African Journal of Education Science and Technology (AJEST) Vol. 9 No. 1 (October, 2026)
University of Eldoret, Kenya, Mount Kenya University, Kenya, Chukwuemeka Odumegwu Ojukwu University, Nigeria, Kyambogo University, Uganda and University of Makeni, Sierra Leone

Machine Learning Approaches to Credit Risk Prediction in Mobile Lending Platforms in Kenya: A Rapid Scoping Review

Kelvin Kipchumba Koech[1], Charles Katila[1] & Alex Kibet[2]
[1]Department of Computer Science, School of Mathematics and Computing
The Co-operative University of Kenya
[2]Department of Computing and Informatics, School of Science and Applied Technology
Laikipia University

Corresponding E-mail: kelvinkipchumba002@gmail.com

Abstract

Research on credit-risk prediction in Kenyan mobile lending uses different populations, outcomes and evaluation designs, making published model scores difficult to interpret. Licensed digital lenders issued about 9.6 million loans by August 2026, so the evidence behind scoring models matters. This rapid scoping review maps model families, data sources, prediction targets and validation approaches and identifies priorities for repayment-risk research. English-language theses, research projects and journal articles published between 30 October 2016 and 30 October 2026 were eligible. Web, repository and citation discovery was supplemented by OpenAlex and Crossref searches, restricted to the first 25 relevance-ranked records for each of eight query–index combinations (200 records; 186 after deduplication). Reporting was informed by PRISMA-ScR; screening and extraction in the update were single-pass and AI-assisted. Ten records were included, nine assessed in full text and one through an institutional abstract. Two analysed observed outcomes from Kenyan lenders: a study of 16,861 microfinance loans reported Random Forest AUC 0.7558, and a scorecard study of 495 digital-lender borrowers reported ROC-AUC 0.825. The other eight used foreign loans, simulations, a household-survey access proxy, an engineered risk label, household risk profiles or a scoring prototype. These tasks and evaluation conditions do not support a common performance ranking, and methodological variety exceeds directly validated evidence of future repayment. Priorities are independently observed repayment outcomes with explicit default definitions, local external validation, borrower-separated and temporal testing, and calibration and class-specific reporting. Practitioners should validate candidate models on their own lending population. Relevance-capped searches, one abstract-only source and the absence of independent human verification limit completeness and confidence; no universally superior algorithm is established.

Keywords: *Kenya; mobile lending; credit risk; machine learning; scoping review*

Introduction

Mobile money has changed how credit reaches Kenyan households and small businesses. Lenders can approve, disburse and collect small loans through a phone, using transaction records and other digital traces in place of conventional collateral. Evidence from M-Shwari links digital loans to household resilience to shocks, although wider credit-market constraints remain (Bharadwaj et al., 2019). Reliable risk assessment therefore matters for access to credit as well as for lender sustainability.

Digital credit has also drawn regulatory attention. The Central Bank of Kenya began licensing and supervising digital credit providers following amendments to the Central Bank of Kenya Act in 2021, and the Digital Credit Providers Regulations took effect in March 2022 (Central Bank of Kenya, 2022). In September 2026, the Bank reported 281 licensed providers and approximately 9.6 million loans valued at KSh 165.1 billion as of August 2026. Its oversight responded to concerns about high costs, unethical debt collection and misuse of personal information (Central Bank of Kenya, 2026).

The scale of digital credit makes lending decisions consequential. Mwesigwa et al. (2026), drawing on the 2025 Global Findex, note that 32% of Kenyan adults borrow from mobile-money providers and that default rates on digital loans have been reported as high as 40%, and they describe how data and risk definitions are negotiated in Kenyan digital lending. Model evaluation must therefore establish which population and outcome a score represents, alongside its predictive performance.

ISSN: 2309:9240, All Rights Reserved for all authors in this Journal

Credit-scoring benchmarks support evaluating ensembles alongside simpler classifiers. A benchmark of 41 classifiers across eight datasets indicates that several modern methods can outperform standard scorecards (Lessmann et al., 2015), whereas an earlier benchmark found that simple methods such as logistic regression also performed very well (Baesens et al., 2003). Mobile-phone records and digital footprints can carry repayment information (Björkegren & Grissen, 2020; Berg et al., 2020). However, findings from other settings do not establish performance for Kenyan mobile-loan products.

Kenyan-focused studies differ in their targets: Murphy (2018) and Omukami et al. (2025) use lender repayment outcomes; Oindi (2024) and Kisutsa (2021) use foreign loans; and Ng'ang'a (2025) predicts self-reported loan denial. Access and repayment are distinct outcomes, and treating them as interchangeable can confuse evidence about who obtains credit with evidence about who repays after disbursement.

Prediction quality includes discrimination, calibration and class-specific errors (Bequé et al., 2017; Saito & Rehmsmeier, 2015). Leakage can inflate results when test information enters development (Kapoor & Narayanan, 2023). A field experiment in Kenya also shows that people may change measured behaviour when decision rules become transparent, which matters for evaluation after deployment (Björkegren et al., 2026).

This review addresses the uncertain relationship between reported performance and the Kenyan repayment decision. Its aim is to map machine-learning studies of credit risk in Kenyan mobile and digital lending, distinguishing observed lender outcomes from indirect evidence, and to identify supported inferences and research priorities rather than rank algorithms.

Methodology

A rapid scoping review was used to map heterogeneous evidence rather than estimate a pooled comparative effect (Peters et al., 2020). Rapid methods streamline selected review steps under resource constraints (Garritty et al., 2021; Tricco et al., 2015). Here, retrieval was relevance-capped, and update screening and charting were single-pass and AI-assisted.

Reporting was informed by PRISMA-ScR (Tricco et al., 2018), with eligibility, sources, selection, charting, synthesis and limitations described below. PRISMA-ScR items map onto this manuscript as follows: rationale and objectives (Section 1), eligibility criteria (Section 2.2), information sources and search (Section 2.3), selection and charting (Section 2.4), synthesis (Section 2.5), characteristics and results (Sections 3.1–3.4), limitations (Section 3.8) and conclusions (Section 4). Figure 1 presents a PRISMA-style selection flow (Page et al., 2021). No prospectively registered protocol or formal critical appraisal was undertaken. Complete PRISMA-ScR checklist compliance is not claimed.

Four questions guided synthesis: (RQ1) Which model families are represented? (RQ2) What populations, data and outcomes are used? (RQ3) How is performance evaluated? (RQ4) What follows for Kenyan lending research and practice? The unit of analysis was a thesis, research project or journal article; the author's thesis experiments were not added as independently discovered evidence.

The population–concept–context framework (Peters et al., 2020) covered borrower or applicant records, including proxies and simulations; credit-scoring development or predictive evaluation; and an explicit Kenyan mobile- or digital-lending context. Eligible English-language theses, projects and journal articles were dated 30 October 2016 to 30 October 2026. Scorecards, neural architectures and hidden Markov classification were eligible. Foreign datasets required an explicit Kenyan mobile-lending focus. Prototypes were classified as development evidence rather than predictive validation.

Association-only surveys, reviews without original model development, generic bank or SACCO studies without a confirmed mobile-lending context, and regional studies without identifiable Kenyan relevance were excluded.

The original search on 30 October 2026 used public web discovery, institutional repositories and publisher records. Terms combined Kenya, mobile lending, digital credit, credit scoring, machine learning, default prediction and repayment. Representative queries were "Kenya mobile lending default prediction machine learning" and "Kenya mobile digital lending machine learning default prediction repayment empirical study credit scoring". Exact-title and author searches supported verification.

ISSN: 2309:9240, All Rights Reserved for all authors in this Journal

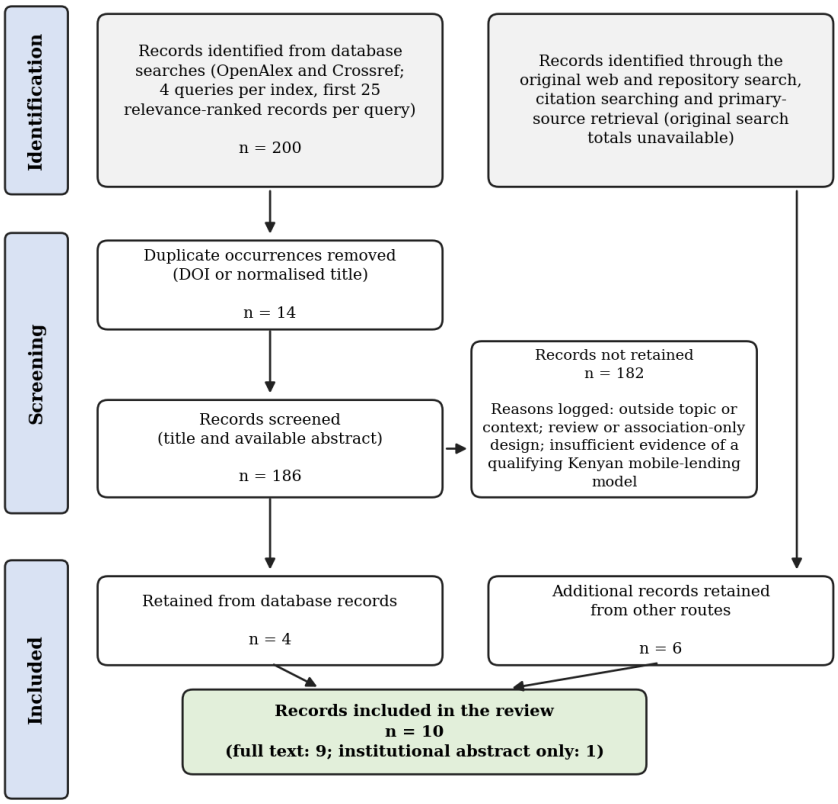
On 2 November 2026, OpenAlex and Crossref API searches used four queries in each index: “Kenya mobile lending”, “Kenya digital credit scoring”, “Kenya credit scoring machine learning” and “Kenya mobile default prediction”. Publication-date filters retained the original window. The first 25 relevance-ranked records per query–index combination yielded 200 occurrences. OpenAlex work search and Crossref’s bibliographic query field were used. Request URLs, match totals, retrieved records and decisions are preserved in the accompanying Search Evidence Pack. Match totals are not screened counts; the retrieval cap limits coverage.

Repository and citation retrieval covered National College of Ireland, Strathmore, University of Nairobi, International Hellenic University, Aga Khan University and Daystar records, plus primary publisher pages. Searches checked alternative-data scoring, social-network default prediction and M-Shwari hidden Markov models. Scopus, Web of Science and IEEE Xplore were not searched directly. Complete historical web-search counts were unavailable, so index counts are reported separately from other discovery routes.

DOIs or normalised titles identified 14 duplicates, leaving 186 index records screened by title and available abstract. Four were included; six came from other discovery routes. Exclusions covered unrelated contexts, reviews, association-only designs and records without a qualifying model. Primary records and accessible methods and results supported eligibility checks.

Independent eligibility screening by a second human reviewer was reported for the original review. The update was single-pass and AI-assisted, without independent human verification or agreement statistics. Nine included records were assessed in full text and one through an institutional abstract. The Ntwiga reports concern distinct simulated M-Shwari and Financial Diaries tasks; related publication versions were not counted separately.

Charted fields were author, year, document type, provenance, target, models, evaluation, results and access. Full-text findings took precedence over abstracts; discrepancies were recorded. Unavailable information did not establish procedural omission. Extraction notes and source locators appear in the Search Evidence Pack. No pooling or numerical risk-of-bias scoring was undertaken.



ISSN: 2309:9240, All Rights Reserved for all authors in this Journal

Figure 1. PRISMA-style flow diagram of record identification, screening and inclusion

Note. Original web and repository search totals cannot be reconstructed and are therefore not shown as a count.

Results address models (RQ1), populations and targets (RQ2), and evaluation (RQ3); discussion addresses implications (RQ4). Records were grouped by data provenance and outcome. Source metrics were retained without pooling or ranking across heterogeneous tasks. Background and methodological references were not counted as included modelling records.

Results and Discussion

Four index records and six from other routes yielded ten included records dated 2018–2026: six theses or projects and four journal articles (Figure 1; Table 1). Nine were assessed in full text; Omukami et al. (2025) remained abstract-only after unsuccessful publisher retrieval.

Table 1. Included records, prediction tasks and evidence access

S. No.	Record	Data and target	Models and reported evaluation	Access
1	Murphy (2018)	16,861 Kenyan MFI loans; unpaid by day 120	RF, regularised LR, SVM, neural nets, spreading activation; 70/30 split; RF AUC 0.7558	Full text
2	Ntwiga et al. (2018), IJEF	Simulated M-Shwari account factors; credit-quality class	HMM and SVD; comparisons of simulated individual and group scores	Full text
3	Ntwiga et al. (2018), JFE	280 Kenya Financial Diaries households; risk profile	HMM; household and peer-group credit-quality comparisons	Full text
4	Karingithi (2020)	Credit-history scoring prototype; independent outcomes unverified	Feed-forward neural architecture; API and acceptance tests; 92% claim not treated as default accuracy	Full text
5	Kisutsa (2021)	Foreign Bondora loans; default	LR, naïve Bayes, decision tree; best accuracy about 64%	Full text
6	Oindi (2024)	Foreign Bondora loans; default	RF and boosting; 80/20 split; best accuracy 86%	Full text
7	Ng'ang'a (2025)	Kenya FinAccess; loan denial	LR and ensembles; 75/25 split; adjusted LR recall 0.687, F1 0.335	Full text
8	Omukami et al. (2025)	495 Mobipesa borrowers; coding direction unclear	Stepwise LR and WoE; 70/30 split; ROC-AUC 0.825	Abstract only
9	Barasa and Wanyonyi (2026)	6,000 simulated digital loans	RF and XGBoost; 70/30 split; reported 10-fold mean AUC 0.9051	Full text
10	Githenya (2026)	30 original and 70 synthetic profiles; engineered class	Multiple classifiers; 80/20 split; gradient-boosting AUC 1.000	Full text

Note. AUC = area under the ROC curve; HMM = hidden Markov model; LR = logistic regression; MFI = microfinance institution; RF = Random Forest; SVD = singular value decomposition; SVM = support vector machine; WoE = weight of evidence.

Model Families Represented (RQ1)

Models included logistic regression, naïve Bayes, decision trees, ensembles, support vector machines, neural architectures, nearest-neighbour methods and hidden Markov models. Random Forest, logistic regression and boosting

ISSN: 2309:9240, All Rights Reserved for all authors in this Journal

recurred; hidden Markov models appeared in the two Ntwiga reports and a neural architecture in Karingithi's prototype.

Populations, Data Sources and Prediction Targets (RQ2)

Murphy (2018) analyses 16,861 Kenyan MFI loans, defining default as not fully repaid by day 120. A stratified 70/30 split and ten-fold training cross-validation yield Random Forest AUC 0.7558; AUC and profit criteria favour different feature sets. Omukami et al. (2025) report ROC-AUC 0.825 for 495 Mobipesa borrowers, but abstract descriptions leave outcome coding unresolved.

Kisutsa (2021) reports decision-tree accuracy around 64% on foreign Bondora loans. Oindi (2024) compares Random Forest with four boosting methods on Bondora, applies SMOTE and reports best accuracy of 86% for Random Forest and CatBoost. Neither establishes Kenyan borrower performance.

Ng'ang'a (2025) predicts FinAccess 2024 self-reported loan denial using four classifiers, a stratified 75/25 split and imbalance adjustments. Barasa and Wanyonyi (2026) evaluate 6,000 simulated digital loans: Random Forest ten-fold mean AUC is 0.9051 (SD 0.047). Their methods describe fivefold tuning and later tenfold robustness analysis.

Githenya (2026) combines 30 original M-Pesa profiles with 70 synthetic profiles, predicting an engineered behavioural risk category. An 80/20 split yields gradient-boosting AUC 1.000; the study discusses initial leakage and corrective preprocessing. This evaluates reproduction of a constructed label rather than independently observed future default.

Ntwiga, Ogutu and Kirumbu (2018) simulate M-Shwari factors and compare individual and peer-group credit-quality scores. Ntwiga, Ogutu, Kirumbu and Weke (2018) classify profiles for 280 Kenya Financial Diaries households. Neither report held-out future-default prediction; credit-quality proportions are not accurate. Karingithi (2020) develops a neural scoring prototype. Its abstract describes 92% borrower differentiation, whereas the testing chapter attaches 92% to API testing. Without an established independently labelled repayment test set, this figure is excluded from predictive comparisons.

Evaluation Designs and Reporting (RQ3)

Reported metrics answer different questions. Lender AUC, foreign-loan accuracy, survey-denial recall and simulation discrimination refer to different tasks; hidden Markov credit-quality proportions describe model outputs; and software tests describe functionality. Where reported, evaluation relied on single random partitions (70/30, 75/25 or 80/20) with or without cross-validation, and the charted results do not report borrower-separated or later-period test data.

Source discrepancies also affect extraction. Ng'ang'a's abstract reports an adjusted logistic-regression F1 of 0.3335, while Table 6.1 in the full text gives 0.335; both give recall of 0.687, and the table value is used here. In that table, Random Forest recall rises from 0.024 to 0.221 after resampling, while accuracy falls from 0.860 to 0.808. This shows how overall accuracy alone can conceal minority-class behaviour. The thesis also reports SHAP analysis and a demographic fairness assessment, so the absence of such analysis cannot be inferred for this record.

Interpretation of the Included Evidence (RQ4)

Two records use observed Kenyan lender outcomes (Murphy, 2018; Omukami et al., 2025). They support lender-specific feasibility; different lenders, outcomes and evaluations prevent claims of algorithm superiority. The other eight provide indirect or developmental evidence. Within these included records, methodological breadth exceeds directly evaluated repayment evidence for a common Kenyan population.

Bondora's accuracy of 86%, simulation AUC 0.9051 and engineered-label AUC 1.000 cannot establish stronger Kenyan default prediction than lender AUC 0.7558. Ranking these scores would confuse task differences with predictive quality.

Implications for Validation Design

Research should connect pre-decision predictors with later repayment in the intended population. Foreign-data and simulation findings require local validation; proxies, engineered labels, household profiles and prototypes require

ISSN: 2309:9240, All Rights Reserved for all authors in this Journal

evaluation against observed repayment. A random holdout does not make a constructed risk category equivalent to future default.

Validation design should follow the lending question. Repeated loans from one borrower can create dependence across training and test partitions, and a later-period test addresses change over time. Resampling, feature selection and tuning should remain inside training folds. These safeguards derive from prediction methodology (Kapoor & Narayanan, 2023) and are not a claim that every included study violated them.

Implications for Reporting and Deployment

Operational evaluation should combine discrimination, calibration and class-specific errors (Bequé et al., 2017; Saito & Rehmsmeier, 2015). Murphy's different AUC-preferred and profit-preferred feature sets show why selection must reflect the lending decision. Thresholds should account for default and unnecessary rejection costs; subgroup evaluation should accompany feature explanations. Because decision rules may change behaviour once they are transparent (Björkegren et al., 2026), a score's performance at validation may not hold unchanged after deployment.

Limitations

Retrieval was limited to 25 results per query per index. Ranking, terminology and index coverage may therefore omit eligible work. OpenAlex and Crossref do not replace specialist databases. Original web search counts are unknown, so Figure 1 is not a complete historical PRISMA flow.

Single-pass AI-assisted update screening and extraction may contain errors. One source is abstract-only; no formal critical appraisal or prospective registration was undertaken. Gathu (2020) and Ngonde et al. (2026) were excluded as association studies without established predictive evaluation at the available access level. These boundaries preclude sector-wide prevalence claims or causal comparisons.

Conclusion and Recommendations

The ten included records cover diverse methods and prediction tasks, preventing identification of an overall best model. Two use observed Kenyan lender outcomes and support lender-specific feasibility; the remaining records provide indirect or developmental evidence. Cross-lender transferability and performance in a common Kenyan repayment population remain unresolved within this evidence set.

The gaps identified support the following recommendations.

- **Observed outcomes.** Proxy, simulated and profile-based studies should be followed by independently observed repayment evaluation, with explicit default definitions and prediction horizons.
- **Local validation.** Foreign-data and simulation findings require evaluation on Kenyan lender data before operational use.
- **Validation design.** Borrower-separated and later-period testing should assess generalisation, with preprocessing and tuning confined to training data.
- **Reporting.** The accuracy–recall discrepancy in Ng'ang'a (2025) and Murphy's decision-cost findings support reporting calibration, class-specific errors, uncertainty and lending costs alongside discrimination.
- **Outcome-based validation.** Prototype and hidden Markov studies need validation against actual loan performance before claims of repayment prediction are made.
- **Practice.** Practitioners should validate candidate models on their own population and not transfer published scores without evaluation.

Future updates should widen retrieval caps, add specialist databases and use independent human screening and extraction. The accompanying Search Evidence Pack supports inspection of the documented update.

References

Baesens, B., Van Gestel, T., Viaene, S., Stepanova, M., Suykens, J., & Vanthienen, J. (2003). Benchmarking state-of-the-art classification algorithms for credit scoring. *Journal of the Operational Research Society, 54*(6), 627–635. https://doi.org/10.1057/palgrave.jors.2601545

ISSN: 2309:9240, All Rights Reserved for all authors in this Journal

Barasa, I. N., & Wanyonyi, S. W. (2026). Predictive modeling of digital credit risk in commercial banks using machine learning algorithms. *Asian Journal of Advanced Research and Reports, 20*(3), 110–132. https://doi.org/10.9734/ajarr/2026/v20i31305

Bequé, A., Coussement, K., Gayler, R., & Lessmann, S. (2017). Approaches for credit scorecard calibration: An empirical analysis. *Knowledge-Based Systems, 134*, 213–227. https://doi.org/10.1016/j.knosys.2017.07.034

Berg, T., Burg, V., Gombović, A., & Puri, M. (2020). On the rise of FinTechs: Credit scoring using digital footprints. *The Review of Financial Studies, 33*(7), 2845–2897. https://doi.org/10.1093/rfs/hhz099

Bharadwaj, P., Jack, W., & Suri, T. (2019). *Fintech and household resilience to shocks: Evidence from digital loans in Kenya* (NBER Working Paper No. 25604). National Bureau of Economic Research. https://doi.org/10.3386/w25604

Björkegren, D., Blumenstock, J. E., & Knight, S. (2026). Manipulation-robust prediction. *American Economic Review, 116*(9), 3263–3293. https://doi.org/10.1257/aer.20241087

Björkegren, D., & Grissen, D. (2020). Behavior revealed in mobile phone usage predicts credit repayment. *The World Bank Economic Review, 34*(3), 618–634. https://doi.org/10.1093/wber/lhz006

Central Bank of Kenya. (2022). *The Central Bank of Kenya (Digital Credit Providers) Regulations, 2022*. https://www.centralbank.go.ke/2022/03/21/central-bank-of-kenya-digital-credit-providers-regulations-2022/

Central Bank of Kenya. (2026, September 30). *Licensing of 29 additional digital credit providers* [Press release]. https://www.centralbank.go.ke/uploads/press_releases/366319113_Press%20Release%20-%20Licensing%20of%2029%20Additional%20Digital%20Credit%20Providers.pdf

Garritty, C., Gartlehner, G., Nussbaumer-Streit, B., King, V. J., Hamel, C., Kamel, C., Affengruber, L., & Stevens, A. (2021). Cochrane Rapid Reviews Methods Group offers evidence-informed guidance to conduct rapid reviews. *Journal of Clinical Epidemiology, 130*, 13–22. https://doi.org/10.1016/j.jclinepi.2020.10.007

Gathu, A. (2020). *The role of alternative data in accurately determining credit score for mobile lending on digital wallets in Kenya* [Master's thesis, Strathmore University]. http://hdl.handle.net/11071/10193

Githenya, A. M. (2026). *AI-based credit risk modelling for financially excluded M-Pesa users in Kenya using alternative data* [Master's thesis, International Hellenic University]. https://repository.ihu.gr/handle/11544/49990

Kapoor, S., & Narayanan, A. (2023). Leakage and the reproducibility crisis in machine-learning-based science. *Patterns, 4*(9), Article 100804. https://doi.org/10.1016/j.patter.2023.100804

Karingithi, L. M. (2020). *A centralized credit scoring prototype for micro lending institutions in Kenya* [Master's thesis, Strathmore University]. http://hdl.handle.net/11071/12172

Kisutsa, G. T. (2021). *Loan default prediction using machine learning: A case of mobile based lending* [Master's project, University of Nairobi]. https://erepository.uonbi.ac.ke/items/caacdfc3-4259-45c2-a4ac-c418970a361a

Lessmann, S., Baesens, B., Seow, H.-V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research, 247*(1), 124–136. https://doi.org/10.1016/j.ejor.2015.05.030

Murphy, D. (2018). *Prediction of loan defaulters in micro finance using social network data* [Master's research project, National College of Ireland]. https://norma.ncirl.ie/3414/

Mwesigwa, D., Jackson, S. J., & Csikszentmihalyi, C. (2026). Risk, data, alignment: Making credit scoring work in Kenya. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26)*. ACM. https://doi.org/10.1145/3772318.3790924

Ng'ang'a, A. A. (2025). *AI-driven credit scoring using alternative data: Unlocking financial access in emerging markets* [Master's research project, National College of Ireland]. https://norma.ncirl.ie/9577/

Ngonde, M. M., Bowen, M., & Nzomoi, J. N. (2026). An empirical analysis of the determinants of digital credit default in Kenya. *International Journal of Research and Innovation in Social Science, 10*(3), 4078–4086. https://doi.org/10.47772/IJRISS.2026.100300303

Ntwiga, D. B., Ogutu, C., & Kirumbu, M. K. (2018). Inclusion of peer group and individual low-income earners in M-Shwari micro-credit lending: A hidden Markov model approach. *International Journal of Electronic Finance, 9*(2), 121–133. https://doi.org/10.1504/IJEF.2018.092195

Ntwiga, D. B., Ogutu, C., Kirumbu, M. K., & Weke, P. (2018). A hidden Markov model of risk classification among the low income earners. *Journal of Finance and Economics, 6*(6), 242–249. https://doi.org/10.12691/jfe-6-6-6

Oindi, B. (2024). *A credit scoring model for mobile lending* [Master's thesis, Strathmore University]. http://hdl.handle.net/11071/15651

Omukami, H., Mwau, C., Wangui, D. K., & Mbuya, A. (2025). Credit scoring in digital lending: A stepwise logistic regression and weight of evidence approach in Kenya. *The Egyptian Statistical Journal, 69*(2), 229–244. https://doi.org/10.21608/esju.2025.415357.1115

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, Article n71. https://doi.org/10.1136/bmj.n71

Peters, M. D. J., Marnie, C., Tricco, A. C., Pollock, D., Munn, Z., Alexander, L., McInerney, P., Godfrey, C. M., & Khalil, H. (2020). Updated methodological guidance for the conduct of scoping reviews. *JBI Evidence Synthesis, 18*(10), 2119–2126. https://doi.org/10.11124/JBIES-20-00167

Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE, 10*(3), Article e0118432. https://doi.org/10.1371/journal.pone.0118432

Tricco, A. C., Antony, J., Zarin, W., Strifler, L., Ghassemi, M., Ivory, J., Perrier, L., Hutton, B., Moher, D., & Straus, S. E. (2015). A scoping review of rapid review methods. *BMC Medicine, 13*, Article 224. https://doi.org/10.1186/s12916-015-0465-6

ISSN: 2309:9240, All Rights Reserved for all authors in this Journal