

# Modeling Access and Barriers to Financial Inclusion

Identifying behavioral barriers to active digital financial usage among account holders using 2025 Global Findex microdata, non-linear ML and SHAP interpretability.

Vincent Zacharias

Malini Sen

Aarav Mathur

# Account ownership has expanded but active digital usage has not

## A C C E S S

73.8%

*of adults in the raw dataset have a bank account.*

**account ownership in our 2025 project dataset**

Before modelling, we filter to ONLY account holders to study the access to usage gap

## U S A G E G A P

35.7%

*of account holders show active digital payment usage in the raw account-holder sample.*

Many account holders do not actively transact digitally.

## O U R Q U E S T I O N

?

*what predicts active digital usage*

Among people who already have access, what predicts whether they use digital payments? and do they differ across countries?

**Access ≠ usage. We study what predicts behavioral adoption, not who owns an account.**

# What the literature establishes

## A C C E S S

### Demirgüç-Kunt et al. Global Findex (2018)

#### M E T H O D

Descriptive cross-country survey statistics; account ownership as inclusion metric.

#### F I N D I N G

Account ownership ~69% globally; rapid expansion driven by DPI and state programs.

#### L I M I T A T I O N

Treats account ownership as inclusion; does not isolate active behavioral usage.

“Global Findex data shows that 69% of adults globally now have a financial account, representing major progress in financial inclusion.”  
Demirgüç-Kunt, A., Klapper, L., Singer, D., Ansar, S., & Hess, J. (2018).

## U S A G E

### RBI — National Strategy for Financial Inclusion 2025–30

#### M E T H O D

Policy framework, demand- and supply-side surveys, Financial Inclusion Index.

#### F I N D I N G

Mandate shifts beyond access to usage, quality, digital literacy and customer protection.

#### L I M I T A T I O N

Strategy-level guidance; no individual-level behavioral prediction or country-specific ranking.

## B A R R I E R S

### Digital payments & explainable-ML literature

#### M E T H O D

Mostly logistic regression on demographics; growing use of tree models with SHAP.

#### F I N D I N G

Digital literacy, trust and infrastructure shape adoption; interactions matter empirically.

#### L I M I T A T I O N

Few studies validate across countries or control for target-equivalent leakage variables.

“Digital financial technologies improve access, but behavioral and structural barriers still prevent many individuals from actively using financial services.”  
SSRN Paper (Fintech & Digital Inclusion)

**Consensus: access has expanded, but usage, quality and country-specific barriers remain uneven and under-modeled.**

# Finding the

## STANDARD APPROACH

### Typical data

- 1. Cross-country tabulations

### Demographic variables

- 1. Typical model
- 2. Logistic regression

### Country fixed effects / pooled sample

- 1. Main weakness
- 2. Limited generalization to unseen economies

## THE GAP

- Linear assumptions miss interactions
- Access ≠ behavioral usage
- Target-equivalent features cause leakage
- Random splits leak geography
- Findings ignore country heterogeneity

“Machine learning techniques can identify hidden behavioral patterns in financial data, enabling better identification of barriers to financial participation.”

JSRA (2024) – AI & Financial Inclusion

## OUR APPROACH

### Cleaner target

- 1. Active payments only
- 2. Passive receipts excluded

### Stronger validation

- 1. Leakage-controlled features
- 2. Leave-One-Country-Out CV

### Interpretability

- 1. XGBoost + SHAP
- 2. Country-specific barrier ranking

**Our contribution is not just higher accuracy — it is validated, interpretable barrier ranking across economies.**

# Dataset - Global Findex 2025 microdata



## TARGET — ACTIVE DIGITAL PAYMENTS

$Y = 1$  if respondent did at least one of:

- paid in-store digitally
- paid utility/bills digitally
- bought something online

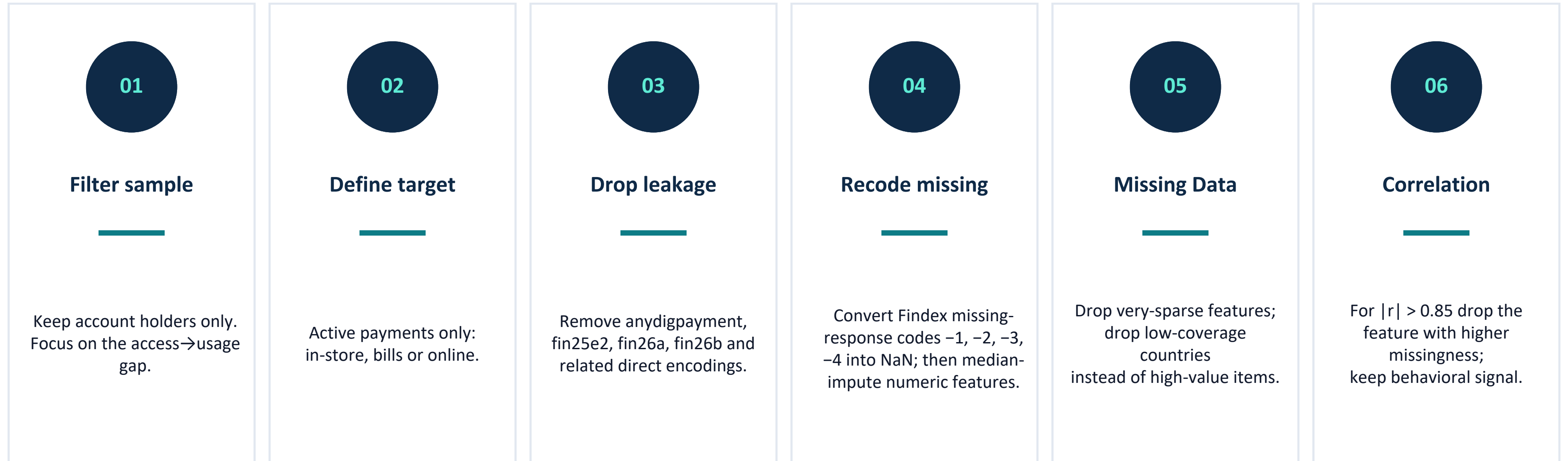
*Excluded -passive digital receipts (wages or government transfers into an account): receipt does not prove behavioral adoption.*

## SAMPLE SCOPE

- Filtered to adult account holders (access already granted)
- Restricted to emerging economies (~69 countries with the relevant payment items surveyed)
- Survey weights were retained for descriptive and country-level interpretation
- Nationally representative face-to-face / phone interviews (World Bank)

**Among account holders we isolate active behavioral usage and exclude passive receipts.**

# Pipeline - leakage prevention is the design choice



**Result** · leakage-controlled feature matrix retains behavioral, digital-capability and demographic signal - without any variable that mechanically encodes the target.

**Leakage control comes before model choice - a small misstep here invalidates every downstream interpretation.**

# Three models, one comparison frame

## BASELINE

### Logistic Regression

*Linear · interpretable · standard*

- Sklearn pipeline with StandardScaler
- L2 regularization, class-weight balanced
- Single coefficient per feature - captures direction but not interactions

## CORE MODEL

### XGBoost

*Gradient-boosted trees · non-linear · regularized*

- `n_estimators = 300`
- `max_depth = 4 · learning_rate = 0.05`
- `reg_lambda = 1.0 · min_child_weight = 2`
- `scale_pos_weight` for class imbalance
- Threshold tuned on F1 over 0.30–0.70

## EXPERIMENTAL

### Deep Neural Network

*Add-on · variance check*

- Small MLP on standardized features
- Reported as sensitivity / triangulation
- Treated as exploratory -not part of the primary inference

**XGBoost is the primary model; LR anchors interpretability and DNN tests robustness.**

# Leave-One-Country-Out cross-validation

## WHY LOCO?

- Each fold holds out one country entirely.
- Train on the remaining ~68 economies; predict on the held-out economy.
- Random splits leak country fingerprints -local digital infrastructure, language, sampling.
- LOCO directly tests generalization to unseen economies.
- Reported metrics are means and standard deviations across all country folds.

## SCHEMATIC — three example folds

*Fold 1 — test: Argentina*



*Train = 69 countries(teal) Test = 1 country(red)*

*Fold 26 — test: India*



*Train = 69 countries(teal) Test = 1 country(red)*

*Fold 68 — test: Zambia*



*Train = 69 countries(teal) Test = 1 country(red)*

**If the model only generalizes within familiar countries, LOCO will expose it -and that is the point.**

# Four-Way Model Comparison

Metric	XGBoost*	Log. Reg.*	MLP (NN)**	RF***
ROC-AUC	0.847	0.830	0.883	—
PR-AUC	0.863	0.851	0.927	—
Precision	0.749	0.734	0.843	—
Recall	0.866	0.851	0.844	—
F1-Score	0.795	0.780	0.843	—

**INTERPRETATION**

**Non-Linear Dominance:** XGBoost consistently beats LR under LOGO-CV: digital inclusion is driven by complex interactions, not simple weights.

**Geographic Generalization:** Mean across 69 separated folds confirms globally valid drivers, not localized overfitting.

**Variance Acknowledged:**  $\pm 0.12$  ROC-AUC std expected  $\rightarrow$  AUC  $> 0.93$  in Brazil/Kazakhstan,  $\sim 0.70$  in Liberia's mobile-money ecosystem.

**Black-Box Tradeoff:** MLP hits 0.883 AUC (single run), but XGBoost was chosen for native SHAP to keep findings policy-interpretable.

★ **LOGO-CV [ \* ]**  
 XGBoost & Logistic Regression metrics computed as mean across 69 Leave-One-Country-Out folds.

★ **GLOBAL BASELINE [ \*\* ]**  
 MLP on single global split. Excluded from LOGO-CV (compute cost + no SHAP).

★ **STRUCTURAL ONLY [ \*\*\* ]**  
 RF

**XGBoost prioritized over MLP for SHAP interpretability and policy relevance.**

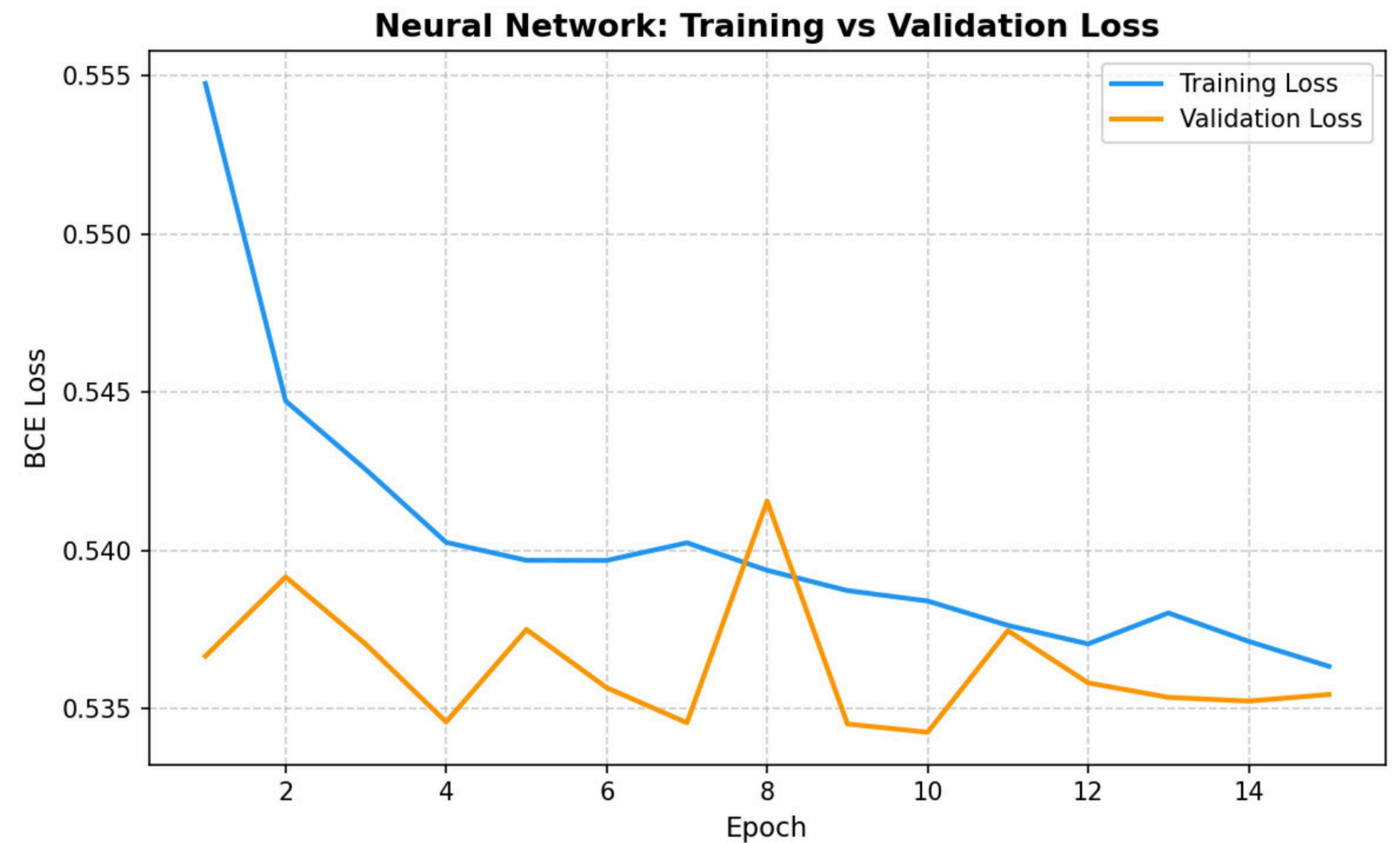
# Deep Feedforward NN as an experimental add-on

## EXPERIMENTAL

### ARCHITECTURE

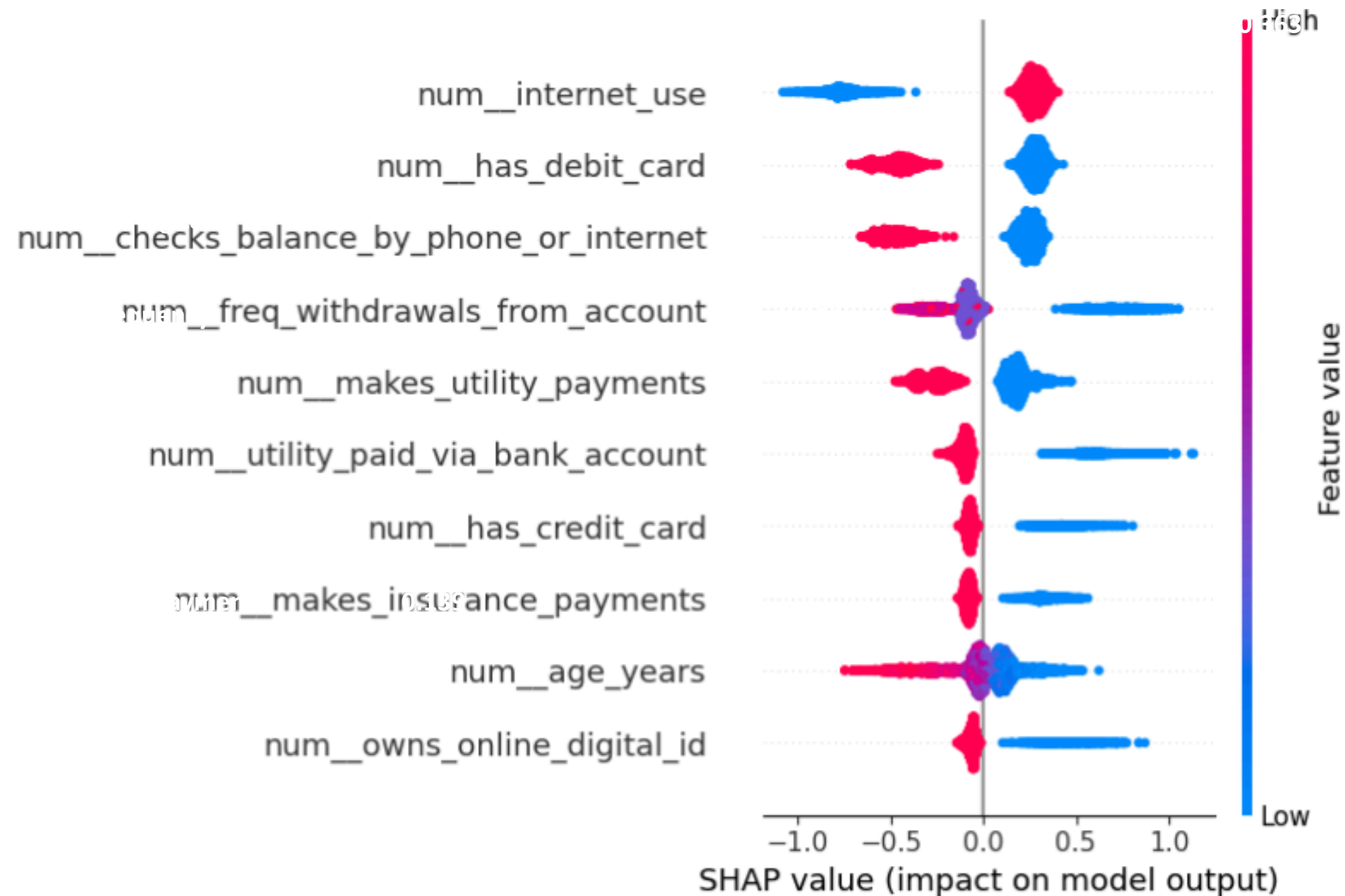
- 4-layer MLP: 128 → 64 → 32 → 16
- BatchNorm + 30% Dropout per layer
- Adam (lr = 0.005), 15 epochs, 80/20 split
- Narrow train-val gap → less chance of overfitting

## TRAINING CONVERGENCE



Despite the Neural Network's strong performance, we prioritized XGBoost: transparent SHAP insights over marginal performance gains

# Global SHAP — what predicts active digital usage



## PREDICTIVE DRIVERS

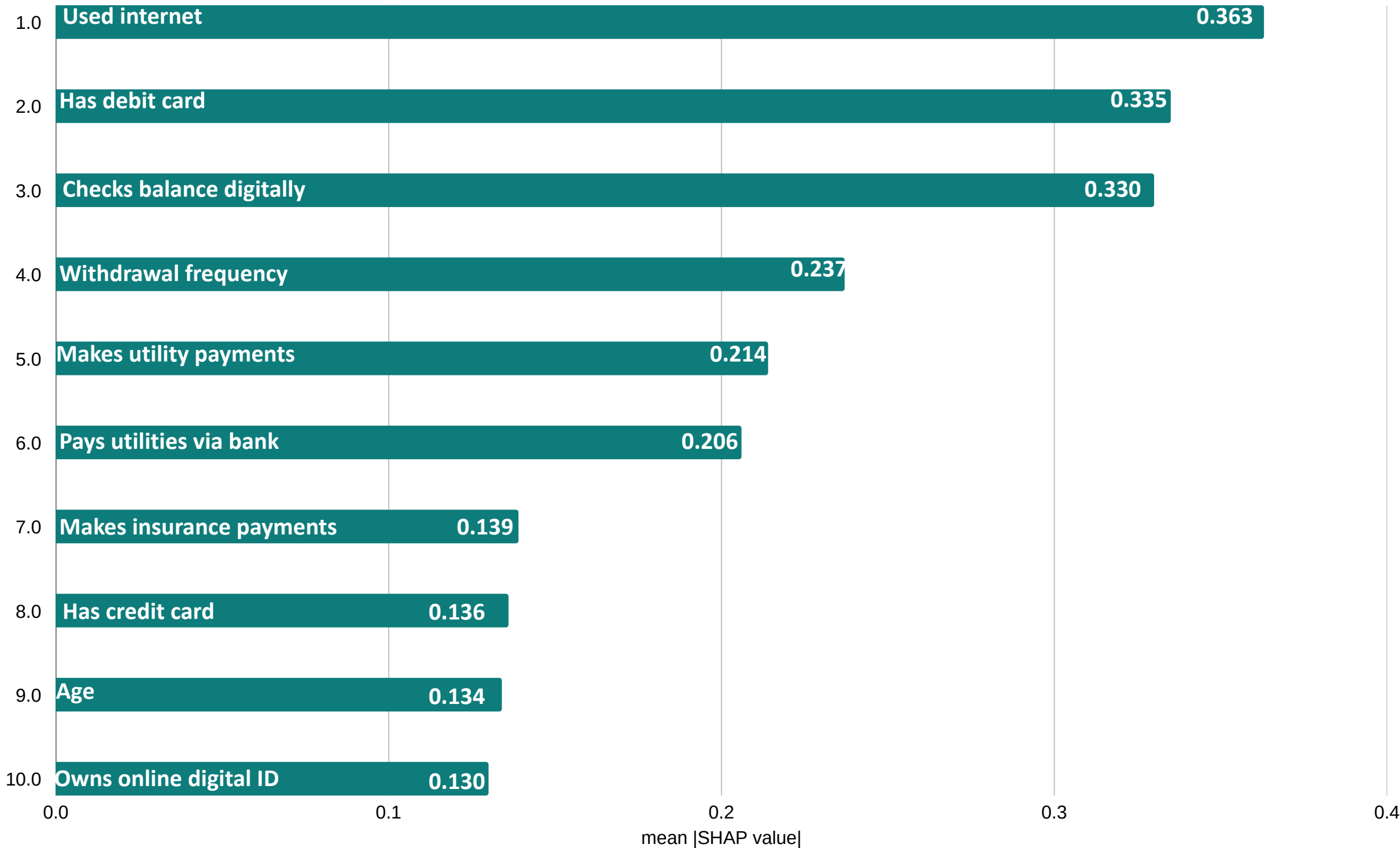
- Internet use and debit-card ownership dominate, followed by digital balance-checking.
- Behavioural items (utility payments, withdrawals) carry more weight than demographics.
- These are predictive barriers / drivers — not causal determinants.

## STABILITY (across 69 LOCO folds)

- Top 9 of 10 features classed “stable” across country folds (CV < 0.5).
- Internet Use is universally the most stable driver across all 69 countries.
- Digital ID ownership fluctuates wildly by country, acting as a highly localized driver.
- Implication: digital-capability features hold across diverse economies.

Digital capability dominates the ranking; behavior outweighs demographics in predicting active usage.

# Global SHAP — what predicts active digital usage



## PREDICTIVE DRIVERS

- Internet use and debit-card ownership dominate, followed by digital balance-checking.
- Behavioural items (utility payments, withdrawals) carry more weight than demographics.
- These are predictive barriers / drivers and not causal determinants.

## STABILITY (across 69 LOCO folds)

- Top 9 of 10 features classed “stable” across country folds (CV < 0.5).
- Internet Use is universally the most stable driver across all 69 countries.
- Digital ID ownership fluctuates wildly by country, acting as a highly localized driver.
- Implication: digital-capability features hold across diverse economies.

Digital capability dominates the ranking; behavior outweighs demographics in predicting active usage.

# Benchmark Comparison under Stricter Validation

## 01 / VALIDATION ARCHITECTURE

### LOCO Cross-Validation

- Tested on entirely unseen countries
  - Prevents geographic leakage
- Measures **true global generalization**

## 02 / TARGET DEFINITION

### Harder Prediction Task

- Restricted to active digital users
  - Excluded passive transfer recipients
- Predicts **meaningful financial participation**

## 03 / LEAKAGE CONTROL

### Aggressive Feature Pruning

- Removed target-equivalent variables
  - Dropped >0.85 correlated features
- **Performance reflects genuine signal**

## SCORE COMPARISON

Benchmark methodology undisclosed → direct comparison is indicative only

## ROC-AUC Comparison

Method	Hajjaji (2026) <sup>†</sup>	Ours
Logistic Regression	<b>0.841</b>	<b>0.830</b>
Grad. Boost	<b>0.937</b>	<b>0.847</b>

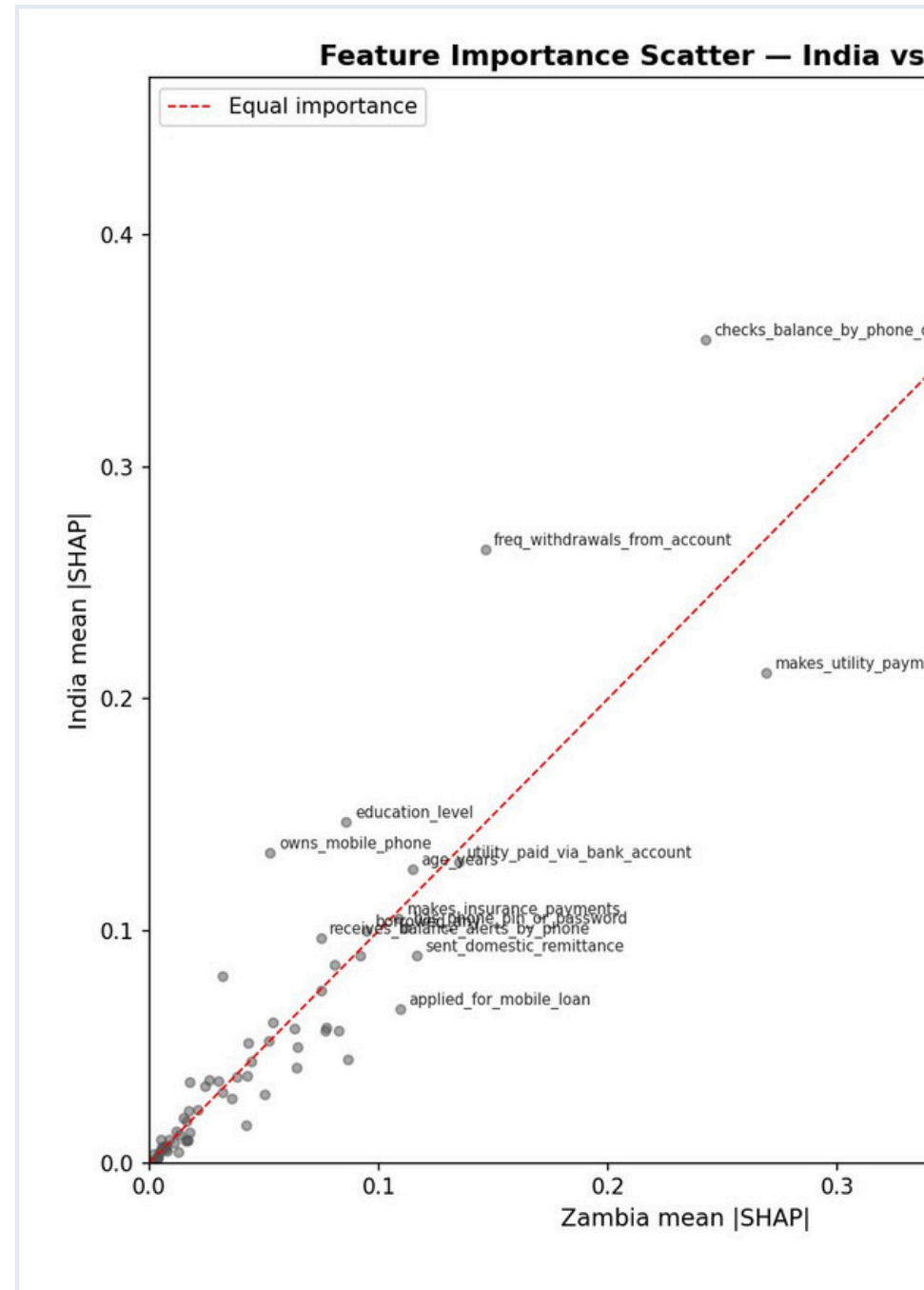
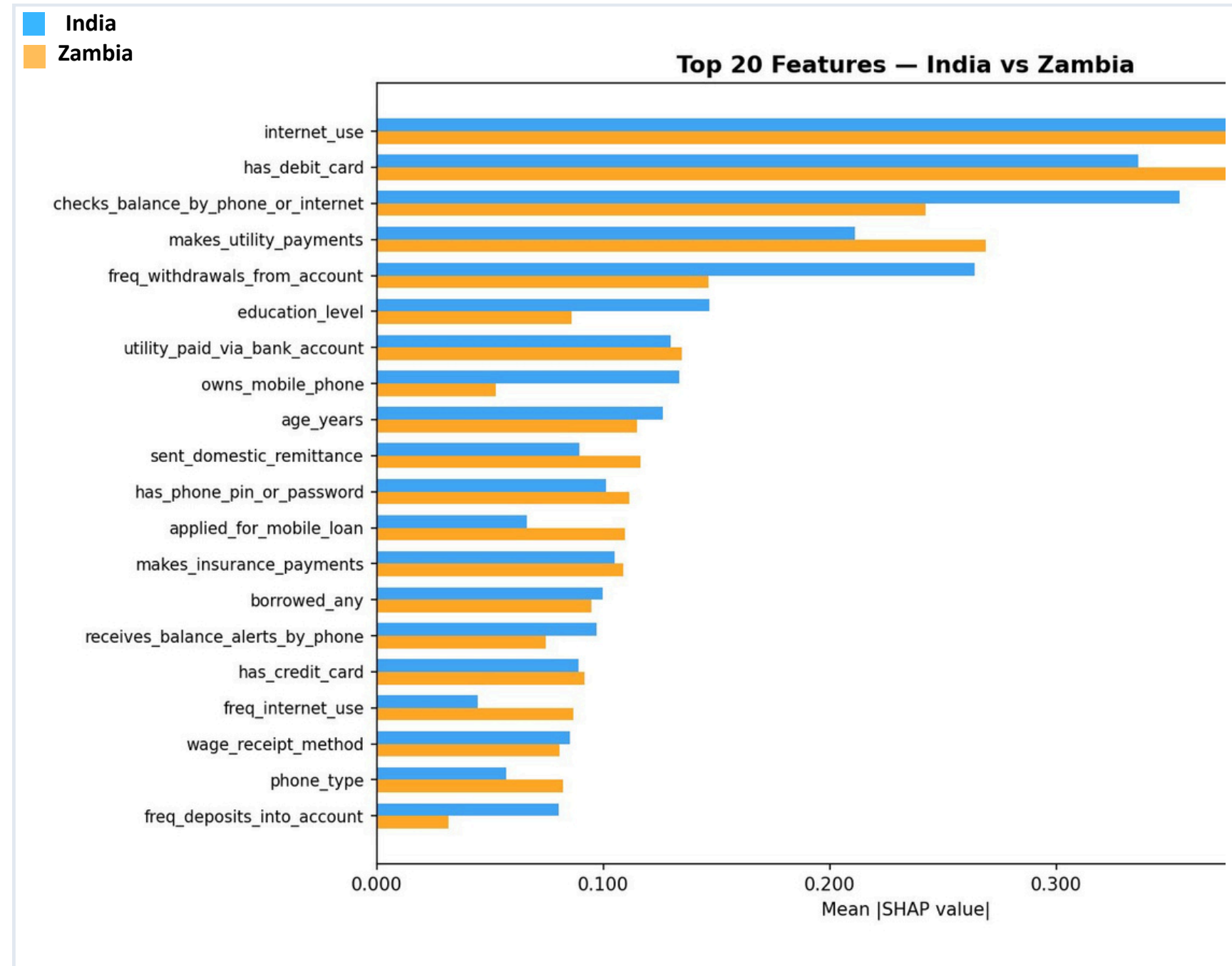
## INTERPRETATION

- Benchmark likely used standard random splits
- Our score obtained under LOCO validation
- Lower AUC reflects stricter evaluation conditions

*Hajjaji, S. (2026). Explaining digital payments adoption with econometrics and explainable machine learning: cross-country evidence from a global household survey. Digital Finance, 8(1). <https://doi.org/10.1007/s42521-026-00183-6> (ResearchGate)*

**Performance remains stable despite geographic holdout validation, leakage suppression, and a harder target definition.**

# India vs Zambia: Revealing Distinct Ecosystem Profiles



**SHARED CORE**

Internet use, debit-card ownership, and digital balance-checking lead in both countries.

**INDIA DIFFERS**

Withdrawal frequency and Education level carry more weight in India.

**ZAMBIA DIFFERS**

Owning debit card, making utility payments, and applying for mobile loan rank higher in Zambia.

**One-size-fits-all inclusion policy is empirically poorly motivated —> country-specific barrier ranking matters!**

# Navigating Challenges

CHALLENGE		SOLUTION
1	Vague Target Variable →	<b>Strict 'Active User' Filtering</b> <i>Excluded passive govt. transfer recipients</i>
2	Geographic Overfitting →	<b>LOCO Cross-Validation</b> <i>Tested on 69 fully unseen economies</i>
3	Extreme Class Imbalance →	<b>Dynamic Threshold Tuning</b> <i>Optimised F1 via class-weight rebalancing</i>
4	Data Leakage →	<b>Target-equivalent feature removal</b> <i>anydigpayment, fin25e2, fin26a, fin26b etc</i>
5	Missing Survey Data →	<b>Explicit Imputation</b> <i>Recoded Findex missing codes as NaN; median-imputed numeric features</i>
6	RAM / Compute Overload →	<b>Strategic Down-Sampling</b> <i>500 rows/fold for feasible SHAP runs</i>

# Limitations, conclusions and policy implications

## L I M I T A T I O N S

- Predictive ≠ causal inference
- Survey-response bias persists
- Heavy imputation may smooth behavior
- Leakage control removed latent signal
- F1 optimization may not reflect policy tradeoffs
- DNN benchmarking limited by compute constraints
- Vulnerable to macroeconomic data drift

## C O N C L U S I O N S

- Active digital usage is globally predictable (ROC–AUC ≈ 0.85)
- Digital capability outweighs demographics
- XGBoost captures non-linear socio-economic interactions
- LOCO validation confirms cross-country generalization
- SHAP heterogeneity reveals need for localized policy

## A L I G N M E N T W I T H R B I N S F I 2 0 2 5 – 3 0

- Usage & quality measured behaviorally, not by ownership
- Digital literacy's top predictors include internet use and balance checks
- Customer protection predictors include phone PIN ownership
- Demand-side, context-specific interventions over one-size-fits-all

**By validating across unseen economies and exposing the behavioral drivers of adoption, we replace policy guesswork with measurable intervention.**