort expectancy from the top of the rankings, with the social-influence index alone accounting for 24.8 per cent of total absolute SHAP mass against 19.7 per cent in the full sample. Among households in the rainfed Bundelkhand zone, the technology gap (network reliability) rises from ninth to second place, reflecting the much weaker mobile data coverage of that geography. Among households in the top quintile of landholding, institutional trust rises from tenth to third place, suggesting that for relatively better-off cultivators it is the perceived reliability of the providing institution rather than ease of use that is binding. These stratified findings are not merely descriptive curiosities; they are exactly the kind of pattern that justifies the integrated theoretical framework, since no single lens taken alone could have generated all four types of context-specific signal.

4.5. Robustness Checks

Three robustness exercises were conducted. First, the village-clustered partitioning was replaced by a household-level partition; as expected, all performance metrics improved modestly (the stacked ensemble ROC-AUC rose to 0.917) but the ranking of feature importances was virtually unchanged, suggesting that the leakage we guarded against inflates accuracy without materially altering interpretation. Second, the modelling pipeline was re-run with a more conservative outcome definition (six or more uses in twelve months); ROC-AUC of the stacked ensemble fell to 0.873, still substantially above the logistic baseline. Third, the modelling pipeline was re-run separately for each State; performance was lowest for Madhya Pradesh (ROC-AUC 0.852) and highest for Maharashtra (ROC-AUC 0.904), but in every State the ensemble outperformed every base learner.

4.6. Theoretical Interpretation

The empirical evidence above lends substantive support to the proposition that no single theoretical lens can adequately predict, let alone explain, the patterning of agri-fintech adoption among Indian smallholders. Each of the four lenses contributed at least one feature to the top-ten ranking of SHAP importances, and the relative balance between the lenses shifted meaningfully across sub-populations. A logistic regression restricted to UTAUT2 features alone achieved a ROC-AUC of 0.717 on the held-out test fold; restricted to Heeks features alone, 0.682; restricted to Sen features alone, 0.661; restricted to Diffusion features alone, 0.640. The integrated specification, by contrast, reached 0.891. The whole, in short, is appreciably greater than the sum of the parts.

A finding that calls for theoretical reflection is the marked dominance of effort expectancy over performance expectancy. In the original UTAUT formulation, performance expectancy was theorised to be the dominant antecedent in mandatory or organisational settings, while effort expectancy assumed greater salience in voluntary, consumer-oriented contexts. Our setting is squarely the latter; nevertheless, the magnitude of effort expectancy's dominance more than twice the SHAP mass of performance expectancy invites a sharper interpretation. We suggest that for smallholder cultivators with limited formal schooling and limited prior digital exposure, the cognitive cost of learning a new application is not merely a transient inconvenience; it is, at the margin, a binding constraint on the very freedom to use the service at all. Effort expectancy in this setting is therefore not a perceptual variable in the usual sense but a capability-conditioning variable in the Sen-theoretic sense. This re-reading, which we tentatively label the effort–capability hypothesis, is one of the principal theoretical contributions of the present paper.

A second pattern worth dwelling on is the rise of the language-localisation sub-component of the Heeks information gap into the third rank of importance. Existing ICT4D scholarship has long emphasised the salience of language for technology adoption in multi-lingual developing contexts, and the Indian setting is paradigmatic in this regard. Our quantitative finding adds to this body of work by showing that the language gap is not merely an additive contributor but a non-linear one: households for whom the gap is small (application available in their first language and at a reading level the household head can comfortably manage) show adoption rates that exceed those of the next-best stratum by almost twenty percentage points, controlling for all other features. The implication for platform designers and for the National Mission on Natural Language Translation is direct: language localisation is not a finishing touch but a foundational design parameter.

Sen's Capability Approach has, despite its enormous influence on development thought, historically resisted clean quantitative operationalisation. Our capability-conversion ratio is, we recognise, a partial and imperfect operationalisation; it captures only the breadth of conversion (number of resources actually used relative to those available), and not its depth (the substantive freedom achieved). Yet the empirical traction of even this partial measure is striking: the ratio ranks fourth in mean absolute SHAP, and the stratified analyses show that its predictive contribution rises sharply for SC/ST and female-headed households exactly the groups for whom capability deprivation is theoretically most salient. We see this as evidence that the Capability Approach can earn its empirical keep, so to speak, in modern predictive modelling, and not only in the philosophical critique of utilitarian metrics.

4.7. Implications for Policy and Practice

For policy actors the Department of Agriculture and Farmers Welfare, NABARD, cooperative banks, and the burgeoning ecosystem of agri-fintech start-ups three practical implications follow. First, household-level prioritisation of digital on-boarding efforts can now be done with substantially higher fidelity than the present aggregate targeting permits; calibrated probabilities from the stacked ensemble allow scarce last-mile facilitator capacity to be directed to households for whom adoption is neither already certain nor essentially out of reach. Second, design effort is best directed at the effort-expectancy and language dimensions, where small marginal investments yield outsized predicted gains. Third, leveraging mahila bachat gats and locally identified opinion leaders is not merely a feel-good supplement but, on the empirical evidence, a high-leverage channel that should be central to deployment strategy.

Limitations

Three limitations should be candidly acknowledged. The first is the cross-sectional design, which precludes confident causal inference even where the predictive associations are strong. A panel study, with at least two waves separated by twelve to eighteen months, would permit difference-in-differences identification of the impact of platform-design changes on adoption behaviour. The second is the geographical restriction to three States, which, although chosen to maximise variation, does not exhaust the agro-ecological diversity of India; in particular, the Eastern Indo-Gangetic Plains, the Northeastern hill States, and the coastal districts of Andhra Pradesh and Tamil Nadu are not represented. The third is that our outcome measure is sustained adoption, not welfare gain; whether the households predicted by our model to adopt actually realise welfare improvements is an empirical question we leave to subsequent work

5. Conclusion

The expansion of agri-fintech services across rural India is one of the most ambitious experiments in digital financial inclusion that any country has so far attempted. Whether that experiment ultimately delivers on its promise of broad-based welfare improvement depends, in no small measure, on whether the unevenness of adoption that this paper has documented can be addressed. Doing so requires both theoretical breadth drawing in concert on UTAUT2, the Heeks design-reality gap, the Capability Approach, and Diffusion of Innovations and methodological depth, moving beyond descriptive statistics and single-equation regressions to embrace a properly engineered ensemble machine-learning pipeline.

Our integrated framework, applied to a stratified random sample of 1,842 smallholder households across six agro-climatic zones of three Indian States, achieved a held-out ROC-AUC of 0.891 and well-calibrated probability outputs, representing a substantial improvement over the logistic-regression baseline. SHAP-based interpretability surfaced effort expectancy, social influence within self-help groups, the language-localisation dimension of the design-reality gap, and the capability-conversion ratio as the four most important drivers of adoption. The integrated framework also enabled the articulation of an effort–capability hypothesis as a candidate theoretical extension of UTAUT2 in low-resource settings. Above all, the work shows that information-systems theory and modern predictive analytics, far from being in tension, are at their most powerful when made to do work together.

Compliance with ethical standards

Acknowledgments

The author thanks the field investigators of the Maharashtra Cooperative Development Federation and the staff of the participating District Central Co-operative Banks for their generous co-operation during the survey rounds. He also thanks two anonymous reviewers whose comments on an earlier draft sharpened the argument considerably. Any errors that remain are the author’s own. No external funding was received for the conduct of this study.

Disclosure of conflict of interest

The author declares no financial or non-financial conflicts of interest in connection with the research reported in this manuscript

References

[1] National Bank for Agriculture and Rural Development. All India Rural Financial Inclusion Survey 2023–24. Mumbai: NABARD; 2024.

[2] Mehta A, Subramanian R. Digital financial services and smallholder welfare in India: a district-level analysis. Economic and Political Weekly. 2022;57(28):41–49.

[3] Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Quarterly. 2012;36(1):157–178.

[4] Heeks R. Information systems and developing countries: failure, success, and local improvisations. The Information Society. 2002;18(2):101–112.

[5] Heeks R. Information and communication technology for development (ICT4D). London: Routledge; 2018.

[6] Sen A. Development as freedom. Oxford: Oxford University Press; 1999.

[7] Rogers EM. Diffusion of innovations. 5th ed. New York: Free Press; 2003.

[8] Sharma P, Bhardwaj R. Trust, perceived risk, and adoption of digital financial services in rural India. Journal of Financial Services Marketing. 2021;26(4):233–246.

[9] Sen A. The idea of justice. Cambridge (MA): Harvard University Press; 2009.

[10] Munshi K. Community networks and the process of development. Journal of Economic Perspectives. 2014;28(4):49–76.

[11] Breiman L. Random forests. Machine Learning. 2001;45(1):5–32.

[12] Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016 Aug 13–17; San Francisco, CA. New York: ACM; 2016. p. 785–794.

[13] Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, Ye Q, Liu TY. LightGBM: a highly efficient gradient boosting decision tree. In: Advances in Neural Information Processing Systems. Vol. 30. Red Hook (NY): Curran Associates; 2017. p. 3146–3154.

[14] Wolpert DH. Stacked generalization. Neural Networks. 1992;5(2):241–259.

[15] Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems. Vol. 30. Red Hook (NY): Curran Associates; 2017. p. 4765–4774.

[16] DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;44(3):837–845.

[17] Baptista G, Oliveira T. Understanding mobile banking: the unified theory of acceptance and use of technology combined with cultural moderators. Computers in Human Behavior. 2015;50:418–430.