International Journal of Data and Network Science 9 (2025) 411–418

Contents lists available at GrowingScience

International Journal of Data and Network Science

homepage: www.GrowingScience.com/ijds

Predictive models based on machine learning to analyze the adoption of digital payments in Latin America and the Caribbean

Jiang Wagner Mamani Lopez[a*], Antonio Víctor Morales Gonzales[a] and Pedro Pablo Chambi Condori[b]

[a]*Universidad Nacional de San Agustín de Arequipa, Peru*
[b]*Universidad Nacional Jorge Basadre Grohmann, Peru*

C H R O N I C L E

Article history:
Received January 22, 2025
Received in revised format February 25, 2025
Accepted March 24 2025
Available online March 24 2025

Keywords:
Digital payments
Financial innovation
Data mining
Bayesian optimization
Hyperparameter Tuning

A B S T R A C T

The use of technology in the financial industry has experienced sustained growth in recent years. However, in many emerging economies, a significant proportion of the population still does not utilize digital solutions for financial transactions. Promoting financial inclusion through digital environments is essential for driving social and economic development. This study aims to develop machine learning models to predict the adoption of digital payments in Latin America and the Caribbean using statistical data from the World Bank's Global Findex Database for 2021. The performance of the Random Forest, LightGBM, XGBoost, and CatBoost algorithms was compared, with the optimal hyperparameter combination identified through Bayesian optimization. The results show that LightGBM achieved the highest performance in predicting digital payments, with an F1-score of 90.25% and a more stable balance between precision and recall compared to the other models. These findings highlight the value of machine learning models in the financial sector, as they enable a more accurate identification of users adopting digital solutions, facilitating the design of strategies to strengthen financial inclusion in the region.

1. Introduction

The development of the financial sector is a key driver of economic growth in emerging economies, strengthening market confidence and expanding access to new financial services (Mollaahmetoğlu & Akcąli, 2019; Sahay et al., 2015). Technological innovation plays a crucial role in this process by facilitating financial inclusion for underserved sectors through efficient and accessible solutions (Bara & Mudzingiri, 2016; Chou, 2007; Qamruzzaman & Wei, 2018). Consequently, financial institutions are offering improved products, services, and innovative processes for the benefit of their users (Qamruzzaman & Jianguo, 2017).

Among the most significant advancements in financial innovation are digital payment systems, designed to simplify transactions through credit and debit cards, online operations, ATMs, and mobile devices (Agarwal et al., 2020; Maherali, 2017). In particular, mobile payments have great potential to reduce barriers to financial services, especially in developing economies, where a considerable portion of the population remains excluded from the formal banking system (Patil et al., 2017).

According to World Bank estimates in Demirgüç-Kunt et al. (2022) approximately 64% of adults globally (aged 15 and older) made or received at least one digital payment in the past year. This indicator varies widely between high-income and developing economies, where the percentage of adults reporting the use of digital payments reaches 95% and 57%, respectively. In Latin America and the Caribbean (LAC), the proportion of adults utilizing digital payments increased from 44% in 2014 to 66% in 2021, signaling a 22-percentage point growth. However, these figures remain lower compared to regions such as Asia and Europe, where financial ecosystems demonstrate a higher degree of digitalization.

* Corresponding author
E-mail address jmamanilo@unsa.edu.pe (J. W. Mamani Lopez)

ISSN 2561-8156 (Online) - ISSN 2561-8148 (Print)

doi: 10.5267/j.ijdns.2025.3.001

The LAC context highlights the opportunity to conduct a study employing machine learning techniques to forecast digital payment usage. Despite the growing interest in this topic, there is a limited number of studies in emerging economies that assess the performance of different algorithms in predicting the adoption of these technologies in financial transactions.

This study aims to develop and evaluate machine learning models to predict the adoption of digital payments in the LAC region. This research contributes to the financial sector, particularly to studies on digital financial inclusion, while also adding to the body of quantitative work that compares the performance of machine learning algorithms in forecasting the use of digital payment methods. These insights can later be integrated into recommendation systems or applications that facilitate financial transactions.

Furthermore, this study highlights the positive impact of innovation in the financial industry, as the use of digital channels reduces operational costs, time, and effort for users (Al-Okaily et al., 2020; Hjelholt & Damsgaard, 2012). Finally, understanding the factors influencing digital payment adoption provides valuable insights for public policy formulation, as empirical evidence supports its positive effects on tax collection and transparency, contributing to the reduction of corruption (Klapper & Singer, 2017; Maherali, 2017; Setor et al., 2021).

The remainder of this paper is organized as follows: Section 2 presents a review of the literature on digital payments. Section 3 describes the methodology used to address the study's objective. Section 4 presents the results of the machine learning models. Section 5 discusses the findings, and finally, Section 6 outlines the study's conclusions.

2. Literature Review

The adoption of digital payments has been analyzed in the literature as a key component of financial inclusion, employing methodologies that range from traditional statistical approaches to machine learning techniques. Various studies have identified sociodemographic, economic, and technological factors as determinants of financial inclusion and digital payment adoption.

At the international level, the research conducted by Lee et al. (2023) in Malaysia applied machine learning models to predict financial inclusion based on socioeconomic factors and attributes related to digital financial services. Their findings revealed that Random Forest had the highest predictive capacity, where the main variables driving financial inclusion were income level, education, the use of digital payments, and online purchases. In the African context, Ismail et al. (2021) found that XGBoost outperformed other models in forecasting financial inclusion probabilities, showing that access to banking infrastructure and income level play a crucial role in the adoption of digital financial service.

Rakipi et al. (2023) studied the application of data mining techniques in digital and electronic payment adoption within Albania's banking sector through a mixed-methods approach. Their findings indicate that although these tools have gained relevance, their implementation in financial institutions remains moderate, yet they hold potential for enhancing the sector's operational efficiency. In Jordan, Obidat et al. (2022) investigated the continued use of mobile wallets through structural equation modeling. Their results indicate that perceived usefulness and ease of use positively influenced users' intention to adopt them on an ongoing basis. Additionally, trust, security, and seamless connectivity emerged as key factors in user retention.

Regarding the impact of tax incentives on digital payment adoption, Allen et al. (2022) examined this effect through the causal forest method combined with machine learning algorithms to identify the main predictors in micro, small, and medium-sized enterprises across countries with different levels of development. The authors highlight that the availability of mobile payment applications and the digitization of transactions—particularly for employee and supplier payments—enhance the likelihood of adoption in businesses. Moreover, tax incentives directed at consumers and merchants contributed to the expansion of point-of-sale terminals in emerging markets.

In the LAC context, financial inclusion and digital payment adoption have been explored through various approaches. In Peru, Maehara et al. (2024) applied machine learning algorithms to analyze the determinants of financial inclusion, identifying educational level and access to banking services as significant factors influencing account ownership and usage. Aurazo & Vega (2021), employing a Heckman probit model, concluded that education and formal employment shape the propensity to conduct electronic transactions.

Despite advancements in modern predictive models, studies analyzing digital payment adoption in LAC through a machine learning approach remain limited. Most existing research focuses on financial inclusion without specifically distinguishing the use of digital services or assessing the performance of different predictive techniques in their estimation. In this regard, this study aims to bridge this gap by comparing machine learning models using various performance metrics to forecast the integration of these payment methods into the regional market.

3. Methodology

The following section outlines the stages of data collection, preprocessing, modeling, and model performance evaluation. As shown in Fig. 1, the process begins with the Extract, Transform, and Load (ETL) phase, during which data was collected from

The Global Findex Database 2021, published by the World Bank. Next, in the Exploratory Data Analysis (EDA) phase, variables are described, and missing values are identified. In the preprocessing stage, continuous variables are normalized, and categorical variables are encoded for integration into the models.

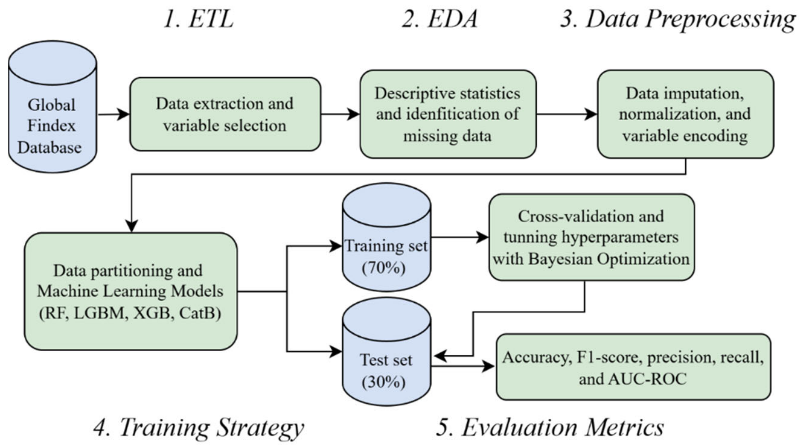

Fig. 1. Methodology diagram

To implement these stages, the Python programming language was used, leveraging fundamental libraries for data management and exploration, including *scikit-learn* (Pedregosa et al., 2011), *pandas* (McKinney, 2010), and *numpy* (Harris et al., 2020). During the modeling phase, four classification algorithms were trained using machine learning libraries: Random Forest, LightGBM (Ke et al., 2017), XGBoost (Chen & Guestrin, 2016), and CatBoost (Prokhorenkova et al., 2018).

Hyperparameter tuning was conducted using a Bayesian optimization approach developed by Gardner et al. (2014) to enhance model performance. Finally, model evaluation was carried out using metrics such as *accuracy*, *f1-score*, *precision*, *recall*, and *area under the receiver operating characteristic curve* (AUC-ROC), allowing the identification of the model with the highest predictive capability for digital payment adoption.

Extract, transform and load (ETL)

Data was obtained from the Global Findex Database, developed by the Development Research Group, Finance Unit and Private Sector Development (2022), which provides information on financial inclusion across 139 countries. This dataset collects insights on financial service access, usage, and barriers, facilitating the assessment of patterns and trends in digital payment adoption.

For this study, the unit of analysis consists of individuals aged 15 and older across 19 LAC countries. Relevant variables influencing digital payment adoption were selected, including sociodemographic and financial characteristics. Table 1 presents the variables used in the model, detailing their type, description, and possible values.

Table 1
Variables used in the digital payment adoption model

Variable	Type	Description	Factor levels or range
Digital payments	Binary	Made or received a digital payment	0=No, 1=Yes
Gender	Binary	Respondent's gender	1=Male, 0=Female
Age	Continuous	Respondent's age	15-99
Education level	Ordinal	Highest level of education attained	1=Primary school or less, 2=Secondary school, 3=Tertiary education or more
Income quintile	Ordinal	Household income quintile	1=Q1, 2=Q2, 3=Q3, 4=Q4, 5=Q5
Work	Binary	Employment status	0=No, 1=Yes
Urban	Binary	Whether the respondent lives in an urban area	0=Rural, 1=Urban
Population	Continuous	Adult (15+) population	2,269,489-168,540,064.
Account at a financial institution	Binary	Respondent has an account at a bank or at another type of financial institution	0=No, 1=Yes
Saved	Binary	Respondent personally saved or set aside money in the past year	0=No, 1=Yes
Borrowed	Binary	Respondent borrowed money in the past year	0=No, 1=Yes
Mobile phone	Binary	Respondent owns a mobile phone	0=No, 1=Yes
Internet Access	Binary	Respondent has access to the Internet	0=No, 1=Yes
Region	Category	Geographical region of the respondent's country	1 = South America, 2 = Central America, 3 = Caribbean, 4 = North America

Exploratory Data Analysis (EDA)

Table 2 displays descriptive statistics for the study variables, including the number of observations, mean, standard deviation, minimum, and maximum values. Key descriptive statistics reveal that 54% of respondents in LAC have made or received a digital payment. Regarding gender distribution, data indicates that 42% of respondents are women and 58% are men, suggesting a slight overrepresentation of males in the sample. The average respondent age is 41 years, with a range spanning from 15 to 99 years, enabling an analysis of generational differences in digital payment usage.

Table 2
Descriptive statistics

Variables	Observations	Mean	Standard Deviation	Min	Max
Digital payments	18519	0.54	0.50	0	1
Gender	18519	0.42	0.49	0	1
Age	18498	41.27	17.31	15	99
Education level	18334	1.87	0.67	1	3
Income quintile	18519	3.24	1.42	1	5
Work	18519	0.74	0.44	0	1
Urban	11513	0.80	0.40	0	1
Population	18519	25489919.70	40368029.49	2269489	168540064
Account at a financial institution	18519	0.59	0.49	0	1
Saved	18519	0.42	0.49	0	1
Borrowed	18519	0.44	0.50	0	1
Mobile phone	18472	0.89	0.32	0	1
Internet Access	18457	0.73	0.44	0	1
Region	18519	1.65	0.85	1	4

Educational attainment is primarily distributed among primary, secondary, and higher education levels, with a higher proportion of individuals having completed at least secondary education. In terms of income, the mean income quintile suggests that most respondents belong to middle-income strata based on economic distribution.

Regarding financial inclusion, 59% of respondents hold an account at a financial institution, while 42% reported having saved in the past year, reflecting relevant financial behavior patterns for the study. Additionally, technology access is significantly high, with 89% of respondents owning a mobile phone and 73% having internet access, indicating a favorable environment for digital payment expansion in the region.

Data Preprocessing

Before training the models, various preprocessing techniques were applied to enhance data quality. For the variable *age*, median imputation was performed, preserving the original distribution without skewing extreme values. For categorical variables such as *education level*, *urban*, *mobile phone* e *internet access*, mode imputation was applied, ensuring that the most frequent values represented missing data coherently with the dataset structure. Additionally, continuous variables such as *age* and *population* were normalized to ensure a uniform scale and prevent models from being influenced by differences in magnitude. Conversely, categorical variables such as *education level*, *income quintile,* and *region* were transformed using one-hot encoding.

Training Strategy

To ensure a reliable evaluation of the models, the dataset was split into 70% for training and 30% for testing, maintaining adequate representation of the target variable in both subsets. Given that the study aims to predict a binary variable, the *f1-score* was selected as the primary metric for model selection. This metric provides a balance between *precision* and *recall*, allowing for performance assessment while considering both false positives and false negatives in a balanced manner.

Hyperparameter selection was performed using Bayesian optimization with 10-fold cross-validation. The implementation utilized the *BayesianOptimization* function developed by Gardner et al. (2014). This method efficiently explores the search space by modeling the relationship between hyperparameter values and model performance. Through iterative adjustments, optimal configurations maximizing predictive capability were identified. Table 3 presents the tuned hyperparameters along with their respective search ranges for each model.

Following the Bayesian optimization process, each model was trained using the best hyperparameters found. The final evaluation was performed on the independent test set using Accuracy, F1-score, Precision, Recall, and AUC-ROC to ensure a comprehensive assessment of model performance.

Table 3
Hyperparameter search space for Bayesian optimization

Model	Hyperparameter	Range	Definition
Random Forest	max depht	(3, 30)	Maximum depth of each decision tree. Controls model complexity and risk of overfitting
	max features	(0.1, 0.999)	Fraction of total features randomly selected for node splitting. Influences diversity in trees
	min samples split	(3, 25)	Minimum number of samples required to split an internal node. Higher values reduce overfitting
LightGBM	num leaves	(20, 100)	Maximum number of leaves per tree. A higher value increases model complexity
	max depth	(3, 15)	Maximum depth of a tree, limiting its growth to prevent overfitting
	feature fraction	(0.7, 0.9)	Fraction of features randomly chosen for each tree construction
	bagging fraction	(0.7, 0.9)	Fraction of data randomly selected for training each iteration, improving generalization
	min split gain	(0.001, 0.5)	Minimum loss reduction required for a node split. Encourages meaningful splits
	min child weight	(5, 50)	Minimum sum of instance weights required in a child node. Helps control overfitting
XGBoost	max depth	(3, 15)	Maximum depth of each tree, affecting model complexity and overfitting risk
	gamma	(0, 3)	Minimum loss reduction required to split a node. Higher values prevent unnecessary splits
	colsamples bytree	(0.5, 1)	Fraction of features randomly selected for constructing each tree
	subsample	(0.2, 0.8)	Fraction of training data randomly sampled before tree growth, reducing overfitting
CatBoost	depth	(4, 10)	Maximum depth of each tree, influencing model complexity and capacity
	l2 leaf reg	(1, 10)	L2 regularization coefficient for leaf values, controlling complexity and reducing overfitting
	bagging temperature	(0, 1)	Controls the level of randomness in data subsampling. Higher values increase diversity
	border count	(32, 255)	Number of splits used to convert numerical features into categorical bins

Evaluation metrics

To measure model performance, commonly used metrics in binary classification problems were employed, described as follows:

- *Accuracy*: represents the proportion of correctly predicted observations relative to the total evaluated cases.
- *F1-score*: represents the harmonic average of precision and recall, serving as a comprehensive metric to assess a model's classification performance by balancing false positives and false negatives.
- *Precision*: indicates the percentage of positive predictions correctly classified, which is particularly relevant in contexts where minimizing false positives is a priority.
- *Recall*: quantifies the proportion of true positive instances that were correctly classified out of all actual positive cases, indicating the model's sensitivity in identifying relevant observations.
- *AUC-ROC*: measures the model's capability to differentiate between classes by evaluating the trade-off between the true positive rate and the false positive rate across various classification thresholds.

4. Results

Table 4 and Figure 2 present the results of the models evaluated on the test dataset. LightGBM achieved the best overall performance, obtaining the highest *f1-score*, which reflects an appropriate balance between precision and recall in classifying digital payment adoption in LAC. Additionally, it recorded the highest *accuracy*, *recall*, and *AUC-ROC*, suggesting a superior ability to correctly identify both users who adopt digital payments and those who do not.

Table 4
Performance comparison of models on the test set

Model	Accuracy	F1-score	Precision	Recall	AUC-ROC
Random Forest	88.34	89.62	87.14	92.25	93.81
LightGBM	**88.91**	**90.25**	86.79	**94.00**	**94.58**
XGBoost	87.87	89.08	**87.58**	90.64	94.00
CatBoost	87.98	89.29	86.87	91.86	94.07

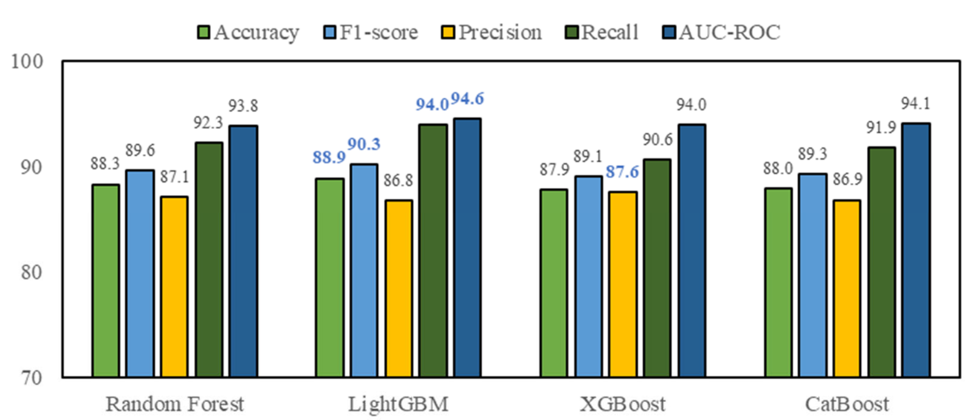

Fig. 2. Comparative evaluation of classification models

Although XGBoost exhibited the highest precision, its recall was lower than that of LightGBM, indicating that it identified fewer positive cases in comparison. Random Forest demonstrated competitive performance, with a solid balance between recall and precision, while CatBoost produced results similar to XGBoost, with close values across all metrics.

These findings underscore the importance of evaluating multiple metrics, as a model with high *precision* does not guarantee optimal classification if its *recall* is low. Since LightGBM achieves the best balance among these metrics while also presenting the highest *AUC-ROC*, it is selected as the most suitable model for the classification task in this study.

Regarding Figure 3, the normalized confusion matrix illustrates the performance of the LightGBM model in classifying digital payment adoption. The results show that 94.00% of users were correctly identified, as well as 82.80% of non-users. However, 17.20% of individuals who did not adopt this payment method were misclassified as users, while 6.00% of positive cases were incorrectly assigned to the category of those who neither made nor received a digital payment.

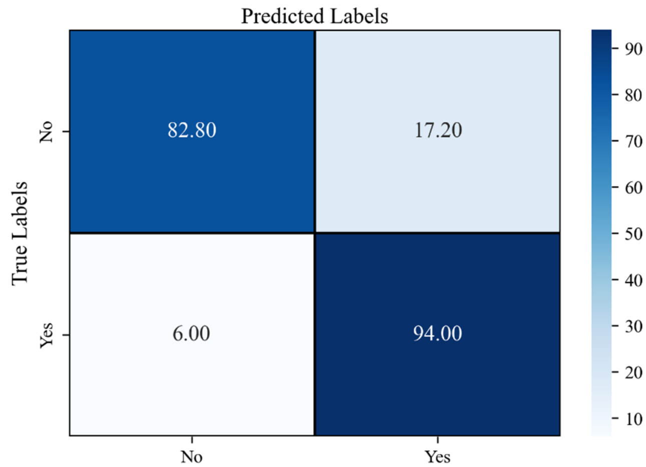

Fig. 3. Normalized Confusion Matrix (%)

4. Discussion

The results of this study highlight the potential of machine learning models in predicting digital payment adoption in LAC. The comparative evaluation of different algorithms revealed that LightGBM achieved the best overall performance due to its balance between precision and recall. Balancing these metrics is essential for research on digital payment adoption, as accurately identifying both users and non-users enables the development of effective strategies to promote financial inclusion.

The application of Bayesian optimization for hyperparameter selection was crucial in enhancing model performance. Automating the search for optimal combinations improved predictive capacity while avoiding manual adjustments or default configurations. Additionally, the findings demonstrate the importance of achieving a proper balance between precision and recall when selecting a classification model. While some algorithms achieved high precision, this did not necessarily translate into optimal identification of positive cases, which could limit their effectiveness in strategies aimed at fostering digital payment adoption. Evaluating multiple metrics provides a comprehensive view of the model's generalization capability, helping to mitigate potential classification biases.

In this regard, implementing models like LightGBM in studies on financial technology adoption can offer valuable tools for designing strategies targeted at banks, fintech companies, and regulatory entities. The model's ability to identify patterns associated with digital payment usage creates opportunities to strengthen financial education campaigns, improve resource allocation, and expand the reach of these solutions in populations with limited access to digital financial services.

5. Conclusion

This study evaluated the performance of various machine learning models in predicting digital payment adoption in LAC. Through a comparative analysis based on multiple metrics, LightGBM demonstrated the best performance by offering an optimal balance between precision and recall, along with superior capability in classifying digital payment users and non-users. The results underscore the importance of considering multiple evaluation metrics when selecting a classification model, as prioritizing precision alone may limit the identification of positive cases. The selection of LightGBM as the optimal model highlights its potential for applications aimed at predicting digital payment adoption, which may be valuable for financial institutions, technology companies, and policymakers seeking to enhance financial inclusion in the region.

Despite these contributions, this study is limited to machine learning models, highlighting the need for further exploration of more advanced architectures. Future research could examine the effectiveness of deep learning approaches, such as neural networks, which may capture more intricate patterns in digital payment adoption. Additionally, integrating explainable AI techniques—such as Shapley Value Analysis (SHAP), Local Interpretable Model-agnostic Explanations (LIME), and Feature Importance methods—could provide deeper insights into the key factors driving digital payment adoption, offering a clearer perspective on the most influential variables in decision-making.

Addressing these aspects will enhance understanding of digital payment adoption in LAC and contribute to the development of innovative strategies for promoting digital financial inclusion in emerging economies.

References

Agarwal, S., Ghosh, P., Li, J., & Ruan, T. (2020). Digital Payments and Consumption: Evidence from the 2016 Demonetization in India. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3641508

Allen, J., Carbo-Valverde, S., Chakravorti, S., Rodriguez-Fernandez, F., & Pinar Ardic, O. (2022). Assessing incentives to increase digital payment acceptance and usage: A machine learning approach. *PLoS ONE*, 17(11 November). https://doi.org/10.1371/journal.pone.0276203

Al-Okaily, M., Lutfi, A., Alsaad, A., Taamneh, A., & Alsyouf, A. (2020). The Determinants of Digital Payment Systems' Acceptance under Cultural Orientation Differences: The Case of Uncertainty Avoidance. *Technology in Society*, 63(101367). https://doi.org/10.1016/j.techsoc.2020.101367

Aurazo, J., & Vega, M. (2021). Why people use digital payments: Evidence from micro data in Peru. *Latin American Journal of Central Banking*, 2(4). https://doi.org/10.1016/j.latcb.2021.100044

Bara, A., & Mudzingiri, C. (2016). Financial innovation and economic growth: Evidence from Zimbabwe. *Investment Management and Financial Innovations*, 13(2), 65–75. https://doi.org/10.21511/imfi.13(2).2016.07

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 13-17-August-2016. https://doi.org/10.1145/2939672.2939785

Chou, Y. K. (2007). Modeling financial innovation and economic growth: Why the financial sector matters to the real economy. *Journal of Economic Education*, 38(1), 78–90. https://doi.org/10.3200/JECE.38.1.78-91

Demirgüç-Kunt, A., Klapper, L., Singer, D., & Ansar, S. (2022). *The Global Findex Database 2021: Financial Inclusion, Digital Payments, and Resilience in the Age of COVID-19*. World Bank. https://doi.org/10.1596/978-1-4648-1897-4

Development Research Group, Finance Unit and Private Sector Development. (2022). *Global Financial Inclusion (Global Findex) Database 2021*. World Bank, Development Data Group. https://doi.org/10.48529/JQ97-AJ70

Gardner, J. R., Kusner, M. J., Xu, Z. E., Weinberger, K. Q., & Cunningham, J. P. (2014). Bayesian optimization with inequality constraints. *ICML, 2014*, 937–945.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, 585(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Hjelholt, M., & Damsgaard, J. (2012). The genesis and evolution of digital payment platforms - A framework for understanding payment transition routes. *ECIS 2012*.

Ismail, Q. F., Al-Sobh, E. S., Al-Omari, S. S., Bani Yaseen, T. M., & Abdullah, M. A. (2021). Using Machine Learning Algorithms to Predict the State of Financial Inclusion in Africa. *12th International Conference on Information and Communication Systems, ICICS 2021*. https://doi.org/10.1109/ICICS52457.2021.9464590

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T. Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems, 2017-December*.

Klapper, L., & Singer, D. (2017). The opportunities and challenges of digitizing government-to- person payments. *World Bank Research Observer*, 32(2), 211–226. https://doi.org/10.1093/wbro/lkx003

Lee, H. S., Kuang, K. S., Liew, P. X., Har, W. M., & Low, C. W. (2023). Supporting Financial Inclusion with Digital Financial Services: Comparative Analysis of Machine Learning Models. *5th IEEE International Conference on Artificial Intelligence in Engineering and Technology, IICAIET 2023*. https://doi.org/10.1109/IICAIET59451.2023.10291296

Maehara, R., Benites, L., Talavera, A., Aybar-Flores, A., & Muñoz, M. (2024). Predicting Financial Inclusion in Peru: Application of Machine Learning Algorithms. *Journal of Risk and Financial Management*, 17(1). https://doi.org/10.3390/jrfm17010034

Maherali, A. (2017). *Financial Inclusion, Digital Payments and Their Impact on Income and Tax Revenue Around the World* [Master's thesis, Harvard University]. https://dash.harvard.edu/handle/1/33826588

McKinney, W. (2010). Data Structures for Statistical Computing in Python. *Proceedings of the 9th Python in Science Conference*, 56–61. https://doi.org/10.25080/Majora-92bf1922-00a

Mollaahmetoğlu, E., & Akcąli, B. Y. (2019). The missing-link between financial development and economic growth: financial innovation. *Procedia Computer Science*, 158, 696–704. https://doi.org/10.1016/j.procs.2019.09.105

Obidat, A., Almahameed, M., & Alalwan, M. (2022). An empirical examination of factors affecting the post-adoption stage of mobile wallets by consumers: A perspective from a developing country. *Decision Science Letters*, 11(3). https://doi.org/10.5267/j.dsl.2022.2.005

Patil, P. P., Dwivedi, Y. K., & Rana, N. P. (2017). Digital payments adoption: An analysis of literature. In Springer (Ed.), *Lecture Notes in Computer Science: Vol. 10595 LNCS* (pp. 61–70). https://doi.org/10.1007/978-3-319-68557-1_7

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn:

Machine learning in Python. *Journal of Machine Learning Research*, 12(85), 2825–2830. https://www.jmlr.org/papers/v12/pedregosa11a.html
Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). Catboost: Unbiased boosting with categorical features. *Advances in Neural Information Processing Systems, 2018-December*.
Qamruzzaman, M., & Jianguo, W. (2017). Financial innovation and economic growth in Bangladesh. *Financial Innovation*, 3(1). https://doi.org/10.1186/s40854-017-0070-0
Qamruzzaman, M., & Wei, J. (2018). Financial innovation, stock market development, and economic growth: An application of ARDL model. *International Journal of Financial Studies*, 6(3). https://doi.org/10.3390/ijfs6030069
Rakipi, A., Shurdi, O., & Imami, J. (2023). Utilization of data mining and machine learning in digital and electronic payments in banks. *Corporate and Business Strategy Review*, 4(4 Special Issue). https://doi.org/10.22495/cbsrv4i4siart5
Sahay, R., Cihák, M., N'Diaye, P., & Barajas, A. (2015). Rethinking financial deepening: stability and growth in emerging markets. *Revista de Economia Institucional*, 17(33), 73–108. https://doi.org/10.18601/01245996.v17n33.04
Setor, T. K., Senyo, P. K., & Addo, A. (2021). Do digital payment transactions reduce corruption? Evidence from developing countries. *Telematics and Informatics*, 60(101577). https://doi.org/10.1016/j.tele.2021.101577