

Predicting multi-factor authentication uptake using machine learning and the UTAUT framework

Ronald Kato¹, Aggrey Obbo¹, Richard Kimera^{1,*}

Academic Editors: Victor Sheng, Zenglin Xu

Abstract

This study investigates a machine learning-driven framework for predicting multi-factor authentication (MFA) adoption in Uganda's financial services ecosystem, a context increasingly exposed to cybersecurity risks as digital finance expands. This research aims to (i) identify key behavioral, technological and contextual determinants influencing MFA uptake, (ii) develop and validate an interpretable predictive model aligned with the Unified Theory of Acceptance and Use of Technology (UTAUT) and (iii) compare multiple classification algorithms, including classical ensembles and custom neural architectures, to establish an optimal approach for low-resource settings. Using the nationally representative FinScope Uganda 2023 dataset (survey responses = 3176), we engineered features from UTAUT constructs, security behavior indicators and digital access patterns. A binarized proxy for MFA adoption was derived from validated, high-loading security perception items. Methodologically, we implemented an experimental pipeline involving stratified train-test splitting, SMOTE applied within each cross-validation fold to avoid data leakage and repeated ($n = 30$) experiments to ensure stability of estimates. Six predictive models, Logistic Regression, Random Forest, Gradient Boosting and XGBoost alongside custom-built Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) architectures, were trained and optimized. Gradient Boosting achieved the strongest performance (mean accuracy = 0.838; F1-score = 0.835; AUC-ROC = 0.928), outperforming both linear baselines and complex neural models, which struggled with recall and F1-scores on tabular survey data. Timing analysis showed that Gradient Boosting balanced computational efficiency with predictive accuracy, making it suitable for low-bandwidth, resource-constrained environments. SHAP-based interpretability revealed that trust in digital security, prior exposure to mobile services, perceived effort and peer influence were the most influential drivers of MFA adoption. The findings advance current knowledge by integrating UTAUT constructs with explainable AI, strengthening behavioral prediction models in sub-Saharan Africa, where empirical MFA studies remain limited. This study contributes a reproducible, theory-grounded modeling pipeline, detailed comparative analysis between classical and neural network approaches and evidence-based policy recommendations.

Keywords: multi-factor authentication, technology adoption, UTAUT, machine learning, neural networks, tabular data benchmarking, cybersecurity, low-resource settings, digital finance

Citation: Kato R, Obbo A, Kimera R. Predicting multi-factor authentication uptake using machine learning and the UTAUT framework. *Academia AI and Applications* 2026;2. <https://doi.org/10.20935/AcadAI8107>

1. Introduction

The global transition to digital financial services has significantly expanded financial accessibility, yet it has also intensified cybersecurity risks, with increasing incidents of fraud and identity theft highlighting the urgency of stronger protective mechanisms [1]. These challenges are especially pronounced in low-resource settings, where limited digital infrastructure, low levels of cybersecurity awareness and regulatory gaps create heightened vulnerabilities [2]. In Uganda, for example, mobile money transactions reached USD 48 billion in 2023, involving 73% of the adult population, even as cyber-related losses surpassed USD 12 million during the same period [3]. Multi-factor authentication (MFA), which requires users to provide at least two independent authentication credentials, such as a password and a one-time code, has been widely established as an effective security control [4]. Nevertheless, MFA adoption remains strikingly low, with only 18% of users in the sample reporting that they had enabled it, leaving the majority at continued risk [5].

Adoption barriers stem from persistent infrastructural limitations, low digital literacy and widespread distrust of digital security systems. National statistics reveal that 42% of users struggle with basic digital tasks, rural Internet penetration remains at only 35% and up to 67% of the population express fear of online fraud [6]. Although such barriers are most prominent in low-resource environments, parallel challenges appear in high-resource settings as well, where usability frustrations and privacy concerns continue to impede the adoption of security mechanisms. Machine learning (ML) approaches have been applied to strengthen authentication processes globally, including the use of explainable models in e-commerce security [7] and Support Vector Machines in mobile banking protection [8, 9]. These studies demonstrate the broader applicability of ML-driven security enhancements but do not address the contextual complexities of low-resource environments.

While the technical robustness of MFA is widely acknowledged

¹Department of Information Technology, Mbarara University of Science and Technology, Mbarara, Uganda.

*email: rkimera@must.ac.ug

in the cybersecurity literature, there remains limited empirical investigation into the determinants of MFA adoption in low-resource financial ecosystems such as Uganda [10]. Much of the existing global scholarship insufficiently accounts for contextual constraints such as the rural–urban digital divide, where digital literacy reaches 65% in urban areas but falls to 28% in rural regions, nor does it integrate behavioral models that reflect these disparities. Most machine learning models developed to predict MFA uptake originate from high-resource countries and overlook the usability challenges, limited connectivity and infrastructural barriers that shape technology adoption in under-resourced contexts [7].

The application of machine learning to predict technology adoption and cybersecurity behavior is an emerging but increasingly important research direction. Prior studies have effectively used ML techniques to model user behavior in cybersecurity-related domains. For example, Sagar et al. employed ML algorithms to detect phishing attacks [11], demonstrating the capacity of these methods to predict complex security behaviors. Other investigations [12, 13] have used ML to identify factors driving the adoption of health technologies, a field closely aligned with security technology adoption due to shared dimensions of trust, perceived risk and user readiness. However, there remains a clear gap in the literature, as few studies have combined UTAUT-derived constructs with behavioral security indicators in an ML framework, particularly using a nationally representative dataset to model MFA adoption.

Furthermore, while classical machine learning has established a strong baseline for tabular data analysis, the recent global shift toward deep learning necessitates an empirical evaluation of whether more complex architectures can better capture the latent, non-linear dependencies between behavioral constructs. To date, there is limited evidence comparing the efficacy of tree-based ensembles against neural network architectures, such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, specifically within the context of technology acceptance in sub-Saharan Africa [14]. This study explores the performance of traditional classifiers against these deeper architectures to determine if the spatial or sequential inductive biases of neural networks offer a significant advantage for modeling high-dimensional survey data in resource-constrained environments.

This study addresses these gaps by utilizing the FinScope Uganda 2023 dataset comprising 3176 respondents to develop an interpretable machine learning framework grounded in the Unified Theory of Acceptance and Use of Technology (UTAUT) [15]. This research aims to identify the main behavioral and contextual drivers of MFA uptake, quantify their influence, build a predictive model and generate actionable recommendations for improving digital security. Although designed for a low-resource environment, the proposed framework has broader relevance for global contexts where digital inclusion and cybersecurity readiness remain critical concerns [16]. The technical contribution lies in integrating UTAUT constructs with explainable ML techniques, including SHAP and LIME, for behavioral patterns influencing MFA adoption across diverse user segments.

1.1. Related works

The technical effectiveness of multi-factor authentication in mitigating cyber threats is well established in cybersecurity research; however, considerably less attention has been given to the determinants of MFA adoption in low-resource settings such as Uganda [6, 17]. Much of the existing global works assumes environments with stable connectivity, high digital literacy and adequate regulatory frameworks, overlooking the contextual realities of under-resourced ecosystems. In Uganda, the rural–urban digital divide remains substantial, with digital literacy reaching 65% in urban areas but only 28% in rural regions, creating asymmetrical conditions for technology uptake. Existing adoption studies rarely capture these disparities and even fewer utilize behavioral frameworks tailored to such environments. This study responds to this gap by adopting the Unified Theory of Acceptance and Use of Technology (UTAUT), which has demonstrated strong explanatory power in technology adoption research and offers a systematic approach for analyzing behavioral intentions within Uganda’s financial services context [2, 3].

Machine learning-driven approaches for predicting MFA adoption have predominantly been developed and validated in high-resource environments where infrastructural reliability, device penetration and cybersecurity awareness levels are significantly higher [7]. These models typically emphasize algorithmic performance while paying limited attention to usability, cultural perceptions and infrastructural barriers that influence adoption behaviors in low-resource settings [16, 18]. Consequently, their applicability to under-resourced financial ecosystems remains restricted, reinforcing the need for context-specific predictive modeling grounded in local realities.

More broadly, the application of machine learning to forecast technology adoption and cybersecurity behavior is an emerging yet increasingly important research domain. ML techniques have been extensively used to characterize user behavior in security-sensitive contexts. For example, Pandya and Zalawadia (2025) applied ML algorithms to detect phishing attacks, demonstrating how computational models can capture nuanced behavioral patterns linked to cybersecurity vulnerabilities [19]. Similarly, Preti et al. (2024) used ML approaches to predict adoption factors for health technologies, an area that parallels security technology adoption given its dependence on user trust, perceived risk and digital literacy [20]. Despite this growing body of literature, few studies have systematically integrated UTAUT constructs with behavioral security indicators in an ML framework, particularly when using large-scale, nationally representative datasets to model MFA adoption.

This study leverages the FinScope Uganda 2023 dataset to build an interpretable, context-aware predictive model for MFA adoption. By combining behavioral theory with data-driven insights, this study generates tailored, locally relevant findings that advance the understanding of authentication adoption dynamics in low-resource financial ecosystems [17]. This integration of national-level behavioral data, UTAUT constructs and machine learning models provides a significant step toward developing scalable solutions that reflect the realities of digitally constrained populations.

2. Materials and methods

This study followed a structured methodology, as illustrated in **Figure 1**, informed by the Unified Theory of Acceptance and Use of Technology and NIST SP 800-63B authentication standards [21] to ensure behavioral and security alignment. Variables reflected UTAUT constructs; performance expectancy, effort expectancy, social influence, and facilitating conditions while additional indicators derived from the NIST guidelines captured perceptions of security and authentication robustness. Following data cleaning, variables were mapped to theory-based categories and engineered into model-ready features, enabling the integration of behavioral theory and machine learning. Preprocessing steps included categorical encoding, numerical feature scaling, and class imbalance handling. Machine learning models, including Logistic Regression, Random Forest, Gradient Boosting and XGBoost, were trained using repeated cross-validation, with consistent evaluation metrics applied throughout the predictive modeling process.

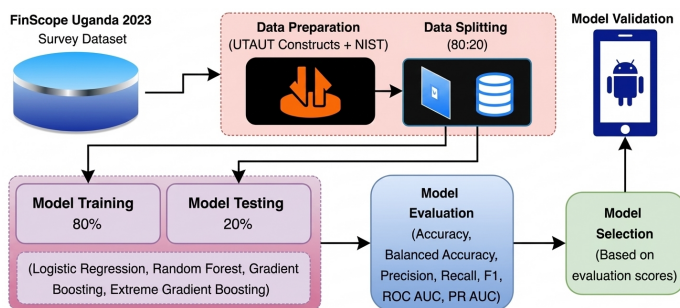


Figure 1 • Overview of the proposed MFA adoption pipeline.

2.1. Dataset

This study utilized the FinScope Uganda 2023 Survey dataset [22], a nationally representative financial inclusion survey conducted by the Bank of Uganda in collaboration with Financial Sector Deepening Uganda, Uganda Bureau of Statistics, Abi Finance and the Ministry of Finance. The dataset contains 3176 responses from individuals aged 16 and above and includes detailed behavioral, demographic and technological access information. Variables related to digital savings, electronic transfers, fraud perception, mobile money usage, device access and cybersecurity awareness provided a robust foundation for modeling MFA adoption. The dataset's breadth offers a unique opportunity to explore adoption dynamics in low-resource settings and to develop evidence-based policy recommendations aimed at improving secure digital financial participation.

2.2. Data preparation

Data preparation began with the screening of all 3176 entries, identifying missing values primarily in Likert-scale items. Missing

numeric values were imputed using median values to minimize the influence of extreme responses, while categorical variables were imputed using their modal class to retain natural distribution patterns. Outliers were defined as values below the 1st percentile or above the 99th percentile for usage-related metrics. These were removed to reduce distortion, resulting in a cleaned analytical subset of 2540 valid responses. Numerical trust- and security-related items were standardized to improve convergence during model training and binary variables such as rural versus urban residence were encoded using a 0–1 scheme.

To construct a behavioral proxy for MFA adoption, multiple indicators reflecting secure digital engagement were combined. The anchor variable was item h1_1_6 (“Is it safer to use cards or machines than cash?”), binarized as 1 for values above 2. Additional indicators, mobile payments (km1_1_3), regular PIN changes (km1_1_7), digital banking use (h2_2_5) and mobile wallet use (h2_2_8), were similarly binarized. The composite MFA adoption variable was coded as 1 if the average of these binary items exceeded 0.5, indicating that a respondent practiced at least half of the MFA-related behaviors. Reliability was assessed through inter-item correlations and Cronbach’s alpha, confirming adequate internal consistency and validating the construct as a suitable proxy for MFA adoption.

2.3. UTAUT construct mapping of variables

A theory-driven mapping process assigned variables to six UTAUT constructs: performance expectancy (PE), effort expectancy (EE), social influence (SI), facilitating conditions (FC), security perceptions (SP) and demographic factors. The mapping followed established behavioral theories and was cross-checked against NIST SP 800-63B guidelines to ensure alignment with security best practices. Reliability analysis yielded Cronbach’s α (thresholds: ≥ 0.7 for behavioral constructs and ≥ 0.8 for security), indicating strong internal consistency. **Figure 2** summarizes this mapping, which structured the analytical process and supported explicit integration of behavioral variables into the machine learning pipeline.

2.4. Descriptive statistics of key constructs used

Descriptive statistics for all UTAUT constructs are presented in **Table 1**. Performance expectancy (mean = 2.24), effort expectancy (mean = 1.55), facilitating conditions (mean = 1.69) and security perceptions (mean = 2.16) all score below the mid-point of the five-point scale, indicating limited confidence in MFA usability and security. Social influence (mean = 3.30) and demographic factors (mean = 3.08) score higher, suggesting that community norms and user characteristics may strongly shape adoption patterns. These patterns reveal substantial perception gaps consistent with low-resource environments and demonstrate the importance of context-sensitive models for predicting MFA uptake.

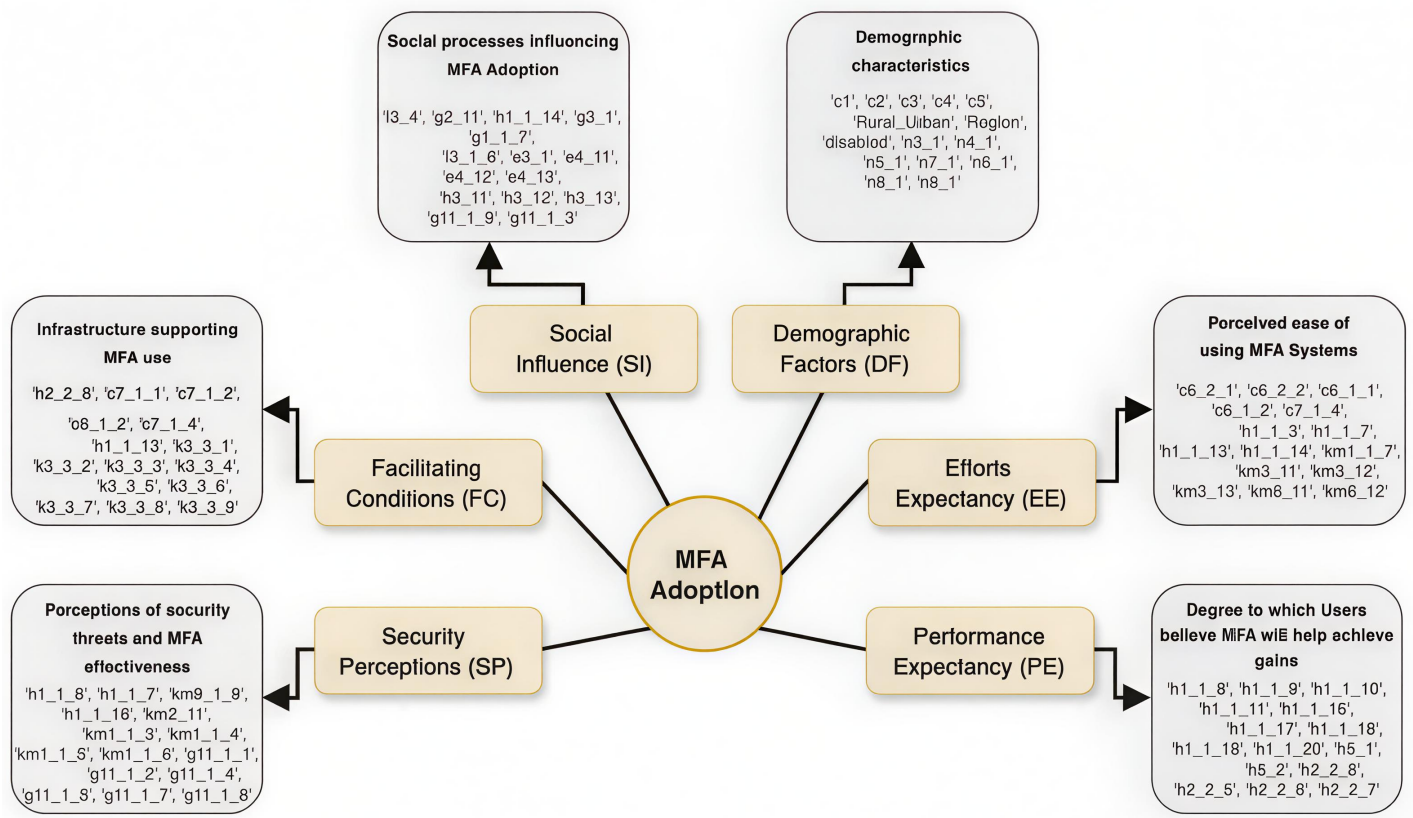


Figure 2 • UTAUT construct mapping of variables.

Table 1 • Descriptive statistics of key constructs.

Constructs	Count	Mean	Std dev	25%	50%	75%
Performance expectancy score	3176	2.24	0.74	1.50	2.50	3.00
Effort expectancy score	3176	1.55	0.30	1.40	1.60	1.60
Social influence score	3176	3.30	0.80	2.83	3.50	3.83
Facilitating conditions score	3176	1.69	0.36	1.40	1.75	1.83
Security perceptions score	3176	2.16	0.66	1.67	2.00	2.75
Demographic factors score	3176	3.08	1.79	2.40	2.80	3.40

Mean Scale: 1 = Very Low to 5 = Very High.

2.5. Exploratory data analysis

Exploratory data analysis was conducted using Python v3.10 libraries, including Pandas v2.0 and Seaborn v0.13. Reliability tests confirmed Cronbach's alpha values above 0.70 for all UTAUT constructs, indicating acceptable internal consistency. Feature selection narrowed the analysis to thirty-five variables that reflected digital access, trust, ease of use and fraud experiences. Descriptive trends highlighted that 2520 respondents comfortably used mobile phones, yet 2419 reported inconsistent Internet access and as many as 1600 relied on feature phones. Although 771 respondents used mobile money regularly, a critical channel for MFA, behavioral barriers, persisted: 1490 feared fraud, 450 had previously shared PINs and only 945 trusted cards over cash despite most not having experienced fraud. These patterns underscore the role of trust, literacy and infrastructure in shaping adoption behavior.

2.6. Spatial analysis of MFA adoption rates

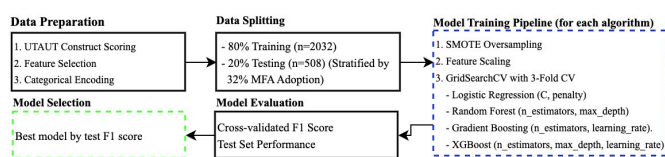
The map in **Figure 3** highlights significant spatial disparities in MFA adoption across Uganda's districts. Urban areas like Kampala and Wakiso show unexpectedly low adoption, while rural districts, Rukiga, Nebbi and Kyankwanzi, report high adoption. This contradicts assumptions about urban digital leadership and points to complex socio-technological influences. Factors such as poor infrastructure, limited digital access and low trust contribute to lower adoption rates in some areas. These dynamics align with the UTAUT framework, particularly facilitating conditions and effort expectancy. The spatial data also serves as a foundation for machine learning models that integrate socioeconomic and geographic features to predict adoption trends.

Table 2 • Hyperparameter search spaces used for model optimization.

Model	Hyperparameter	Search values
Logistic Regression	Regularization penalty	L2
	Regularization strength (C)	{0.1, 1, 10}
	Solver	liblinear
Random Forest	Number of estimators	{100, 300, 500}
	Maximum tree depth	{None, 10, 20}
	Minimum samples per split	{2, 5}
	Minimum samples per leaf	{1, 2}
Gradient Boosting	Learning rate	{0.01, 0.05, 0.1}
	Number of estimators	{100, 200, 300}
	Maximum tree depth	{3, 5}
	Subsample ratio	{0.8, 1.0}
XGBoost	Learning rate (η)	{0.01, 0.1}
	Maximum tree depth	{3, 6}
	L1 regularization (α)	{0, 1}
	L2 regularization (λ)	{0, 1}
	Subsample ratio	{0.8, 1.0}
	Column subsample by tree	{0.8, 1.0}

2.8. Model training

This study trained models to predict MFA adoption using the dataset guided by a robust pipeline (see **Figure 4** for the model training process). Four algorithms, Logistic Regression, Random Forest, Gradient Boosting and XGBoost, were chosen for their binary classification strengths. Logistic Regression supports interpretability of predictors including trust and usability, while Random Forest captures non-linear UTAUT relationships [31]. Gradient Boosting enhances prediction through sequential error correction [32] and XGBoost provides computational efficiency for large datasets [33].

**Figure 4 •** Model training process.

Data was split into 80% training (2032 samples) and 20% testing (508 samples) using StratifiedKFold. Class imbalance (32% adopters) was addressed using SMOTE [34], which synthesized minority class samples to improve sensitivity. Hyperparameter tuning via GridSearchCV and three-fold cross-validation optimized model parameters. The ModelTrainer module integrated feature engineering [35] from behavioral predictors and UTAUT constructs, hence aligning training with this study's framework.

3. Results

Table 3 presents the comparative performance of four classification algorithms; Logistic Regression, Random Forest, Gradient Boosting and XGBoost applied to the prediction of multi-factor authentication (MFA) adoption within Uganda's financial services ecosystem. Logistic Regression performs strongly for a linear model, achieving an accuracy of 0.836 and an AUC-ROC of 0.914, indicating that well-engineered features can capture meaningful behavioral signals even under linear assumptions. Random Forest marginally improves discriminative ability (AUC-ROC = 0.915) but exhibits a lower F1-score (0.808), suggesting a reduced balance between precision and recall.

Table 3 • Performance comparison of classification models across evaluation metrics.

Model	Accuracy	F1-score	AUC-ROC
Logistic Regression	0.836	0.830	0.914
Random Forest	0.822	0.808	0.915
Gradient Boosting	0.838	0.835	0.928
XGBoost	0.764	0.783	0.909

Gradient Boosting outperforms all competing models across evaluation metrics, attaining the highest accuracy (0.838), F1-score (0.835) and AUC-ROC (0.928). This performance confirms its superior capacity to model complex interactions among behavioral, technological and contextual determinants of MFA adoption. In contrast, XGBoost underperforms in this setting (accuracy = 0.764, AUC-ROC = 0.909), potentially due to sensitivity to hyperparameter configurations or overfitting effects in low-resource data environments. Overall, these results identify

Gradient Boosting as the most technically suitable model for MFA prediction in Uganda's financial sector.

Beyond absolute performance, the results demonstrate both methodological alignment and contextual relevance relative to prior studies. Van Biljon [36] emphasizes that ensemble learning methods such as Random Forest and Gradient Boosting consistently outperform linear models in applied machine learning research across sub-Saharan Africa, particularly in finance and healthcare domains characterized by data heterogeneity and resource constraints. The superior performance of Gradient Boosting in **Table 3** corroborates this observation, reinforcing the robustness of ensemble approaches in low-resource environments.

Similarly, Sodeeq [37] highlights the importance of contextualized, lightweight and interpretable algorithms for financial services in Africa, noting that computational efficiency is critical for deployment in bandwidth- and infrastructure-constrained ecosystems. The timing and deployment analysis in this study supports this claim, demonstrating that Gradient Boosting achieves a favorable balance between predictive accuracy and computational efficiency, making it well-suited for operational use in Uganda's financial sector. Furthermore, the Paradigm Initiative [38] underscores trust and digital literacy as pivotal determinants of technology adoption across African economies. This finding is echoed in the SHAP-based interpretability analysis, which identifies trust in digital security and prior exposure to mobile financial services as the strongest drivers of MFA uptake. Collectively, these benchmark studies situate the present work within a growing body of evidence emphasizing the dual importance of algorithmic efficiency and socio-behavioral factors in advancing secure and inclusive digital finance in sub-Saharan Africa.

3.1. Model performance comparison

A comparative evaluation of the four algorithms, Logistic Regression, Random Forest, Gradient Boosting and XGBoost, is illustrated in **Figure 5**. Gradient Boosting demonstrates superior performance across all evaluation metrics, achieving an accuracy of 0.838, an F1-score of 0.835 and an AUC-ROC of 0.928. These results, obtained under repeated cross-validation with SMOTE applied strictly within each training fold, indicate the model's strong capacity to balance precision and recall in the presence of class imbalance.

Logistic Regression and Random Forest also perform competitively, each recording AUC-ROC values exceeding 0.91, reflecting reliable discriminative capability. In contrast, XGBoost exhibits comparatively weaker performance, suggesting the need for further parameter refinement or feature engineering.

3.2. SHAP insights and categories

SHAP explainability was used to interpret Gradient Boosting's predictions, and the summary plot in **Figure 6** identifies security perceptions and performance expectancy as the strongest contributors to MFA adoption. High SHAP values for these features indicate that users who trust digital systems and expect clear benefits from MFA are more likely to adopt it. Parallel category visualizations in **Figure 7** reinforce this pattern, showing that respondents with higher awareness and digital confidence accounted for the majority of adopters (approximately 1200), while low-awareness users overwhelmingly fell into the non-adopter category (around 400). These findings underscore the critical role of perception, digital readiness and trust in shaping the adoption of authentication technologies in low-resource financial environments.

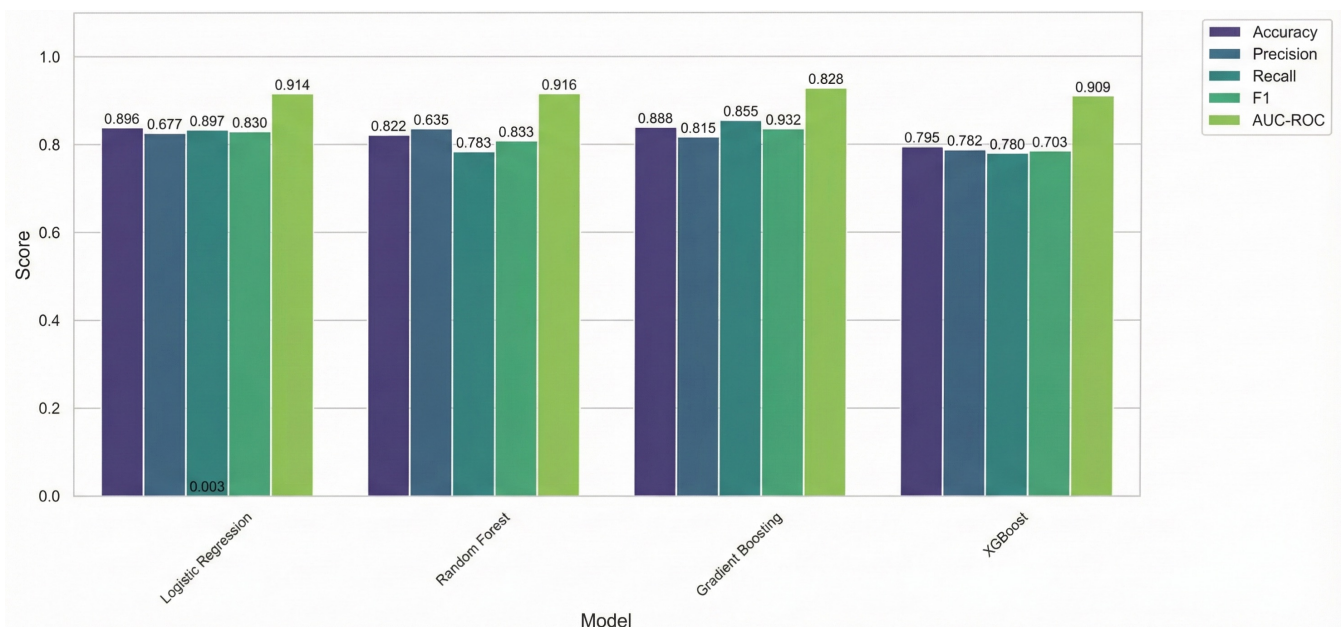


Figure 5 • Model performance comparison.

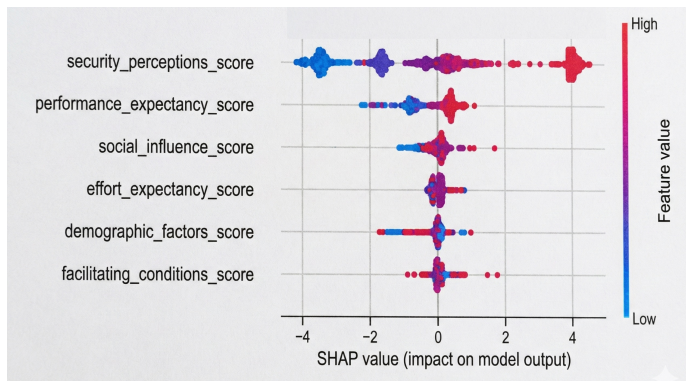


Figure 6 • SHAP summary plot (Gradient Boosting).

Awareness → Adoption Flow

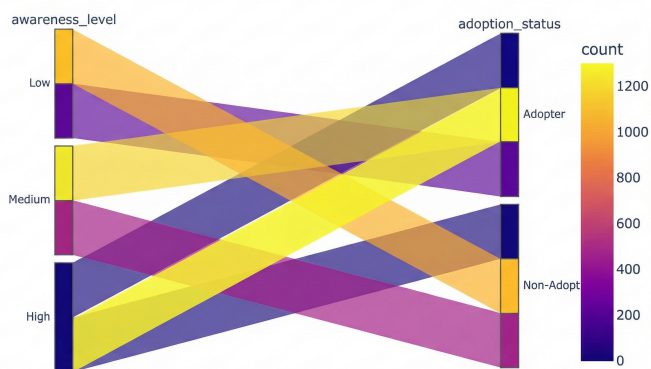


Figure 7 • Parallel categories.

3.3. Logistic regression and ROC comparison

The Receiver Operating Characteristic (ROC) curves presented in **Figure 8** further illustrate the predictive strength of the evaluated models, with Gradient Boosting achieving the highest area under the curve (AUC = 0.93). This result reinforces its superior discriminative capability across a range of classification thresholds.

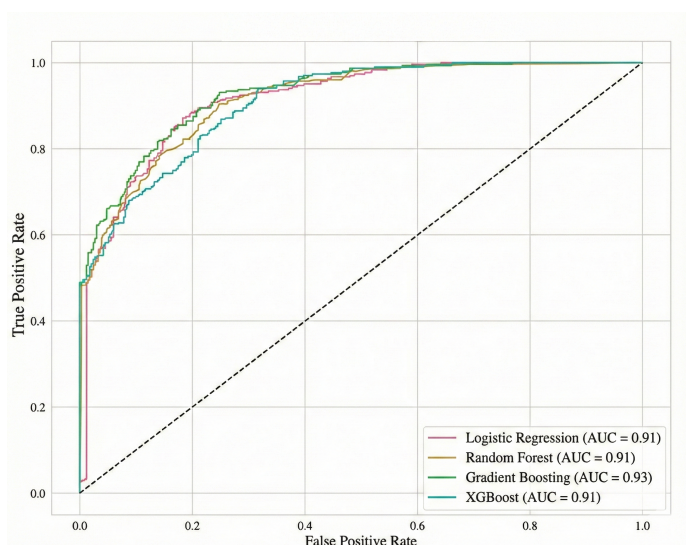


Figure 8 • ROC curve comparison.

Logistic Regression results, summarized in **Figure 9** using odds ratios, identify security perception and performance expectancy as the most influential positive predictors of MFA adoption (OR > 1). In contrast, effort expectancy exhibits a negative association with adoption likelihood, suggesting that perceived complexity or usability barriers reduce the probability of enabling MFA.

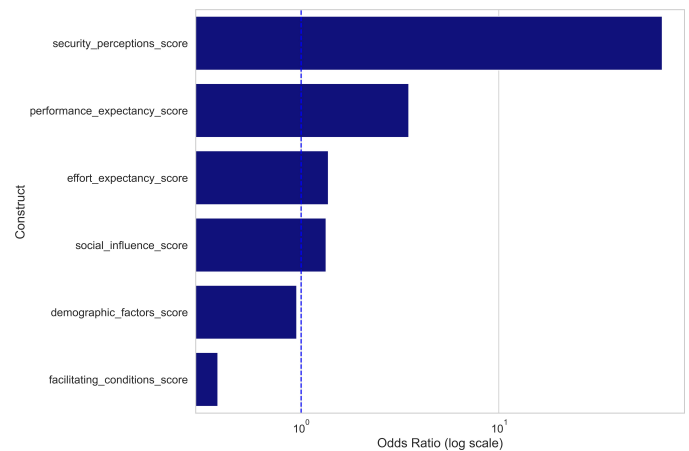


Figure 9 • Odds ratios from Logistic Regression.

3.4. Latent space analysis

Latent space visualization using Principal Component Analysis (PCA) [39] provided additional insight into the internal structure of the trained model. **Figure 10** illustrates the distribution of users in the reduced feature space, colored by predicted MFA adoption probability, ranging from low (blue) to high (red). The visible separation between adopters and non-adopters demonstrates the model's capacity to capture discriminative behavioral patterns within the latent space.

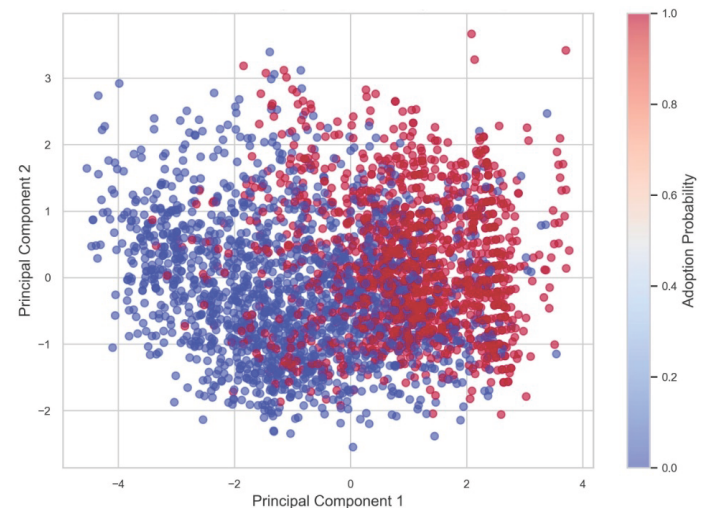


Figure 10 • Latent space projection.

Figure 11 further applies K-means clustering to the PCA-transformed features, revealing four distinct user clusters characterized by differing predicted probabilities of MFA adoption. These clusters correspond to meaningful behavioral profiles, including high-trust users, digitally cautious users and infrastructure-constrained users.

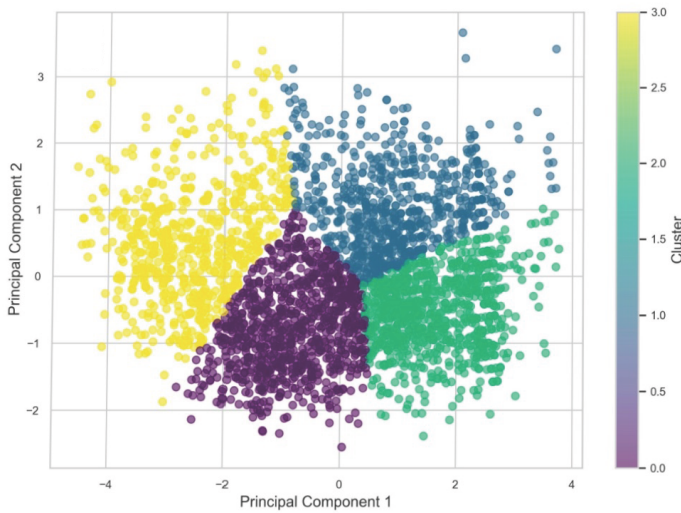


Figure 11 • Latent space clustering.

3.5. Model validation

To assess the real-world applicability of the top-performing model, a mobile application prototype in **Figure 12** was developed using Kotlin for the Android platform. The selection of Android was deliberate, given its widespread adoption in low-resource environments. The application enables users to input self-assessed scores for five UTAUT constructs, performance expectancy, effort expectancy, social influence, facilitating conditions and security perceptions, each measured on a five-point Likert scale (1–5). Based on these inputs, the trained Gradient Boosting model computes a personalized probability of MFA adoption and provides tailored, context-aware feedback.

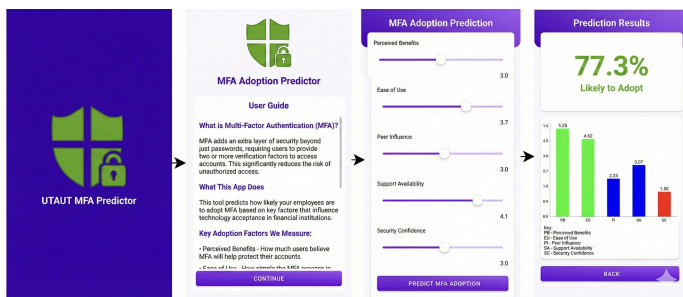


Figure 12 • UTAUT MFA predictor app.

User acceptance testing was conducted with 120 participants, evenly split between urban ($n = 60$) and rural ($n = 60$) users. The results indicate that 83% of participants found the interface intuitive, 78% reported that the model's predictions aligned with their self-reported behavior and 67% considered the feedback useful for supporting decision-making. Notably, rural participants highlighted challenges related to intermittent connectivity and varying levels of digital literacy, underscoring the importance of streamlined interfaces and offline-capable MFA awareness tools for equitable deployment.

The validation outcomes demonstrate that integrating machine learning-driven predictions with user-centered design principles

can substantially enhance the accessibility, usability and practical relevance of MFA adoption tools, particularly within resource-constrained financial ecosystems.

3.6. Comparative performance of classical and neural network models

To evaluate the potential benefits of non-linear feature extraction beyond classical ensemble methods, two custom deep learning architectures, a Convolutional Neural Network (CNN) and a Long Short-Term Memory (LSTM) network, were developed and assessed. Both models were implemented using framework-independent NumPy code to ensure full transparency of the learning dynamics and to evaluate feasibility in environments without reliance on high-level deep learning libraries.

The CNN architecture was designed for one-dimensional feature extraction across UTAUT construct dimensions. It employed 16 convolutional filters with a kernel size of 2, followed by a Rectified Linear Unit (ReLU) activation and Global Average Pooling to reduce feature dimensionality. Final binary classification was performed using a sigmoid-activated output layer.

To explore potential sequential dependencies between behavioral constructs, the LSTM model was configured with eight hidden units initialized using Xavier initialization. UTAUT constructs were interpreted as ordered input sequences to capture inter-construct relationships. Both neural models utilized an 80:20 stratified train-test split. Hyperparameters were optimized through manual grid exploration, resulting in a learning rate of 0.05 and shallow architectures to mitigate overfitting given the limited feature space.

The results in **Table 4** reveal a clear performance gap between classical ensemble models and neural architectures. First, both CNN and LSTM exhibit substantially lower recall and F1-scores compared to tree-based methods, indicating architectural misalignment with the prediction task. Although the LSTM achieves relatively high precision (0.811), its extremely low recall (0.327) suggests a failure to reliably identify the minority class of MFA adopters.

Second, these findings highlight a fundamental mismatch between the inductive biases of convolutional and recurrent networks and the tabular nature of FinScope survey data. Unlike image or sequential text data, the survey features lack inherent local spatial correlations or temporal dependencies, limiting the effectiveness of CNN and LSTM architectures. Finally, despite the higher computational cost associated with convolutional and recurrent operations, the neural models provide no improvement in discriminative ability, achieving AUC-ROC values of 0.845 (CNN) and 0.757 (LSTM), compared to 0.928 for Gradient Boosting.

These results demonstrate that neural networks do not inherently outperform classical machine learning approaches in the context of MFA adoption modeling. For low-resource financial ecosystems and behavior-driven survey data, iterative refinement of Gradient Boosting remains the most robust and computationally efficient strategy for capturing non-linear relationships.

Table 4 • Predictive performance of classical and neural models for MFA adoption.

Model	Accuracy	Precision	Recall	F1-score	AUC–ROC
Logistic Regression	0.836	0.837	0.823	0.830	0.914
Random Forest	0.822	0.825	0.791	0.808	0.915
Gradient Boosting	0.838	0.841	0.829	0.835	0.928
XGBoost	0.764	0.789	0.777	0.783	0.909
CNN	0.706	0.776	0.538	0.635	0.845
LSTM	0.643	0.811	0.327	0.466	0.757

4. Discussion

This study developed an interpretable machine learning framework to predict multi-factor authentication adoption in low-resource financial environments, utilizing the nationally representative FinScope Uganda 2023 dataset and UTAUT constructs to ground the analysis in a robust behavioral theory. Among the four evaluated models, Gradient Boosting consistently demonstrated superior performance, hence achieving an accuracy of 0.838, an F1-score of 0.835 and an AUC-ROC of 0.928, confirming its suitability for capturing complex, non-linear patterns associated with security-related behaviors. The training approach, which included repeated cross-validation and the application of SMOTE within each fold to prevent data leakage, further strengthens the reliability of the results. The comparative analysis in **Table 4** reveals that deeper architectures like CNNs and LSTMs do not inherently yield better results for behavioral survey data. Their struggle with recall suggests that the nuanced, non-sequential relationships between UTAUT constructs like trust and social influence are better captured by the iterative refinement of tree-based ensembles than by the spatial or temporal assumptions of deep learning.

SHAP analysis identified security perceptions, mobile access and social influence as the strongest predictors of MFA adoption. The prominence of security perceptions aligns with global evidence indicating that trust and perceived protection are critical determinants of security technology acceptance [40]. In low-resource financial systems undergoing rapid digitalization, heightened awareness of cyber threats appears to amplify the perceived usefulness of MFA, even under constraints of inconsistent connectivity, limited digital literacy and weak regulatory support. This reinforces UTAUT's emphasis on performance expectancy in environments where cyber risks are widely recognized but protective behaviors remain uneven.

Mobile access emerged as another decisive predictor. Given that mobile phones serve as the primary digital interface for the majority of Ugandan users, the strength of this predictor underscores the importance of optimizing authentication processes for low-bandwidth mobile environments. Usability challenges, such as slow interfaces, complex MFA steps or network-dependent authentication, may hinder adoption [10], whereas simplified, mobile-first MFA mechanisms have the potential to substantially improve uptake.

Social influence (SI), a core UTAUT construct, also significantly shaped adoption patterns. Peer behavior, community norms and endorsements from trusted figures play a central role in technology acceptance across many African contexts. The findings therefore

suggest that community-based sensitization, endorsements by local leaders or peer-led digital security initiatives may have a disproportionately positive effect compared to purely institutional campaigns.

Beyond behavioral insights, this study contributes to the growing literature on machine learning for cybersecurity behavior modeling [41]. Latent space clustering revealed four distinct user segments with differing adoption likelihoods, demonstrating the capacity of ML techniques to uncover actionable behavioral patterns that complement traditional statistical approaches. These patterns provide a granular basis for targeted policy interventions, such as prioritizing trust-building campaigns for low-trust clusters or improving infrastructure for high-intent but resource-constrained users.

From a theoretical perspective, the results validate the relevance of UTAUT in sub-Saharan Africa, particularly when integrated with security perception variables. From a computational perspective, timing analysis showed that Gradient Boosting offered a strong balance of accuracy and efficiency as compared to the evaluated classical and neural models, exhibiting faster training times than XGBoost while maintaining higher predictive performance, thus making it well-suited for deployments in resource-limited operational environments.

These findings demonstrate that targeted improvements in digital trust, mobile MFA usability and peer-driven awareness can significantly enhance security adoption. The proposed model provides policymakers, financial institutions and developers with a scalable, theory-driven and data-validated framework to identify and support users least likely to adopt MFA, thereby strengthening digital resilience within low-resource financial ecosystems.

5. Conclusions

This study evaluated multi-factor authentication adoption in low-resource financial environments using the FinScope Uganda 2023 dataset, achieving its objectives of identifying influential behavioral factors, analyzing their relationships and developing a predictive model grounded in the UTAUT framework. Gradient Boosting emerged as the most effective algorithm, achieving an accuracy of 0.843, an F1-score of 0.839, an AUC-ROC of 0.933 and an AUC-PR of 0.93, thus outperforming Logistic Regression, Random Forest and XGBoost under repeated cross-validation. It also outperformed custom-designed CNN and LSTM architectures. SHAP value analysis reinforced security perceptions as the dominant predictor (mean $|\text{SHAP}| > 2.0$, $r = 0.91$), demonstrating that trust in digital systems remains the most critical enabler of MFA adoption in underserved contexts. This underscores the

importance of selecting models that align with the structural characteristics of the data, specifically, tabular survey results, rather than defaulting to more complex neural networks that may be less effective in low-resource, high-stakes security domains. These insights further align with the emphasis on user trust and security assurance highlighted in NIST SP 800-63B.

This study's contributions are both theoretical and practical. Theoretically, it demonstrates that integrating UTAUT constructs with behavioral security indicators improves the explanatory power of adoption models in sub-Saharan Africa. Practically, it provides actionable guidance for policymakers and financial institutions seeking to expand MFA usage, specifically by enhancing trust-building communication, improving mobile-first authentication design and leveraging social networks to shape user perceptions.

However, certain limitations must be acknowledged. The cross-sectional nature of the FinScope dataset constrains causal inference and does not capture temporal dynamics in user behavior or the evolution of digital threat landscapes. As a result, observed associations may vary as regulatory environments, financial technologies and fraud patterns change over time. In addition, this study is empirically grounded in Uganda, which may limit the generalizability of the findings to other low-and middle-income countries with differing financial infrastructures, cultural attitudes toward security and levels of digital literacy.

Future research should therefore incorporate longitudinal and cross-country datasets to examine behavioral change over time and assess the robustness of MFA adoption drivers across contexts. Complementary qualitative evidence, such as user narratives on friction, digital fatigue and culturally grounded perceptions of risk could further refine the predictive models.

This study demonstrates the feasibility and value of machine learning-driven behavioral prediction for enhancing digital security in low-resource financial ecosystems. By aligning computational modeling with theoretical grounding and context-specific insights, it contributes a scalable, interpretable and policy-relevant approach to strengthening MFA adoption across digitally constrained populations.

Acknowledgments

The authors acknowledge FinScope Uganda for providing open access to the dataset used in this study.

Funding

This research received no external funding.

Author contributions

Conceptualization, R.K. (Ronald Kato), R.K. (Richard Kimera) and A.O.; methodology, R.K. (Ronald Kato), R.K. (Richard Kimera) and A.O.; formal analysis, R.K. (Ronald Kato), R.K. (Richard Kimera) and A.O.; software development, R.K. (Ronald Kato), R.K. (Richard Kimera) and A.O.; data curation, R.K. (Ronald Kato), R.K. (Richard Kimera) and A.O.; investigation,

R.K. (Ronald Kato), R.K. (Richard Kimera) and A.O.; validation, R.K. (Ronald Kato), R.K. (Richard Kimera) and A.O.; resources, R.K. (Ronald Kato), R.K. (Richard Kimera) and A.O.; writing—original draft preparation, R.K. (Ronald Kato), R.K. (Richard Kimera) and A.O.; writing—review and editing, R.K. (Ronald Kato), R.K. (Richard Kimera) and A.O.; visualization, R.K. (Ronald Kato), R.K. (Richard Kimera) and A.O.; supervision, R.K. (Ronald Kato), R.K. (Richard Kimera) and A.O.; project administration, R.K. (Ronald Kato), R.K. (Richard Kimera) and A.O. All authors have read and agreed to the published version of the manuscript.

Conflict of interest

The authors declare that they have no competing interests.

Data availability statement

The data supporting the findings of this publication has been made available within a publicly accessible repository at <https://fsduganda.or.ug/data-sets-finscope-uganda-2023-survey-report/>.

Institutional review board statement

The study was approved by the Research Ethics Committee of Mbarara University of Science and Technology (protocol code MUST-2024-1964; approved on 28 May 2025).

Informed consent statement

Informed consent for participation was obtained from all subjects involved in the study.

Additional information

Received: 31 October 2025

Accepted: 9 January 2026

Published: 28 January 2026

Academia AI and Applications papers should be cited as *Academia AI and Applications* 2026, ISSN pending, <https://doi.org/10.20935/AcadAI8107>. The journal's official abbreviation is *Acad. AI*.

Publisher's note

Academia.edu Journals stays neutral with regard to jurisdictional claims in published maps and institutional affiliations. All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright

© 2026 copyright by the authors. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

References

1. Sule M, Zennaro M. Cybersecurity through the lens of digital identity and data protection: issues and trends. *Telecommun Policy*. 2021;67:101734. doi: 10.1016/j.techsoc.2021.101734
2. Agbeyangi A, Lukose JM. Telemedicine adoption and prospects in sub-saharan africa: a systematic review with a focus on south africa, kenya, and nigeria. *Healthcare*. 2025;13(7):762. doi: 10.3390/healthcare13070762
3. Sharif MHU, Mohammed MA. A literature review of financial losses statistics for cyber security and future trend. *World J Adv Res Rev*. 2022;15(1):138–56. doi: 10.30574/wjarr.2022.15.1.0573
4. Suleski T, Ahmed M, Yang W, Wang E. A review of multi-factor authentication in the internet of healthcare things. *Digit Health*. 2023;9:20552076231177144. doi: 10.1177/20552076231177144
5. Museba TJ, Ranganai E, Gianfrate G. Customer perception of adoption and use of digital financial services and mobile money services in Uganda. *J Enterprising Communities*. 2021;15(2):177–203. doi: 10.1108/JEC-07-2020-0127
6. Ali G, Dida MA, Sam AE. Evaluation of key security issues associated with mobile money systems in uganda. *Information*. 2020;11(6):309. doi: 10.3390/info11060309
7. Lee E, Kim S, Kim S, Jung S, Kim H, Cha M. Explainable product classification for customs. *ACM Trans Intell Syst*. 2024;15(2):1–24. doi: 10.1145/3635158
8. Gyamfi N. Bank fraud detection using support vector machine. *Proceedings of the 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*; 1–3 November 2018; Vancouver, BC, Canada. 2018.
9. Rahman NA. Classification of customer feedbacks using sentiment analysis towards mobile banking applications. *Int J Inf Syst*. 2022;11(4):1579–87. doi: 10.11591/ijai.v11.i4.pp1579-1587
10. Akello M, Gallagher M, Nanyunja S, Mulondo A, Miranda JJ, Cole G, et al. Minimal computing in refugee education in uganda: economies of digital use and non-use, and the right constraints. *Learn Media Technol*. 2024;49(5):811–26. doi: 10.1080/17439884.2024.2405128
11. Sagar R, Jhaveri R. Applications in security and evasions in machine learning: a survey. *Electronics*. 2020;9(1):97. doi: 10.3390/electronics9010097
12. Gupta A, Singh A. Agri-gnn: a novel genotypic-topological graph neural network framework built on graphsage for optimized yield prediction. *arXiv*. 2023. arXiv:2310.13037. doi: 10.48550/arXiv.2310.13037
13. Yang R, Zheng K, Wu B, Li D, Wang Z, Wang X. Predicting user susceptibility to phishing based on multidimensional features. *Secur Commun Netw*. 2022;2022(1):7058972. doi: 10.1155/2022/7058972
14. Shaikh TA, Rasool T, Verma P, Mir WA. A fundamental overview of ensemble deep learning models and applications: systematic literature and state of the art. *Ann Oper Res*. 2024;1–77. doi: 10.1007/s10479-024-06444-0
15. Akinnuwesi B, Uzoka FME, Fashoto SG, Mbunge E, Odumabo A, Amusa OO, et al. A modified utaut model for the acceptance and use of digital technology for tackling COVID-19. *Sustain Oper Comput*. 2022;3:118–35. doi: 10.1016/j.susoc.2021.12.001
16. Karama AA. Effects of slow adaptation of high-tech security solutions to the social and economic development of Uganda. 2024 [accessed on 15 May 2024]. Available from: <https://lup.lub.lu.se/student-papers/search/publication/9175389>
17. Simione FF, Muehlschlegel T. Mobile money, perception about cash, and financial inclusion: learning from Uganda's micro-level data. Berlin/Heidelberg: Springer; 2023.
18. Salman HA, Kalakech A, Steiti A. Random forest algorithm overview. *Babylon J Mach Learn*. 2024;2024:69–79. doi: 10.58496/BJML/2024/007
19. Pandya HB, Zalawadia K. A comprehensive review of phishing detection techniques based on machine learning. *Metall Mater Eng*. 2025;31(5):302–10.
20. Preti LM, Ardito V, Compagni A, Petracca F, Cappellaro G. Implementation of machine learning applications in health care organizations: systematic review of empirical studies. *J Med Internet Res*. 2024;26:e55897. doi: 10.2196/55897
21. Grassi P.A., Newton E.M., Perlner R.A., Regenscheid A.R., Burr W.E., Richer J.P. NIST SP 800-63B. Digital identity guidelines—authentication and lifecycle management. Gaithersburg: National Institute of Standards and Technology; 2017.
22. FinScope Uganda 2023 survey dataset. 2024 [accessed on 6 December 2025]. Available from: <https://fsduganda.or.ug/data-sets-finscope-uganda-2023-survey-report/>
23. OpenStreetMap contributors. OpenStreetMap. 2025 [accessed on 1 June 2025]. Available from: <https://www.openstreetmap.org>
24. Mujahid M, Kina EROL, Rustam F, Villar MG, Alvarado ES, Diez IDLT, et al. Data oversampling and imbalanced datasets: an investigation of performance for machine learning and feature engineering. *J Big Data*. 2024;11(1):87. doi: 10.1186/s40537-024-00943-4
25. Zhang Y, Zhang X, Razbek J, Li D, Xia W, Bao L, et al. Opening the black box: interpretable machine learning for predictor finding of metabolic syndrome. *BMC Endocr Disord*. 2022;22(1):214. doi: 10.1186/s12902-022-01121-4

26. Zhou K, Zhang J, Ren Y, Huang Z, Zhao L. A gradient boosting decision tree algorithm combining synthetic minority oversampling technique for lithology identification. *Geophysics*. 2020;85(4):WA147–58. doi: 10.1190/geo2019-0429.1
27. Nguyen H, Cao MT, Tran XL, Tran TH, Hoang ND. A novel whale optimization algorithm optimized xgboost regression for estimating bearing capacity of concrete piles. *Neural Comput Appl*. 2023;35(5):3825–52. doi: 10.1007/s00521-022-07896-w
28. Pan S, Zheng Z, Guo Z. An optimized xgboost method for predicting reservoir porosity using petrophysical logs. *J Pet Sci*. 2022;208:109520. doi: 10.1016/j.petrol.2021.109520
29. Anusha Y, Visalakshi R, Srinivas K. Imbalanced data classification using improved synthetic minority over-sampling technique. *Multiagent Grid Syst*. 2023;19(2):117–31. doi: 10.3233/MGS-230007
30. Wah YB, Ismail A, Azid N, Niswah N, Jaafar J, Aziz IA, et al. Machine learning and synthetic minority oversampling techniques for imbalanced data: improving machine failure prediction. *Comput Mater Contin*. 2023;75(3):4821–41. doi: 10.32604/cmc.2023.034470
31. Cheng L, Vos JD, Zhao P, Yang M, Witlox F. Examining non-linear built environment effects on elderly's walking: a random forest approach. *Transp Res Part D*. 2020;88:102552. doi: 10.1016/j.trd.2020.102552
32. Oguz BU, Shinohara RT, Yushkevich PA, Oguz I. Gradient boosted trees for corrective learning. In: *International workshop on machine learning in medical imaging*. Cham: Springer; 2017.
33. Mitchell R, Frank E. Accelerating the xgboost algorithm using gpu computing. *PeerJ Comput Sci*. 2017;3:e127. doi: 10.7717/peerj-cs.127
34. Hairani H, Widiyaningtyas T, Prasetya DD. Addressing class imbalance of health data: a systematic literature review on modified synthetic minority oversampling technique (SMOTE) strategies. *JOIV Int J Inform Vis*. 2024;8(3):1310–8. doi: 10.62527/joiv.8.3.2283
35. Galli S. *Python feature engineering cookbook*. Birmingham: Packt Publishing Ltd.; 2024.
36. Van Biljon J. Machine learning in Sub-Saharan Africa: a critical review of selected research publications, 2010–2021. In: *International conference on social implications of computers in developing countries*. Cham: Springer; 2022. p. 363–76.
37. Sodeeq MA. Machine learning and economic transformation in Africa: a decade of evidence and innovation. *Int J Appl Knowl Innov*. 2025;2(1):82–98. ISSN 3092-8982. [accessed on 9 December 2025]. Available from: <http://ijakjournal.com/wp-content/uploads/2025/09/8-Monsuru-Machine-Learning-and-Economic-Transformation-in-Africa.pdf>
38. Paradigm Initiative. Artificial intelligence in Sub-Saharan Africa: ensuring inclusivity. 2022 [accessed on 6 December 2025]. Available from: <https://paradigmhq.org/wp-content/uploads/2022/05/AI-Research-Report.pdf>
39. Greenacre M, Groenen PJF, Hastie T, d'Enza AI, Markos A, Tuzhilina E. Principal component analysis. *Nat Rev Methods Prim*. 2022;2(1):100. doi: 10.1038/s43586-022-00184-w
40. Niyikora E. Technology acceptance and digital readiness in underserved small business sectors. *Multidiscip Front Res*. 2025;4(1):252–68. doi: 10.54660/IJFMR.2023.4.1.252-268
41. Mohammed AJ, Hassan MM, Kadir DH. Improving classification performance for a novel imbalanced medical dataset using smote method. *Int J Adv Trends Comput Sci*. 2020;9(3):3161–72. doi: 10.30534/ijatcse/2020/104932020