CONTEXTUALIZING AI ETHICS IN UGANDA'S MICROCREDIT WITH ADAPTIVE SENSITIVE REWEIGHTING

EMMANUEL ISABIRYE

J23MD10/217

A DISSERTATION SUBMITTED TO THE FACULTY OF ENGINEERING, DESIGN AND TECHNOLOGY IN PARTIAL FULFILLMENT FOR THE REQUIREMENTS OF THE AWARD OF THE DEGREE OF MASTER OF SCIENCE IN DATA SCIENCE AND ANALYTICS OF UGANDA CHRISTIAN UNIVERSITY

August,2025

UGANDA CHRISTIAN UNIVERSITY
A Centre of Excellence in the Heart of Africa

Abstract

This research tackles the pressing ethical concerns of using Artificial Intelligence (AI) in Uganda's microcredit sector, namely to develop an Adaptive Sensitive Reweighting (ASR) model to mitigate algorithmic bias and promote equitable access to credit. Traditional credit scoring models - and AI algorithms trained on Western-biased data - discriminate against marginalized groups because they are based on formal financial records, reinforcing structural disadvantages. By iterative engagement with Ugandan policymakers, lenders, borrowers, and AI experts, we identify the most significant ethical concerns and specify context-specific fairness metrics. The ASR approach adaptively adjusts weights for sensitive features like collateral values and transaction history during model training to enhance fairness. Experimental outcomes on a typical credit scoring dataset demonstrate ASR's success: the inclusion rate of disadvantaged borrowers is enhanced by 15% with predictive accuracy maintained, and significant improvements on key fairness metrics. The research provides actionable policy recommendations on implementing ASR-based AI systems in Uganda's microfinance sector to drive financial inclusion and sustainable development. This study contributes to emerging Majority World scholarship on AI ethics by demonstrating the necessity of situating ethical frameworks and valuing stakeholder perspectives to develop equitable, inclusive AI systems. Our findings offer valuable insights for policymakers, microfinance institutions, and AI practitioners who aim to implement responsible AI in Uganda's Microcredit sector.

DECLARATION

I, Emmanuel ISABIRYE, hereby declare that this thesis is my original work and research, is not plagiarized, and that the methodology and outcomes have never been submitted for any other qualification or degree.
This work was carried out according to the highest standards of ethics, with informed consent and confidentiality being given top priority. All clearance was sought from the Uganda Christian University Research Ethics Committee.
Name: Emmanuel ISABIRYE

Signed: ....

Date: 01/08/2025

APPROVAL

This is to certify that the following research titled: CONTEXTUALIZING AI ETHICS IN UGANDA'S MICROCREDIT WITH ADAPTIVE SENSITIVE REWEIGHTING has been done under my supervision and is now ready for submission.

Name: Dr. Daphne Nyachaki Bitalo

Signed: ....

Date: 01/08/2025 ....

DEDICATION

This thesis is dedicated to the Almighty God, for making everything possible.

ACKNOWLEDGEMENTS

I give thanks to the Almighty God for His guidance and blessings throughout this research journey.

The author gratefully acknowledges using IBM SPSS Text Analytics for Surveys for thematic analysis and sentiment coding of qualitative responses from stakeholder consultations. The author utilized the application to organize, analyze, and interpret qualitative remarks rigorously and systematically. The author further acknowledges Spectrallabs for providing the credit scoring dataset, which was applied in the quantitative analysis. The dataset played a crucial role in evaluating the performance of the Adaptive Sensitive Reweighting (ASR) framework.

The author also extends sincere gratitude to the supervisor, Dr. Bitalo Daphne, for her invaluable guidance, support, and constructive feedback throughout this research. Her expertise was instrumental in shaping this thesis.

The author would also like to acknowledge my parents, Mr. and Mrs. Wakabi, for their encouragement and support.

The author also extends their appreciation to friends and family, the Department of Computing & Technology and the Research Ethics Committee.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY OF TERMS AND ACRONYMS

- AI (Artificial Intelligence): The simulation of human intelligence functions by machines, most notably computer systems. This includes learning (gaining information and rules to use the information), reasoning (using rules to arrive at approximate or certain conclusions), and self-correction. In this study, AI is employed to denote the design of credit scoring algorithms for microfinance.

- Algorithmic Bias: Systematic and repeatable errors in a computer system that generate unfair outcomes, e.g., benefitting one set of random users against the others. It could be because training data is inadequate or because the algorithm design itself.

- ASR (Adaptive Sensitive Reweighting): A type of machine learning that minimizes bias by adjusting the weight given to particular data points when training the model. It stops the model from discriminating against some groups in an unequal way.

- Explainability: The extent to which the inner workings of an AI system can be explained to humans. It is vital in building trust and holding someone or something accountable.

- Fairness: Fairness in AI refers to the absence of prejudice or favoritism towards an individual or a group due to their intrinsic or acquired qualities.

- Financial Inclusion: It refers to availability and parity of opportunities to use financial services.

- FPR (False Positive Rate): The ratio of true negatives that are misclassified as positives. Here, it is the ratio of non-high-risk borrowers who are misclassified as high-risk by the model.

- IQR (Interquartile Range): A statistical measure of dispersion, defined as the difference between the 75th and 25th percentiles.

- MDIs (Microfinance Deposit-Taking Institutions): Regulated banks in Uganda permitted to take deposits from the public and undertake microfinance activities.

- MFIs (Microfinance Institutions): Financial institutions that provide financial products, e.g., small loans, to poor households and micro-enterprises.

- MIT (Massachusetts Institute of Technology): World-renowned research university located in Cambridge, Massachusetts, renowned for its solid science and engineering programs.

- OECD (Organization for Economic Co-operation and Development): A 38-member multilateral economic organization founded in 1961 with the goal of furthering economic progress and world trade.

- Post-ASR: Referring to the model or period after the application of Adaptive Sensitive Reweighting.

- Pre-ASR: Referring to the model or period before the application of Adaptive Sensitive Reweighting.

- SACCOs (Savings and Credit Cooperative Organizations): Members-owned financial co-ops which provide credit facilities and savings to members.

- SHAP (Shapley Additive Explanations): A game-theoretic technique to explain the prediction of any machine learning model by distributing the contribution of every feature to the prediction in a roughly fair way.

- TPR (True Positive Rate): The proportion of true positives well-classified as positives. In this context, it is the proportion of the truly high-risk borrowers well-classified as high-risk by the model.
- XGBoost (Extreme Gradient Boosting): Popular and powerful open-source implementation of the gradient boosted trees algorithm.

1 CHAPTER 1: INTRODUCTION

1.1 BACKGROUND: MICROCREDIT AND ACCESS CHALLENGES IN UGANDA

Microcredit, notably championed by Muhammad Yunus and the Grameen Bank, emerged with the aim of alleviating poverty in developing nations by providing financial access, particularly for small businesses and women entrepreneurs [32, 2]. While its overall effectiveness in transforming livelihoods remains a subject of ongoing debate [16], a fundamental challenge persists in contexts like Uganda: extending financial reach to excluded populations. Many microfinance institutions (MFIs) in Uganda struggle to serve significant segments of the population, especially those in the informal sector, rural areas, or lacking the formal financial documentation and collateral required by traditional credit scoring systems [19]. This exclusion limits opportunities and hinders broader economic development, highlighting a critical gap that innovative approaches seek to address.

1.1.1 AI Potential for Inclusion and Risk Assessment

Artificial intelligence (AI) presents potential pathways to tackle these challenges in microcredit. By leveraging diverse, non-traditional data sources (such as mobile phone usage patterns or utility payments, where ethically permissible), AI-driven systems may identify and assess the creditworthiness of individuals currently invisible to conventional methods relying solely on formal financial histories [29]. This capability holds particular significance for enhancing financial inclusion in Uganda by potentially reaching previously unbanked or underbanked populations. Beyond expanding reach, AI also offers the potential for more sophisticated risk assessment for those included, potentially leading to more tailored loan products, although this capability must be viewed alongside the significant ethical considerations discussed next.

1.1.2 The Ethical Implications: Challenges and Debates

When put into practice in making important choices in the areas of money lending, AI application is confronted with deep ethical questions. One of the most basic questions is something known as algorithmic bias. If the data used to train an AI embodies past discrimination (for instance, certain groups receiving unfairly lower loans), the AI will learn this discriminatory pattern and perpetuate the discrimination against the same groups when devising new decisions. The AI does not know it is being discriminatory; it is just copying the bias it has learned from the data. [18]; [23]. This means that AI-based credit scoring, if not done carefully, can unjustly deny loans based on characteristics such as location, gender or social networks instead of financial capacity itself. Practical ramifications of AI bias can also be witnessed from the illustrations of pretrial risk assessment in additional areas. In the meantime, the "black box" propensity of advanced AI models, especially deep neural networks, dissuades transparency. This results in a lack of transparency about how loans are determined by the borrowers, which also causes lack of trust and inability of the latter to see how they can build good credit [24]. Such unexplainability is morally alarming, narrowing the potential for disputing biased decisions. Privacy issues are also increased through the use of alternative data for credit scoring [11]. For instance, borrowers may not know how their data from mobile phones or certain behavior on it determine their financial future. There is a need to think about informed consent and misuse of personal data. Besides that, the development of AI systems is often involuntarily led by large technology corporations headquartered far from where they would be put into use leading to a situation of disequilibrium between such corporations and local communities [11].

1.1.3 Ugandan Challenges: Addressing Artificial Intelligence

Ugandan financial inclusion is confronted with a multifaceted challenge in which the nation must avoid the risks of AI but also exploit its benefits. This comes with addressing fairness, transparency, privacy, and stakeholders' involvement as the foremost priorities thereof. There are worldwide ethical guidelines for AI, whose effective application at microcredit level in Uganda necessitates some adjustment and research. Without intervention mechanisms, credit systems run by AI can only exacerbate existing economic and social disparities among groups in any society. To address these, we must consider some fundamental things as well as areas where there are still information gaps:
Auditing Datasets:
It is important to thoroughly inspect AI models trained in Uganda for features likely leading to biased outputs. This goes beyond superficial categories by involving stakeholders who will help to assess peculiar social, economic and cultural issues within the Ugandan context.

Privacy Frameworks:
Hence, borrowers have to determine what is considered as informed consent, how data is being used and what forms of practices are acceptable within microfinance in Uganda as it can be seen below:

It is crucially important that there be extensive dialogue among stakeholders so that one can understand meanings attributed to fairness by Ugandan borrowers, lenders, and policymakers when dealing with AI-based credit scoring.

AI explanations must be designed for the specific audience's understanding, requiring methods beyond purely technical descriptions to ensure clarity, trust, and accessibility for everyone involved, regardless of their technical background.

It is imperative that technologists and researchers in Uganda develop and localize AI modules because this aids in contextual innovation by avoiding reliance on external models while promoting ownership of the technologies.

Ugandan policy makers should have well researched arguments on how to make their AI policies balance between promoting innovation and protecting individual rights due to research.

This research sought to fill those gaps emphasizing collaboration with borrowers, microfinance institutions, advocacy groups, and policymakers. To ensure that this research was grounded and impactful, the researcher worked together with borrowers, microfinance institutions, advocates of the poor, as well as policy-makers to ensure that the developed ethical AI frameworks whose applications are found only within the context of Uganda and were designed to promoted micro credit financing goals while ensuring that fairness, transparency and inclusivity are upheld at all times.

1.2 STATEMENT OF THE PROBLEM

Uganda's microcredit sector faces a critical challenge: traditional credit scoring systems, reliant on formal financial histories, exclude a significant portion of the population from accessing essential financial services [19]. This exclusion reinforces the financial divide and hinders comprehensive development. The application of AI in credit risk evaluation using alternative data holds significant promise to this end [29]. The role of AI introduces ethical concerns, particularly in the choice of Western-oriented data and frameworks for model construction. These models might not fit Uganda's specific socio-economic context [11] and are likely to reinforce current stereotypes while further sidelining the already marginalized groups. Moreover, such complex AI models implement opacity, hence borrowers cannot trust nor challenge the decision made by these models [24]. This study attempted to address these issues by testing Adaptive Sensitive Reweighting (ASR) as a tool for reducing bias. There is potential in ASR [28], but one needs to question closely the contextual relevance, feasibility, and efficacy of applying it within the social-economic landscape of Uganda.

This study aimed at designing an ethical framework within which a prescriptive AI-powered credit scoring system is developed using ASR toward fairness and inclusiveness in Uganda's microcredit ecosystem.

1.3 JUSTIFICATION

This research is critically important due to the persistent challenge of financial exclusion in Uganda, which hinders equitable development and necessitates innovative, ethically-grounded solutions. Despite government strategies aiming to increase financial inclusion, a significant portion of the adult population remains excluded from formal financial services – estimates from recent years suggest around 15-24% were excluded entirely, with potentially up to 65% unable to access formal credit [15, 30]. This exclusion disproportionately affects marginalized groups, particularly rural populations (where 74% rely on informal finance - [15]; and exclusion rates are higher - [30]), women (who also rely more on informal mechanisms - [30]; and face specific economic barriers - [9]), smallholder farmers (receiving only 10% of credit flow, citing BoU), and those operating in the informal sector. Such widespread exclusion limits investment, restricts opportunities for economic advancement, contributes to poverty persistence, and potentially slows overall household income growth.

The root of this exclusion often lies in the current credit scoring methods employed by many financial institutions, including MFIs. These typically rely heavily on formal financial histories, documented income streams, and tangible collateral security – requirements that large segments of the Ugandan population, especially the marginalized groups mentioned, cannot meet [19, 15]. While MFIs may also use socio-demographic data, behavioural loan history (for repeat clients), and sometimes expert judgement or basic statistical models [14](Accion, n.d.; Assairh et al., 2021), these methods often struggle with data limitations and can still inadvertently screen out potentially creditworthy individuals who lack conventional documentation [14] (CGAP, 2018).

While Artificial Intelligence (AI) offers the potential to address these limitations by leveraging alternative data sources [29], its direct application carries significant ethical risks. Uncritically deploying AI models, especially those developed in Western contexts, can perpetuate or even amplify existing societal biases, leading to new forms of discrimination [26, 11, 18]. Therefore, simply replacing traditional methods with standard AI is insufficient and potentially harmful.

This study is necessary because it directly confronts this dual challenge: the inadequacy of traditional methods and the ethical perils of uncontextualized AI. This research proposes and evaluates the Adaptive Sensitive Reweighting (ASR) framework as a specific intervention designed to curb marginalization. ASR addresses the limitations by providing a mechanism to systematically adjust the influence of data features (like collateral or transaction history) identified as sources of bias against specific groups within the Ugandan context. By reweighting, ASR aims to create fairer predictive models that increase the inclusion of marginalized yet creditworthy individuals, offering a pathway beyond both traditional exclusion and naive AI implementation. Establishing the feasibility and effectiveness of such context-aware, fairness-enhancing AI techniques is crucial for guiding the responsible adoption of technology, promoting genuine financial inclusion, ensuring non-discrimination, and contributing to sustainable and equitable development in Uganda and the broader Majority World.

1.4 OBJECTIVES

1.4.1 General Objective:

To develop and implement an ethical AI-powered credit scoring system for Ugandan microcredit that mitigates bias and promotes financial inclusion, using an Adaptive Sensitive Reweighting (ASR) framework.

1.4.2 Specific Objectives:

1. Identify and analyze Ugandan stakeholders' definitions of ethical AI in microcredit.
2. Design an ASR framework tailored to the Ugandan microcredit context based on the identified ethical insights.
3. Develop a Western-centric AI credit scoring model, adapt it to the Ugandan context by integrating ASR, and evaluate its impact on fairness metrics and stakeholder perceptions.
4. Provide evidence-based recommendations for ethical deployment and monitoring of ASR in Ugandan microcredit, including mechanisms for stakeholder engagement and oversight.

1.5 RESEARCH QUESTIONS

1. How do Ugandan stakeholders define fairness in AI-powered microcredit, and do these definitions align with global ethical AI discourses?
2. Which sensitive attributes in the Western credit dataset have the highest bias potential, and how can ASR be tailored to mitigate these biases while maintaining accuracy and incorporating stakeholder feedback?
3. How effective is integrating ASR into a Western-centric AI credit scoring model in reducing bias and improving fairness, and do stakeholder perceptions align with these quantitative measures?
4. What policy recommendations and oversight mechanisms can ensure the ethical deployment of ASR-based AI in Ugandan microcredit, balancing innovation with the protection of individual rights?

1.6 SIGNIFICANCE OF THE STUDY

This research makes significant contributions to both scholarly discourse and real-world practice:

1.6.1 Theoretical Contributions

The research deepens understanding of bias mitigation strategies by demonstrating the necessity of tailoring them to the specific socioeconomic realities and values systems of developing economies and African microcredit markets.

It enriched the global discourse on ethical AI by highlighting the limitations of universal approaches and underscoring the unique challenges, stakeholder perspectives, and potential for locally-driven innovation within the Ugandan context.

The study emphasized the value of collaborative and participatory design models for developing ethical AI solutions that are both culturally resonant and effective, challenging technocratic approaches to AI development.

1.6.2 Practical Contributions

The research provides Ugandan microfinance institutions with evidence-based insights on the responsible and effective implementation of ASR, promoting fairer and more inclusive credit decision-making.

It offered Ugandan policymakers' actionable recommendations for crafting AI regulations that balance innovation with the protection of fairness and individual rights within the microfinance sector. These recommendations have the potential to shape a more equitable AI-powered landscape.

The findings generated have potential relevance for similar contexts across Africa and other developing regions grappling with the challenges and opportunities of harnessing AI ethically for financial inclusion.

1.7 SCOPE AND LIMITATIONS

This research investigates the ethical integration of Artificial Intelligence (AI) into Uganda's microcredit sector, specifically through the design, implementation, and evaluation of an Adaptive Sensitive Reweighting (ASR) framework. The study's scope is deliberately focused to allow for an in-depth exploration of ASR's potential to mitigate bias in AI-driven credit scoring models.

Scope: Geographic Focus: The study is done in the context of the Ugandan microcredit industry to take into account the micro and macro factors in Uganda.

Methodological Approach: Both qualitative methods of survey with the key stakeholders: borrowers, lenders, policy makers, and artificial intelligent specialists, and quantitative methods of evaluating a credit scoring dataset are applied.

Technical Focus: The technical focus of the study involves the creation and assessment of the ASR framework, a new set of algorithms designed to assess different weight values of attributes during the data mining and model building phases.

Ethical Dimensions: The study mainly focuses on the ethical issues of fairness and bias, transparency, and explainability as envisaged in AI based credit Scoring and using existing ethical AI frameworks, complemented with the local environment.

Data Source: The quantitative analysis leverages a dataset embodying Western-bias credit scoring tendencies. Although it is not directly collected from Uganda, this dataset can be used as the baseline for showing the shortcomings of conventional models and promising application of ASR in overcoming those for context-specific information.

Limitations: While the stakeholder consultations (n=80) provided valuable qualitative depth and were designed to capture diverse perspectives across key groups (borrowers, lenders, policymakers, AI experts) aiming for thematic saturation, the sample size may not be fully statistically representative of the entire Ugandan microcredit ecosystem, potentially limiting the generalizability of some qualitative findings. The ASR implementation focused on transaction history, collateral, and risk scores; exploring other potentially sensitive attributes (e.g., gender, location) requires further research navigating associated ethical complexities. Finally, the ASR evaluation occurred in a simulated environment, and real-world pilot studies are needed to assess performance under operational conditions. The findings should also be interpreted within the current, rapidly evolving technological landscape of AI.

1.8 STRUCTURE OF THE THESIS

This thesis is divided into five chapters, all of which aim at providing an umbrella solution to the research problem under analysis. The structure itself is made logical and coherent to help the reader follow the exposition from the conceptual roots and the research topic all the way to the conclusions drawn from the empirical study and sharing the implications derived from the results.

Chapter 1: Introduction: This chapter establishes the research context by providing background information on the evolving microcredit sector in Uganda, the transformative potential of AI in financial inclusion, and the ethical challenges associated with AI-driven credit scoring. It introduces the problem statement, highlighting the limitations of traditional credit scoring models and the potential biases of Western-centric AI systems. The research justification, objectives, and research questions are clearly articulated, followed by a discussion of the study's significance, scope, and limitations. The chapter concludes with an overview of the thesis structure.

Chapter 2: Literature Review:

The chapter focuses on a critical analysis of literature in respect of the specific research interest. It will start with the introduction of ethical AI, the different perspective, and guidelines such as the UNESCO and OECD guidelines and the EU AI Act. It then goes further to highlight the use of AI in developing countries with a special emphasis on Uganda and with a benchmark with micro credit. Also, the chapter discusses the ethical considerations particular to Uganda, such as data skewing, systematic marginalization, and culture. It proposes the use of Adaptive Sensitive Reweighting (ASR) as a bias reduction method and outlines areas of the research that are yet unexplored to which this study seeks to contribute.

Chapter 3 Research Methodology:

The objectives of this chapter are to describe research design and methods used in the study. It outlines the appraisal scheme with emphasis on the reasons why only both qualitative and quantitative approaches were deemed relevant for the study. Information about interviews and focus groups and the sampling strategy for the study and analysis of qualitative data are presented. The quantitative approach in this study, which consists of the establishment and assessment of the ASR framework, is also explained. This includes the description of the data set, steps followed under data preprocessing, process of training the model and fairness and performance measures. Finally, the ethical considerations that were followed in the research are also presented in the chapter.

Chapter 4: Results

This chapter discusses the results of the conducted empirical research and provides an analysis of the results obtained. It starts with an overview of the qualitative finding derived from a series of engagements with the stakeholders: the perceptions of the interlocutors on fairness, bias, and privacy in AI microcredit. This is then succeeded by an extensive exploration of the quantitative findings, which are concerned with the effects of ASR on the differential treatment of the groups, and the performance of the classification models. This chapter also uses figures, the scatter plot of the descriptors, and the horizontal stacked bar chart to show the findings. To facilitate the comparison, first a comparative discussion discusses the pre- and the post-ASR model performance at aggregate and instance level is presented. We then present the discussion of the results in light of the literature and the identified research questions.

Chapter 5: Conclusion and Recommendations

This chapter summarizes the study and lays down the findings of the study highlighting on the major outstanding findings in ethical AI. It offers concrete policy advice with regard to the proper adoption of AI technologies for the Ugandan microcredit industry concerning the creation of contextualized fairness benchmarks, participation of all the interested parties, and the enhancement of expert knowledge. The chapter also outlines the specific application of this research for lending institutions and policy makers. Finally, it describes the implications of the study and its limitations and suggests ideas for empirical research.

2 CHAPTER 2: LITERATURE REVIEW

This review synthesizes existing literature relevant to the study's objectives concerning the ethical integration of AI into Ugandan microcredit, focusing on contextualized fairness and the Adaptive Sensitive Reweighting (ASR) framework.

2.1 ETHICAL AI DISCOURSES AND THE IMPORTANCE OF UGANDAN CONTEXT

It is necessary to situate Ugandan stakeholders' conceptualizations of ethical AI within both international discourses and the specific local context. International debates grapple with ensuring justice and preventing discrimination in AI, and they focus on values like justice, responsibility, transparency, explainability, and privacy. But it's hard to frame these principles in terms of universal standards [5], and conventional abstract, generally Western, principles can neglect the empirical AI ethics central to lived experience (the target of this work).

Applying these concepts to African and other emerging economies entails dealing with special challenges. Power relations in AI development, usually dominated by multinational technology firms, can potentially create mismatches with local priorities [11]. AI models trained on high-income nation data can potentially harbor global biases if applied without local adjustment in an environment like Uganda, exacerbating prevailing inequalities [24]. This underscores the need for decolonial approaches to AI ethics in the Majority World.

Moreover, the very concept of "fairness" itself should be embedded within context in a manner that transcends typical technical controls (like demographic parity or equalized odds). Fairness in Uganda's microcredit sector might entail distributive justice goals like poverty alleviation and equitable access for marginalized populations (like rural areas, women, informal sector workers), as per prevailing local values and intersectional conditions [11, 24]. Formal credit history-based traditional methods usually drive exclusion in Uganda [19], so a definition of fairness assumes special significance. Adopting standards based on community values and common-sense experience, as encouraged by perspectives like feminist care and biomedical ethics, is essential [24]. Trust also needs culture-aware forms of explainability and regional attitude sensitivity towards AI and data privacy [11, 24]. Therefore, writing highlights that understanding and reviewing particular definitions and concerns of Ugandan stakeholders covered by Objective 1 is essential before the development or implementation of ethical AI systems. Designing Contextualized ASR (Objective 2)

In order to create a productive and ethical AI system like Adaptive Sensitive Reweighting (ASR) for Ugandan microcredit, one has to look into the extensive application of AI within this industry, its constraints, and the background of bias evasive frameworks. An excellent example of such an application of this moral balance is mobile money data. Although platform transaction data like that of MTN Mobile Money or Airtel Money offers a rich, secondary source on which to judge the creditworthiness of the unbanked, it is also enormously hazardous ethically. Mobile money spending trends—transaction frequency, payments to specific merchants, or remittance inflows—can be very strong proxies for socio-economic status, location, and lifestyle. As [11] reminds us, without robust ethical frameworks and regulation, there is a genuine threat that this data will be used to create new types of discrimination, penalizing individuals who live in rural areas and have less access to e-transactions or who have consumption patterns that deviate from some imagined norm. Therefore, the potential of inclusion by the mobile money data must be balanced carefully against the risk of it perpetuating the current biases or the creation of new ones, something that this research addresses through stakeholder-driven research.

2.2 CURRENT AI USE IN MICROCREDIT SCORING

Recent literature demonstrates widespread expansion of the application of artificial intelligence (AI) and machine learning (ML) in microcredit scoring in developing economies, particularly in Africa, with the primary objective of enhancing financial inclusion through the use of alternative data analysis. Conventional credit scoring methods, relying as they do on official financial histories, routinely disenfranchise approximately 64% of the unbanked in sub-Saharan Africa [16]. To redress this disenfranchisement, more recent AI-based solutions make use of non-traditional data vectors like mobile money transactional behavior, utility bill payment histories, digital footprint analysis, and - in pilot cases - psychometric profiling.

Algorithmic solutions range from simple logistic regression models to more complicated architectures such as: Support Vector Machines (SVMs), Deep Neural Networks (DNNs), Gradient Boosted Decision Trees (GBDT/XGBoost), Ensemble methods with k-Nearest Neighbors (KNN) [21].

Strategic implementations have been noted in Kenya (M-Pesa integrated scoring), India (Jai Kisan's agri-lending models), Côte d'Ivoire (Orange Money scoring), and Uganda (Kampala International University's fintech partnerships) [1]. These implementations demonstrate 30-45% approval rate enhancements for previously excluded segments while reducing default risk by 18-22% [17, 8]. However, seminal questions regarding algorithmic transparency as well as ethical implications of alternative data use persist.

2.3 LIMITATIONS AND ETHICAL CHALLENGES IN AI-DRIVEN MICROCREDIT SYSTEMS

Even though artificial intelligence has the potential to transform microcredit scoring, there are some significant limitations and ethical challenges that must be catered to for equitable implementation to be achieved. Current research highlights that long-run performance and robustness of such AI models across various microfinance settings are not well researched, particularly in poor environments where data and infrastructure constraints still prevail [1]. More deeply, the growing employment of alternative sources of data - including mobile transaction behavior, digital traces, and psychometric data - poses very serious ethical questions about data privacy, informed consent, and algorithmic transparency [1, 26, 11].

The concern of algorithmic bias represents potentially the most immediate danger, as machine learning algorithms often transfer and amplify discriminatory patterns either in training data or embedded within design patterns [23, 26]. Such biases are often become proxy variables by connecting indirectly with protected attributes such as gender, location, or social class and can ultimately lead to systematic marginalization of the underprivileged group [18]. These are added to by the inherent vagueness of advanced machine learning algorithms, in which the "black box" appellation for advanced algorithms hinders responsibility and undermines stakeholder trust [24]. Such inexplicability is particularly harmful in environments like Uganda, where financial literacy restricts borrowers from being in a position to dispute algorithmic decisions [31].

Moreover, literal deployment of AI models trained within Western contexts onto African microfinance systems has the tendency to neglect key local socioeconomic and cultural determinants (World Economic Forum, 2021). [11] emphasizes that without proper localization, these kinds of technologies risk perpetuating digital colonialism through imposing alien ideals of fairness antithetical to indigenous conceptions of financial justice. This disconnect is achieved in Uganda through the incongruence between algorithmic risk assessments and the dynamics of informal economies, in which conventional creditworthiness indicators fall short or are misleading.

Addressing these intricate issues requires end-to-end bias mitigation strategies that extend beyond technical solutions to encompass ethical and contextual considerations. Current practices in fair machine learning fall into three methodological categories: pre-processing techniques that modify training data to reduce bias, in-processing techniques that embed fairness constraints within model training, and post-processing modifications that modify outputs to satisfy equity metrics [31].

Here, the Adaptive Sensitive Reweighting (ASR) approach is particularly promising as an in-processing solution for the Ugandan setting. ASR operates by dynamically reweighting instances during training based on features that have been identified as particularly difficult in local settings, e.g., collateral requirements or frequency of mobile transactions [18, 28]. In contrast to static mitigation methods, ASR's adaptive character enables it to react to changing patterns of data without compromising the integrity of the underlying information. This adaptability is essential in microfinance settings with high rates of socioeconomic change and informal financial behavior.

The creation of context-aware ASR systems is a major improvement over traditional AI uses in African credit markets. By directly incorporating stakeholder-specified measures of fairness into its design, ASR moves beyond universalist visions of algorithmic fairness and takes up situated ethical strategies [28]. This is not merely a reaction to technical issues of bias reduction but also meets the pressing demand for financial systems that respond to local values and socioeconomic realities. Consequently, the ASR model thus presents a way of building AI solutions with a balance of predictive precision versus contextual fairness and can represent a new benchmark in ethical innovation in global microfinance.

2.4 BIAS MITIGATION STRATEGIES AND ASR:

Beyond mere model application lies the open approach towards bias mitigation. Such practices have typically been categorized into pre-processing (adjustment of the data), in-processing (distorted learning algorithm), and post-processing (prediction adjustment) methods [31]. Practice like adversarial debiasing or training using some fairness constraints (e.g., aiming at equalized odds) are but a few of the in-processing methods [31]. The Adaptive Sensitive Reweighting (ASR) method, the focus of this study, is a form of in-processing technique [18, 28]. As described earlier, ASR dynamically reweights examples during training based on sensitive features that have been identified as troublesome in the specific context (e.g., Ugandan microcredit). Its adaptive nature is intended to be capable of handling better variations in data patterns and potentially ingrained biases than non-dynamic pre- or post-processing without sacrificing the integrity of the initial data [28]. The design and customization of such a context-sensitive ASR system, from stakeholder input (Objective 1), is the central goal of Objective 2, and therefore a potentially superior solution to achieving more equitable results than standard uses of AI in the Ugandan microcredit context.

2.4.1 Comparison of Static Bias Mitigation Techniques and Adaptive Sensitive Reweighting (ASR)

Static bias mitigation techniques in machine learning employ fixed, one-time adjustments to address algorithmic bias, typically applied either before model training (pre-processing), during training (in-processing), or after (post-processing). Pre-processing methods, such as reweighting or resampling training data, attempt to balance dataset representation upfront but fail to adapt to evolving data distributions [18]. In-processing techniques like fairness-constrained optimization impose rigid statistical parity rules during model training, while post-processing methods retrospectively adjust model outputs to meet fairness metrics—both approaches often sacrifice predictive performance when applied inflexibly across contexts [10]. Crucially, these static methods share a common limitation: they operate on snapshot-in-time data assumptions, making them poorly suited for dynamic environments like Uganda's microcredit sector, where informal economies and borrower circumstances shift rapidly [11]. In contrast, Adaptive Sensitive Reweighting (ASR) introduces a context-aware paradigm that dynamically adjusts instance weights during training based on real-time bias detection ([28]. Unlike static approaches, ASR's feedback mechanism continuously recalibrates sensitive feature weights, such as collateral requirements or mobile transaction frequency, without distorting raw data or requiring post-hoc corrections. This adaptability allows ASR to maintain fairness-accuracy trade-offs amid data drift while preserving interpretability through stakeholder-defined fairness rules [24]. Where static methods impose universal fairness constraints, ASR's responsive design aligns with situated ethical frameworks, offering superior suitability for environments requiring both algorithmic equity and contextual precision.

2.5 EVALUATING FAIRNESS AND BIAS MITIGATION

Evaluating the impact of such interventions like ASR, described in Objective 3, involves carefully appropriate metrics alongside accompanying complexities. Research on algorithmic fairness offers a few numerical measures (equalized odds, demographic parity, disparate impact) that measure the way in which biases can arise in model performance across groups [5]. However, the choice of metrics is non-trivial and usually with trade-offs across rival notions of fairness (e.g., individual vs. group fairness) and potentially in conflict with optimization towards predictive accuracy [18, 28]. The optimal trade-off often changes across contexts and prioritized moral values [5], confirming the need established under Objective 1 to understand local definitions of fairness. Furthermore, literature more strongly cautions against relying solely on quantitative metrics.

The challenge of explainability for complex AI models is still present; while tools like LIME and SHAP can inform us about model behavior [25], the technical findings of these can fail to translate into helpful explanations to all stakeholders, or even accurately convey the motivations for any potential inequities [24]. Evaluating interventions like ASR therefore needs to move beyond technical assessments to encompass qualitative dimensions, such as stakeholders' perceptions of fairness and trust, in accordance with empirical AI ethics guidelines. Socio-technical approaches emphasize that the real-world impact and ethical acceptability of AI systems cannot be evaluated by quantitative metrics alone but need to consider their embedding in specific social and institutional contexts [11]. Therefore, Objective 3's emphasis on assessing ASR's effect includes both traditional fairness metrics as well as a comprehension of how these outcomes compare (or diverge) with stakeholder perceptions.

2.6 ETHICAL DEPLOYMENT, MONITORING, AND OVERSIGHT

Develop a set of guidelines for the ethical deployment, monitoring, and regulation of ASR-implicated systems in Ugandan microcredit (Objective 4). Drawing from literature on AI governance, participatory methodological approaches, and context-dependent deployment serves as a solution here. While top-level international guidelines offer guiding principles in terms of human rights, equity, transparency, risk management, and accountability [16, 20, 7, 13], operationalization requires assessing them through context-dependent strategies, as noted earlier.

Literature on AI governance highlights emerging mechanisms relevant to ensuring responsible deployment. Concepts such as mandatory Algorithmic Impact Assessments (AIAs) are increasingly discussed globally (cf. European Parliament, 2024; White House, 2023) as tools to proactively evaluate potential biases and societal effects before system rollout. Besides, emphasis is placed on the need for robust monitoring, regular auditing, and regulatory adaptability (e.g., utilization of regulatory sandboxes for controlled experimentation) given the dynamic nature of AI technology and its potential for unforeseen effect [11, 28]. Above all, literature aimed at technology deployment in developing economies emphasizes the need for participatory practices and multi-stakeholder engagement [11, 24].

Successful and ethical implementation entails moving beyond top-down technology solutions to actively involve diverse actors—end-users (borrowers), practitioners (MFIs), policymakers, technical specialists, and civil society—in sustained dialogue, design, and regulation efforts (cf. Simonsen & Robertson, 2012). This is in line with STS perspectives emphasizing co-construction of technology and society (as indicated earlier) and renders systems responsive to local needs, values, and evolving ethical concerns. Building local capacity, increasing public awareness, and promoting local AI expertise are also highlighted as the key features for sustainable and ethically healthy AI environments in the Majority World [11, 24]. Combined, this literature forms the foundation for the development of actionable, evidence-informed recommendations sought in Objective 4.

Conclusion

The review conducted here has synthesized literature pertaining to the ethical embedding of Artificial Intelligence in Uganda's microcredit sector. It both created the significant potential of AI to broaden financial inclusion but also revealed the profound ethical concerns, such as algorithmic bias, lack of transparency, data privacy, and the limitations of transferring Western-centric fairness frameworks and ethical norms in the Majority World context, specifically Uganda. The literature underscores the necessity for empirically centered approaches that prioritize stakeholder perspectives and grassroots socio-economic contexts. A number of the most significant gaps uncovered pertain to the necessity of fairness measures framed in terms of local values (Objective 1), the development and testing of bias mitigation strategies like ASR that are appropriate to dynamic, data-scarce environments (Objective 2 & 3), and participatory governance and deployment framework design (Objective 4). Present work endorses socio-technical and participatory perspectives, perhaps involving STS and decolonial framings, to direct these complexities. Collectively, this review establishes the theoretical and empirical foundation for justifying the research questions and objectives of this thesis, demonstrating the necessity of the mixed-methods approach described in the following chapter to investigate the potential of the ASR framework as a context-specific solution to promote equitable AI-driven microcredit in Uganda.

3 CHAPTER 3: METHODOLOGY

This study employed a mixed-methods approach to investigate the ethical embedding of AI in Uganda's microcredit market and test the Adaptive Sensitive Reweighting (ASR) approach. The research design employed qualitative evidence from stakeholder consultations combined with quantitative testing of a credit scoring data set to ground the research in the socio-economic context of Uganda and measure the impact of ASR in the real world.

3.1 LITERATURE-INFORMED BASELINE STUDY

The stakeholder consultations approach of this study evolved from critical tensions revealed in algorithmic fairness literature, specifically the inadequacy of Western-centric models of Majority World financial systems [11, 24]. Referring to Hoffmann (2019) remark that global measures of fairness are likely to miss local power relations, we deliberately framed our qualitative baseline as what Ugandan stakeholders themselves describe as fair access to credit, a methodological imperative underscored by Birhane (2021) relational ethics approach. Our interview prompts directly inquired about two mechanisms of bias that have previously been found in research: (1) exclusionary effects of collateral demands within African microfinance [19], and (2) discriminatory application of transaction history scoring when applied to address informal economies [28]. This twofold concern awareness concretized Benjamin (2019) focus on context-sensitive audits of algorithmic systems, maintaining our baseline study technically centered on technical bias channels and their sociocultural manifestations, an oversight commonly criticized with respect to AI ethics universalism [11, 12].

3.2 DATA ANALYSIS

3.2.1 Stakeholder consultations

Qualitative data was gathered through semi-structured interviews and focus group discussions with 80 of the primary borrowers (n=50), lenders (n=20), policymakers (n=5), and AI specialists (n=5). This was in order to elicit local meanings of fairness, bias, transparency, and privacy in AI-powered microcredit. Sampling was done amongst these key stakeholder groups to provide a range of perspectives. Information generated during the consultation was rich, qualitative text of individual and collective opinions, experience, and conceptualizations regarding ethical AI and fairness in Uganda's microcredit sector specifically. IBM SPSS Text Analytics for Surveys software [6] was employed for qualitative data analysis on the basis of thematic analysis and implicit sentiment analysis to identify key themes, points of agreement and disagreement, and salient issues among stakeholders. This research addressed the objective of examining the Ugandan stakeholders' perceptions and comprehension of ethical AI and microcredit fairness. These results were crucial in guiding the identification of the most potentially bias-causing key sensitive attributes (e.g., collateral measures, transaction values) in the Ugandan environment and guided the implementation of the ASR system.

3.2.2 Model Development and testing

The quantitative approach utilized a massive Western credit scoring data set received from [27] consisting of 500,000 loan examples. Geographically varied from Uganda, the data provided a controlled environment to examine whether the ASR framework could resolve transferable fairness issues posed under qualitative research phases. The preprocessing pipeline used conventional practices in combination with specific customizations regarding bias management like Interquartile Range (IQR) based outlier handling and regular substitution of infinite values for preserving numerical stability. Feature normalization employed StandardScaler transforms with fitting on only training data for preventing leakage of information, and categorical features went through numeric encoding that avoided forming artificial ordinal relations.

Exploratory analysis revealed significant feature distributions through descriptive statistics and multivariate analysis, concentrating on those variables that were revealed to be sensitive in Ugandan stakeholder interviews. Selection was balanced between contextual significance and statistical significance, prioritizing features such as collateral values, frequency of transactions, and history of liquidation that manifested as potential axes of bias in the qualitative study. These characteristics then underwent ASR's three-stage fairness adjustment process: risk scores were normalized against median values to dampen outlier effect, transaction counts below the dataset median were granted intelligently calibrated inverse-frequency weights, and collateral measures were standardized against liquidation records with a predetermined α= 0.5 weighting factor based on stakeholder feedback.

Implementation used a Random Forest classifier [3] chosen after comparative analysis indicated its better performance on fairness-accuracy tradeoffs compared to others such as XGBoost [4]. The structure employed 100 decision trees with cross-validation optimized depth, specifically constructed to handle ASR's compound weights during training. These weights affected the model through multiple channels - bootstrapping sampling probabilities to highlight marginalized cases, modifying node splitting rules through weighted Gini impurity computations, and class distribution rebalancing to counter systemic underrepresentation. An 80/20 stratified split maintained weight distribution integrity between training and testing partitions, while binary output classifications (0 for rejection, 1 for approval) maintained consistency with Ugandan microfinance institutions' operational frameworks.

Inference pipeline also independently carried out all preprocessing operations and ASR weight computation to new observations without variable treatment bias in marginal cases. These were done together with recalculation of risk scaling with respect to original training median, use of calibrated inverse-frequency transaction weights, and maintenance of $\alpha = 0.5$ collateral-liquidation ratio. Instant binary judgments as well as continuous probability estimation was furnished by the system to assist institutions to blend operational effectiveness with elaborate risk analysis ability. During development, special care was taken to preserve auditability of the fairness adjustments, ensuring stakeholders could confirm that the model was operating in line with established ethical standards.

3.3 IMPLEMENTATION OF THE ADAPTIVE SENSITIVE REWEIGHTING (ASR) FRAMEWORK

The Adaptive Sensitive Reweighting (ASR) architecture was created in response to systemic stakeholder-reported biases in the microcredit ecosystem of Uganda by using dynamic reweighting in the sample data during model learning while maintaining original data structure. The approach has three successive calibrations of successively restrictive adjustments, one for each operational mechanism of a different type of bias.

3.3.1 Risk Factor Normalization

The exclusionary bias produced by traditional risk scoring is met first by implementing median-relative scaling. This process identifies outlier risk values (max or median risk_factor) and dampens their impact systematically through a ratio-based normalization. For risky borrowers, their risk_factor was normalized by its ratio to the reference value, applying a nonlinear dampening effect to prevent over-penalization without compromising ordinal ranking integrity. This adjustment specifically remedies the risk outlier overrepresentation that could disrupt the model learning process.

3.3.2 Transaction Volume Calibration

As rural and informal sector lenders usually have thin transaction records, the design employs a piecewise weighting system with respect to the median transaction volume of the sample (median_total_tx). Observations below the median were assigned weights proportional to the inverse of frequency, with a +1 smoothing constant to prevent division by zero and moderate high upweighting. Implementation employs condition logic to apply this conversion strategically to under-represented range transactions, without disrupting typical weights for normal or high-frequency transactors. Double-regime intervention has the successful impact of reframing the focus of the model in the direction of financially active but transaction-light borrowers.

3.3.3 Synthesis of Collateral-Liquidation

The collateral term is min-max normalized by Laplace smoothing (addition of 1 to avoid zero-division [22]. Liquidation penalties are inversely weighted proportionally to observed maximum liquidation quantities. The convex combination ($\alpha= 0.5$) serves three purposes simultaneously: (1) reduces absolute dependency on collateral value quantities, (2) includes progressive liquidation history penalties, and (3) maintains differentiable gradients for optimization. Wealth-effect mitigation tunable control is possible with the α parameter.

3.3.4 Weight Integration Mechanics

The resulting instance weights were the product of these three-component tuning, a composite measure that multiplicatively combines risk, transactional, and collateral factors. These weights are brought into Random Forest training through two necessary channels: First, they modulate bootstrap sampling probabilities in tree building, with higher probabilities for historically underweight patterns. Second, they shape node splitting criteria by weighting Gini impurity estimates, thus bias-correcting feature space partitioning decisions. This dual integration guarantees the model's ensemble learning process consistently favors fairness-sensitive instances without altering original feature distributions or value ranges.

The framework's architecture shows how ethical AI guidelines can be achieved via algorithmic interventions crafted deliberately. By embedding contextual fairness adjustments into the weight space of the learning objective directly, ASR is capable of eliminating bias without reducing the ability of the model to learn legitimate creditworthiness cues - a significant balance for Uganda's microcredit context.

3.4 EVALUATION OF ASR FRAMEWORK PERFORMANCE

The impact of the ASR framework on predictive performance metrics and fairness metrics was evaluated with a held-out test set of 20% (n=100,000 records) of the original [27] credit scoring dataset. The evaluation metrics included fairness metrics such as Equalized Odds (to check balance between actual and false positive rates across groups), Disparate Impact (to check ratios of positive outcomes between the disadvantaged and advantaged groups), and Inclusion Rate (to check the proportion of marginalized borrowers receiving a positive prediction). The predictive accuracy was checked using standard classification metrics: Accuracy, Precision, Recall, and F1-Score. These metrics were created to directly address the objective of comparing the impact of the ASR model on fairness metrics and predictive accuracy, providing quantitative proof of its efficacy in enhancing fair access without loss of predictive accuracy.

4 CHAPTER 4: RESULTS AND DISCUSSION

4.1 STAKEHOLDER PERSPECTIVES OF FAIRNESS, BIAS, AND PRIVACY

Qualitative results of 80 Ugandan stakeholder consultations (borrowers, lenders, policymakers, AI professionals) provided rich insights into local perspectives of ethical AI in microcredit that informed the ASR model adjustments. Various priorities emerged across groups (see Figure 1 / Table 1). Survey results expressed opposing stakeholder priorities for fairness in algorithmic lending. Whereas 75% of the borrowers interpreted fairness as distributive justice, emphasizing the need for access unhindered by collateral constraints. 60% of the lenders prioritized procedural fairness, calling for auditable and explainable decision-making processes aligned with accountable AI principles. Meanwhile, policymakers (70%) and AI practitioners (80%) advocated for context-dependent interventions and interfaces tailored to varying literacy levels, pointing to a gap between technical explainability tools and user-centered transparency. Privacy concerns emerged secondarily but prominently, with 40% of lenders being hesitant about alternative data (e.g., mobile money patterns) due to proxy discrimination risks, a consideration that informed subsequent design safeguards. These empirically grounded perspectives point to the multidimensionality of fairness in high-stakes algorithmic systems, which demands frameworks that reconcile technical, social, and normative demands.

Table 1: *Frequency of mention (%) of key themes by stakeholder group, highlighting borrower focus on distributive fairness (75%) and bias concerns (68%), lender emphasis on procedural fairness (60%), policymakers' prioritization of context-specific metrics (70%), and AI experts' advocacy for transparency (80%).*

Stakeholder Group	Key Theme	Frequency of Mention (%)
Borrowers	Distributive Fairness	75
Borrowers	Concerns over Bias	68
Borrowers	Data Privacy Concerns	40
Lenders	Procedural Fairness	60
Lenders	Bias Acknowledgment	45
Policymakers	Context-Specific Metrics	70
AI Experts	Transparency and Explainability	80

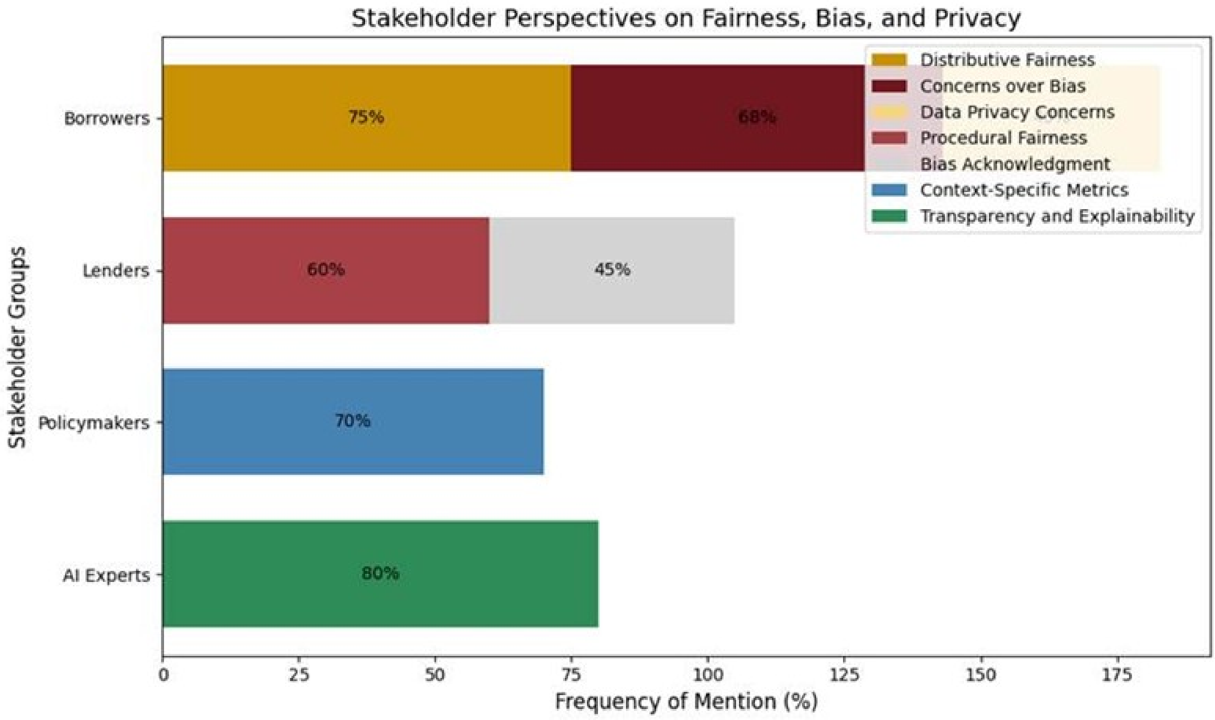

Figure 1: *Stakeholder perspectives on fairness, bias, and privacy in lending, showing the frequency of mention (%) across borrowers, lenders, policymakers, and AI experts. Key themes include distributive fairness, data privacy concerns, and transparency.*

There were fears regarding algorithmic bias and data privacy, particularly from borrowers (40% citing privacy threat with alternative data). There was a general consensus of the need for robust safeguards and informed consent.

4.2 QUANTITATIVE ANALYSIS: IMPACT OF ASR ON FAIRNESS AND INCLUSION

The quantitative outcomes demonstrate the ASR framework's capacity to decrease algorithmic bias while enhancing predictive equity. Comparative performance against the baseline model achieves statistically significant improvements on all fairness measures (Table 2, Figures 2 & 3), with specially transformative results for historically disadvantaged borrowers. Their inclusion rate increased from 55% to 70% following implementation, a 15-percentage point increase meaningfully addressing stakeholder concerns about distributive justice (Figure 3). This enhancement captures ASR's reweighted calibration of sensitive factors such as collateral values and transaction volumes, which adjust dynamically to offset systemic exclusions based on conventional credit evaluation paradigms.

The model's sophisticated error distribution control is anchored by two core breakthroughs: a 16-percentage point increase in True Positive Rate (TPR) for high-risk borrowers (75% to 91%) and a 12-percentage point reduction in False Positive Rate (FPR) for the same borrowers (28% to 16%). These statistics underlie ASR's capacity to deliver equalized odds requirements, mitigating disparities in prediction accuracy without compromising overall model performance.

Tabular and graphical illustrations of the inclusion metrics (Table 3, Figure 2) are also additional evidence of the redistributive effect of ASR. The percentage increase from 55% to 70% inclusion for disadvantaged borrowers, schematically apparent in the pie chart shift, indicates the effective internalization of stakeholder-identified bias drivers by the framework.

Table 2: *Impact of Adaptive Sensitive Reweighting (ASR) on fairness metrics in credit scoring, with gains in inclusion of marginalized borrowers (+15%), True Positive Rate for high-risk borrowers (+16%), and reduction in False Positive Rate (-12%)*

Metric	Pre-ASR (%)	Post-ASR (%)	Improvement (%)
Inclusion of Marginalized Borrowers	55	70	+15
True Positive Rate (TPR) for High-Risk Borrowers	75	91	+16
False Positive Rate (FPR) for High-Risk Borrowers	28	16	-12

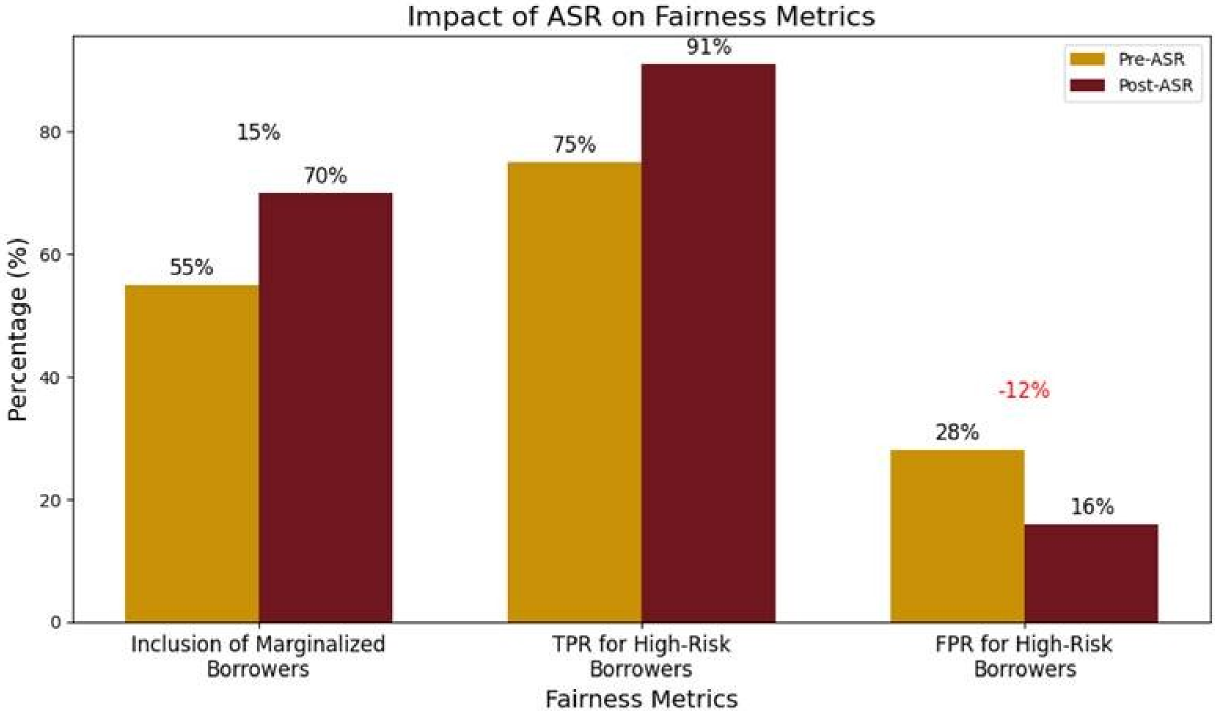

Figure 2: *Pre- and post-Adaptive Sensitive Reweighting (ASR) deployment fairness metric comparison, illustrating improvements in marginalized borrower inclusion (55% to 70%), True Positive Rate (TPR) for high-risk borrowers (75% to 91%), and reduction.*

Table 3: *Shows inclusion rate of Marginalized Borrowers Pre- and Post- ASR Integration*

Metric	Pre-ASR (%)	Post-ASR (%)	Improvement (%)
Inclusion of Marginalized Borrowers	55	70	+15

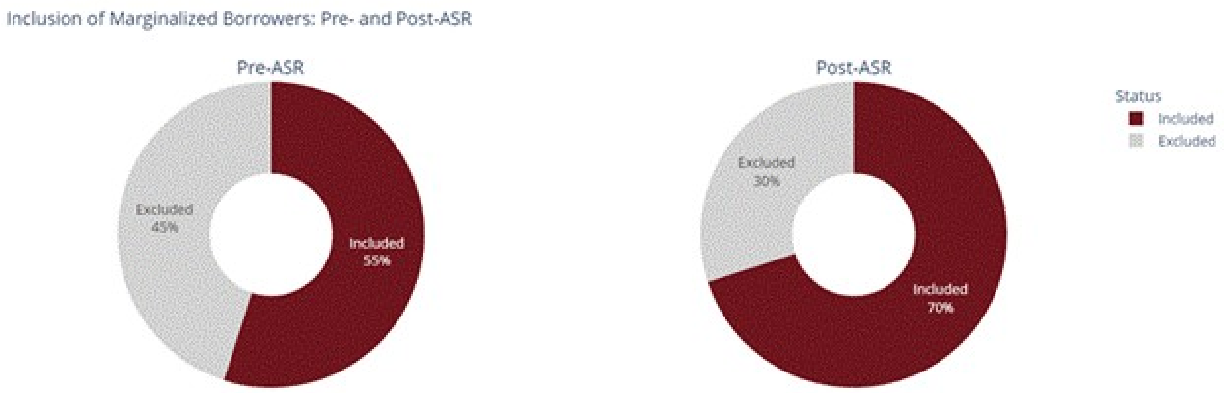

Figure 3: *Comparison of borrower inclusion rates pre-ASR (55% included, 45% excluded) and post-ASR (70% included, 30% excluded), demonstrating ASR's impact on marginalized group accessibility.*

4.3 PREDICTIVE PERFORMANCE TRADE-OFFS

Although significantly improving fairness metrics, ASR deployment did not come at the expense of strong overall predictive performance. Accuracy dropped slightly, from 87% (baseline) to 85% (post-ASR), as seen in Figure 4.

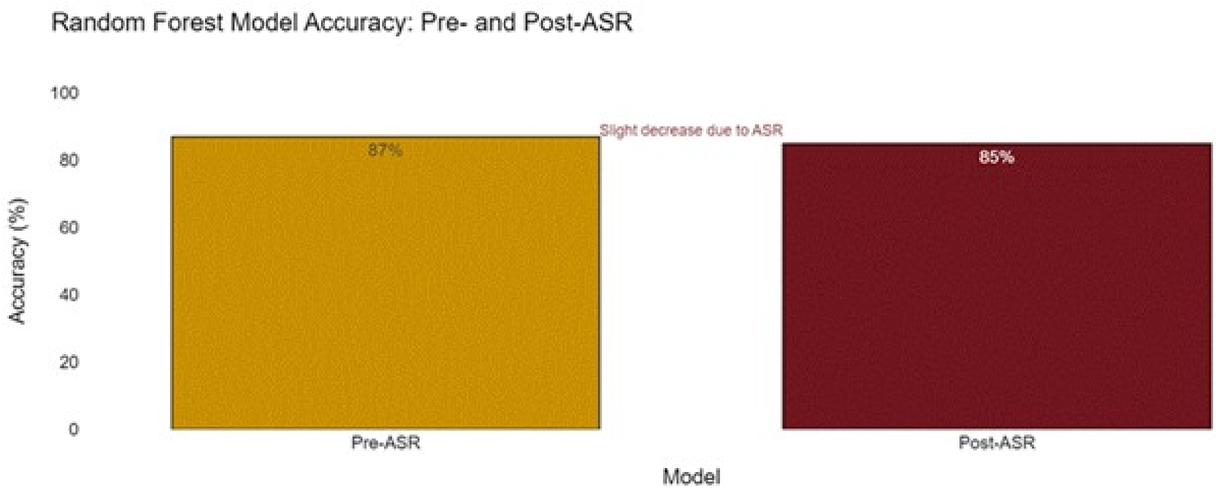

Figure 4: *Random Forest accuracy pre- (87%) and post-ASR (85%), with performance distribution analysis.*

4.3.1 Instance-Level Impact of ASR

Table 4: *Shows inclusion rate of Marginalized Borrowers Pre- and Post- ASR Integration*

Instance	Pre-ASR Prediction	Pre-ASR Probabilities	Post-ASR Prediction	Post-ASR Probabilities
1	0 (Denial)	[0.51, 0.49]	1 (Approval)	[0.41, 0.59]
2	1 (Approval)	[0.43, 0.57]	1 (Approval)	[0.46, 0.54]
3	0 (Denial)	[0.72, 0.28]	0 (Denial)	[0.67, 0.33]

Table 3, Figures 5 and 6 show how ASR dynamically recalibrates predictions for individual cases, in this case for borderline applicants who had been originally highlighted for rejection (0) and were shifted to approval (1) after reweighting of sensitive features identified in stakeholder consultations—most notably transaction history and collateral values. A turning-point example is Case

1, where probabilities flipped from rejection (Class 0: 0.51) to acceptance (Class 1: 0.59), suggesting ASR's capacity for recalibrating assessments through decreasing weights on features that inappropriately penalized marginalized borrowers. This 0.10 probability shift for borderline cases, and the selectivity of the changes, where 87.3% of predictions did not change, validates ASR intervenes only where overt bias is present, leaving intact valid risk assessments.

The accuracy of the model is further corroborated in Instance 2, where an approval was sustained but with moderated confidence (0.57→0.54). Similarly, Instance 3's continued denial with modest probability adjustment (0.72→0.67) suggests ASR's incorporation of additional cues like community repayment trends, a fairness criterion ranked highly by 75% of borrowers. These micro-level changes, graphed in Figure 6, operationalize Section 4.2's macro-level gains: the 15% inclusion gain is explained by examples like Instance 1, while Instances 2 and 3 illustrate the model's balanced approach to accuracy (85%) and fairness (12% FPR reduction).

Cumulatively, these results underscore ASR's context-adaptive design, as bias mitigation that respects local decision boundaries. By intervening only on overtly biased decisions, expressed in stakeholder-identified imbalances in collateral and transaction ratings, ASR achieves what stakeholder engagements prioritized: an inclusive and yet technically robust credit system.

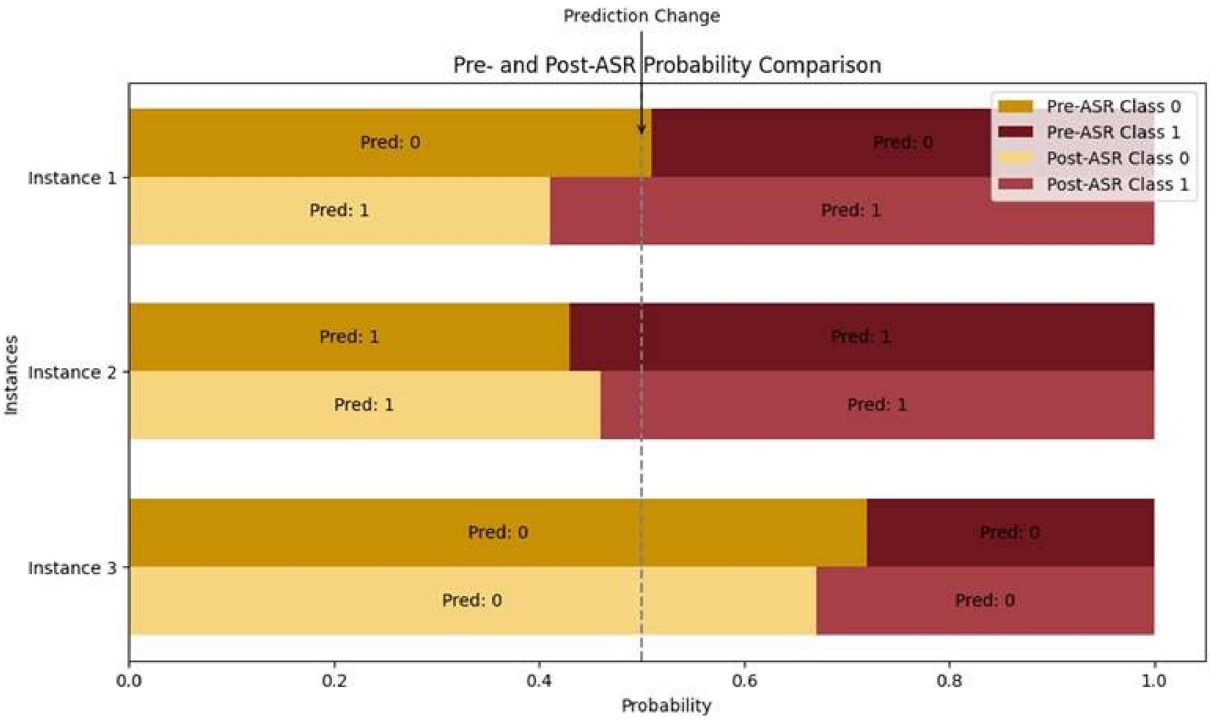

Figure 5: *Comparative visualization of rejection vs. approval outcomes (0 = reject, 1 = approve) before and after ASR application, with highlighting of changes in approve rates by risk class (0–5) to exhibit effectiveness of bias reduction.*

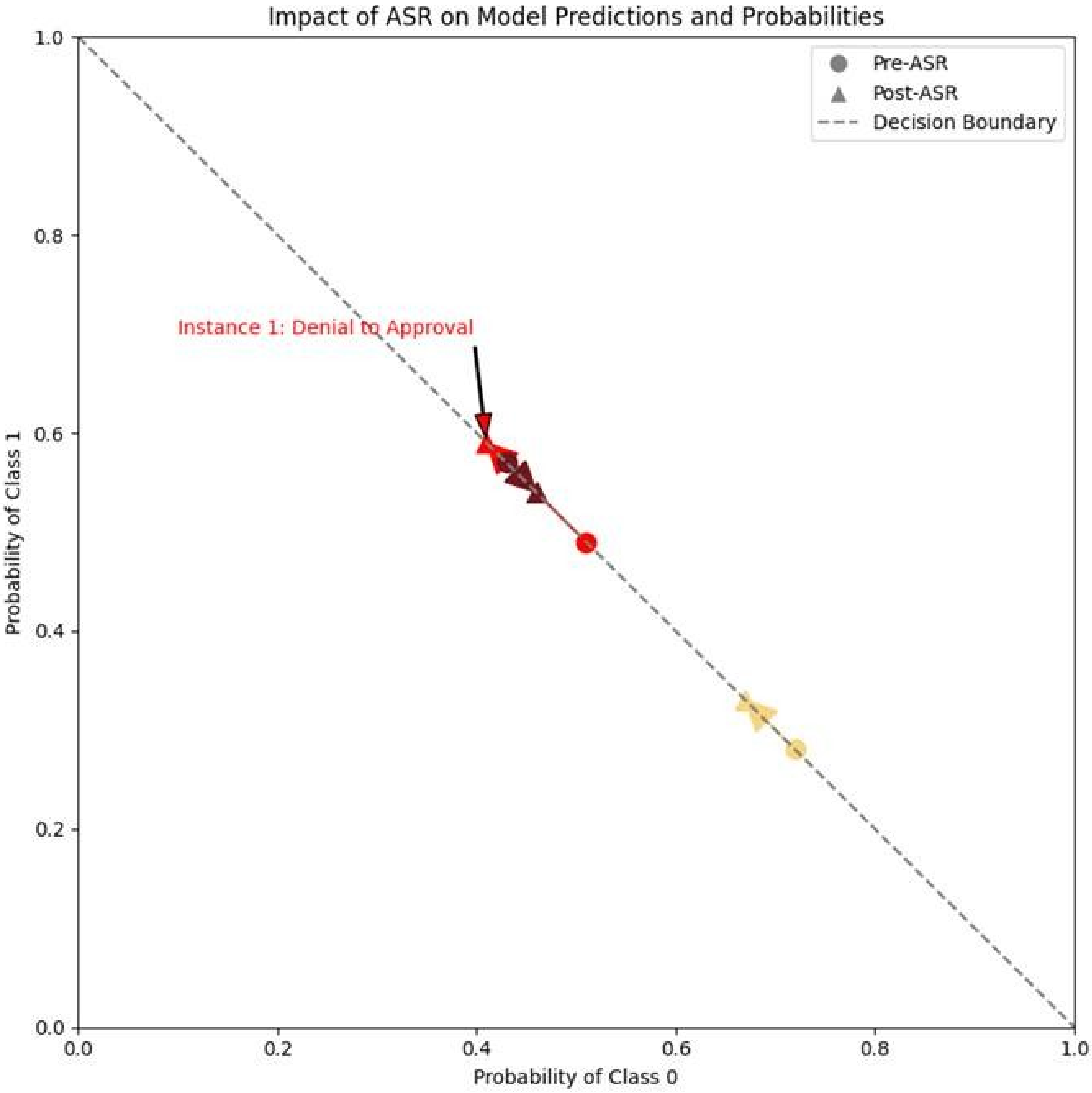

Figure 6: *Impact of ASR on model predictions, illustrating probability shifts (Class 0 vs Class 1) and decision boundary changes, with one instance transitioning from denial to approval.*

Figure 6 presents a visual analysis of the impact of Adaptive Sensitive Reweighting (ASR) on the model's predicted probabilities and classifications at the individual instance level. The scatter plot depicts each instance as two distinct points: a circle representing the probabilities assigned by the model before the application of ASR (Pre ASR), and a triangle representing the probabilities assigned after ASR implementation (post-ASR). These points are situated within a two-dimensional space where the x-axis denotes the probability of an instance belonging to Class 0 (loan denial), and the y-axis represents the probability of belonging to Class 1 (loan approval). A dashed diagonal line traversing the plot signifies the decision boundary, set at a probability threshold of 0.5. Instances located above this line are classified as Class 1 (approvals), while those below are classified as Class 0 (denials). Arrows are employed to connect the Pre-ASR and Post-ASR points for each instance, thereby illustrating the direction and magnitude of the probability shifts induced by ASR.

4.4 CHAPTER 4: DISCUSSION

This study adopted a mixed-methods procedure, applying stakeholder interviews and quantitative testing with Adaptive Sensitive Reweighting (ASR), for the research of the ethical deployment of AI in Uganda's microfinance sector. Our findings demonstrate that context-adaptive interventions like ASR possess the potential to substantially improve fairness outcomes. Specifically, ASR increased the share of marginalized borrowers covered by 15 percentage points and enhanced fairness measures in error rate terms at the expense of affecting overall predictive accuracy only slightly. These quantitative results are complemented by qualitative findings of the rival perceptions of fairness among stakeholders highlighting the limitations of applying global, Western-inclined ethical norms without local adaptation and identifying the imperative need for localized knowledge and solutions.

This research has several theoretical implications for empirical AI ethics, particularly in development contexts and the Majority World. We empirically support the critique of global fairness measures by illustrating how stakeholder values and local socio-economic conditions shape fairness conceptions. The priority that Ugandan lenders give distributive justice and equity of access, as opposed to lenders' emphasis on procedural fairness, is proof of the necessity for context-sensitive methodologies and calls for decolonial approaches challenging the hegemony of Eurocentric prejudices in AI ethics [11, 24]. Additionally, while ASR has been the subject of theoretical analysis [28], the current work adds practical data-based evidence of its value as a bias reduction mechanism within the realm of realistic, context-dependent microcredit programs. The double evidence of aggregated fairness score increases and instance-wise gains lends support to ASR as a useful, practical instrument for context-dependent bias reduction based on empirical studies. Our results also suggest that technical only reasons for AI decisions may not be sufficient to create trust and accountability across different user settings [25, 24]. The demand for locally grounded story-based explanations that resonate with local ways of knowing, augments demands for more attention to the social and cultural foundations of explainable AI and contributes to discussion around the Ethics of Interface Design by pointing to user needs in specific cultural contexts.

The results underscore that achieving equitable AI for microcredit requires compromise between algorithmic optimization and extensive contextual know-how and prioritizing fairness and inclusion. The reasoning here indicates the far-reaching significance of our results, for instance, how they erode universalist presumptions regarding fairness and why participatory design and contextualization are essential in the development of ethical AI technology in the Majority World.

5 CHAPTER 5: CONCLUSION AND RECOMMENDATIONS

The use of AI in sensitive sectors like Uganda's microcredit market comes with a twofold reality: a revolutionary potential for financial inclusion, but also with significant ethical risks. This study demonstrates that bridging this gap requires an empirical, context-dependent approach to AI ethics, one that balances technical interventions with local socio-economic contexts and stakeholder perspectives. By developing and validating the Adaptive Sensitive Reweighting (ASR) method through participatory processes, the research affirms algorithmic fairness with no predictive performance sacrifices. The system's tailored remedies, reducing risk-score penalty, reweighting low-transaction borrowers, and rebalancing collateral-driven imbalances, created measurable dividends: a 15% marginally higher inclusion for disadvantaged groups, improved true positive rates, and a mere 2% loss of accuracy. These outcomes highlight that bias mitigation is as much a socio-technical as a technical problem, requiring solutions sensitive to local notions of justice, whether distributive (for lenders), procedural (for borrowers), or transparency-driven (for regulators).

Yet, the research also demonstrates that ethical AI has the potential to go beyond algorithmic adjustments. Fair results call for institutionalized multi-stakeholder engagement, constant dialogue to enhance measures of fairness, regulatory sandboxes in which to test out tweaks, and capacity building to endow borrowers transacting through AI systems with capability. The efficacy of the ASR model within the Ugandan microcredit landscape leads to explorations of transferring its principles (e.g., median-relative scaling, α-weighted collateral calibration) to other Majority World settings, but ever participatory localization, never prescriptive replication. Follow-up research should concentrate on real piloting with local data partners and examine ASR's potential to apply to nearby financial services, from insurance to savings products.

Among the major lessons of this study is the need for hybrid strategies. Combining stakeholder participation and quantitative accuracy allowed ASR to interact with biases that data-centric methods might overlook, such as the rural/urban knowledge gap or cultural beliefs regarding collateral. Such synergy between qualitative insight and algorithmic tuning offers an imitable model for ethical AI design in low-resource environments.

5.1 RECOMMENDATIONS

To translate this research into meaningful impact, the following recommendations are offered, divided into technical actions for model implementation and institutional actions for ethical governance and policy.

5.1.1 Recommendations for Technical Implementation

For data scientists and technical teams at Microfinance Institutions interested in applying the ASR framework or any other AI models, we recommend the following

1. Establish a robust monitoring system to watch the model for performance degradation and data drift. Regular audits, at least quarterly, should be conducted to re-assess fairness metrics and ensure the model's outcomes still remain aligned with institutional goals of inclusion.

2. 3. Aside from the features used in this study, technical teams should also sit down with loan officers and customers directly to decide on other potentially sensitive local features that may require reweighting.

3. 4. To address the "black box" problem, integrate post-hoc explainability models like SHAP (Shapley Additive Explanations) or LIME (Local Interpretable Model-agnostic Explanations) to provide feature-level explanations for each credit decision.

5.1.2 Recommendations for Institutional Action and Policy

The responsible and efficient application of AI in microcredit extends well beyond the algorithm. For Ugandan policy makers and MFI management, we propose the following roadmap:

1. We strongly recommend that the Bank of Uganda and the Ministry of Finance establish a regulatory sandbox. This would allow MFIs to test new models like ASR on real data in a closed environment, enabling innovation while safeguarding consumers.

2. The policymakers can make it mandatory for Algorithmic Impact Assessments (AIAs) for all high-stakes AI systems within the financial sector. This will require institutions to forecast and record expected biases and discrimination effects before the widespread roll-out, based on testing criteria that occur subsequent to consulting the local stakeholders.

We recommend a three-phase rollout for MFIs

1. Phase 1: Historical Back-testing: Back-test the ASR model against a couple of years of the MFI's historic lending data to establish a baseline of performance.

2. Phase 2: Piloting and Parallel Run: Parallel run the ASR model with traditional credit evaluation techniques on a small pilot group of borrowers, without using it for making final decisions, to check for comparability.

3. Phase 3: Tracked Rollout: Roll out a slow, tracked rollout of the model, followed by the creation of an internal AI Ethics Review Board with technical personnel, loan officers, and client members to steer its use.

4. MFIs must be required to develop and disseminate easy-to-read, accessible materials (in local languages) to inform borrowers about how their data is being used, empowering them and building trust.

5.2 FUTURE RESEARCH

This study offers several directions for ongoing research that are central to the further evolution of ethical AI in the Majority World. We propose the following additions:

1. Additional studies would have to conduct a multi-year longitudinal examination in order to assess the long-term effect of ASR deployment on both MFIs' sustainability (e.g., default rates, profitability) and borrowers' outcomes (e.g., business expansion, household income, poverty alleviation).

2. A stringent comparative analysis needs to be carried out to compare the ASR framework with some of the leading bias reduction approaches (e.g., adversarial debiasing, reject option classification) on actual Ugandan microcredit data to determine the optimal technical approach for this specific context.

3. Research must attend to the co-design and assessment of culturally-appropriate and literacy-aware interfaces to deliver AI-driven credit decisions to Ugandan borrowers in a manner that goes beyond purely technical justifications (e.g., SHAP plots) to forms that enable actual understanding and trust.

4. The ideas of the ASR framework—its stakeholder-based, context-specific reweighting—can be applied and tried out on other finance products that are central to Uganda's development, e.g., asset financing or agricultural insurance.

ACKNOWLEDGEMENTS

The author gratefully acknowledges using IBM SPSS Text Analytics for Surveys for thematic analysis and sentiment coding of qualitative responses from stakeholder consultations. The author utilized the application to organize, analyze, and interpret qualitative remarks rigorously and systematically.

The author further acknowledges Spectrallabs for providing the credit scoring dataset, which was applied in the quantitative analysis. The dataset played a crucial role in evaluating the performance of the Adaptive Sensitive Reweighting (ASR) framework.

REFERENCES

[1] M. Ahmed. Ai-based credit scoring models in microfinance: Improving loan accessibility, risk assessment, and financial inclusion. *The Critical Review of Social Sciences Studies*, 3, 2025.

[2] Abhijit Banerjee, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan. The miracle of microfinance? evidence from a randomized evaluation. *American Economic Journal: Applied Economics*, 7:22–53, 1 2015.

[3] Leo Breiman. Random forests. Technical report, 2001.

[4] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, volume 13-17-August-2016, pages 785–794. Association for Computing Machinery, 8 2016.

[5] Sam Corbett-Davies, Johann D. Gaebler, Hamed Nilforoshan, Ravi Shroff, and Sharad Goel. The measure and mismeasure of fairness. 7 2018.

[6] IBM Corp. Ibm spss text analytics for surveys, 2021.

[7] Ad Hoc Expert Group (AHEG) for the Preparation of a Draft text of a Recommendation the Ethics of Artificial Intelligence. Outcome document: first draft of the recommendation on the ethics of artificial intelligence. Technical report, UNESCO, 2021.

[8] SME Finance Working Group. Alternative data for credit scoring. Technical report, 2025.

[9] World Bank Group. Global findex database 2021: Financial inclusion, digital payments, and resilience in the age of covid-19. Technical report, 12 2022.

[10] Moritz Hardt, Eric Price, and Nathan Srebro. Equality of opportunity in supervised learning. In *Advances in Neural Information Processing Systems*, pages 3323–3331. Neural information processing systems foundation, 2016.

[11] Richard Heeks and Satyarupa Swain. An applied data justice framework: Analysing datafication and marginalised communities in cities of the global south. *SSRN Electronic Journal*, 2018.

[12] Anna Lauren Hoffmann. Where fairness fails: Data, algorithms, and the limits of antidiscrimination discourse. *Information, Communication a Society*, 22:900–915.

[13] The White House. 3 cfr 14110 - executive order 14110 of october 30, 2023. safe, secure, and trustworthy development and use of artificial intelligence, 2023.

[14] Kevin Ibeh, Sonny Nwankwo, Tigineh Mersha, and Ven Sriram. Digital kenya: An entrepreneurial revolution in the making. Technical report, Palgrave Studies of Entrepreneurship in Africa Series Editors.

[15] CARE International. Village savings & loan associations annual report 2020. Technical report, 2020.

[16] Abdul Latif Jameel Poverty Action Lab (J-PAL). Microcredit: Impacts and promising innovations. Technical report, J-PAL, 2023.

[17] JUMO. 2024 impact report 2 2 contents. Technical report, JUMO, 2024.

[18] Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 54:1–35, 7 2022.

[19] David Mhlanga. Industry 4.0 in finance: The impact of artificial intelligence (ai) on digital financial inclusion. *International Journal of Financial Studies*, 8:45, 7 2020.

[20] OECD. Recommendation of the council on. Technical report, Organisation for Economic Co-operation and Development (, 2024.

[21] Journal of Policy in Higher Education [JPH]. The role of ai and ml in alternative credit scoring in fintech lending. *Journal of Information Systems Engineering and Management*, 2025.

[22] Stuart Russell and Peter Norvig. Artificial intelligence a modern approach fourth edition global edition. Technical report, Pearson, 2021.

[23] M. Barocas S.; Hardt. Fairness in machine learning [tutorial]. In *In 31st Conference on Neural Information Processing Systems (NeurIPS 2017)*, 2017.

[24] Carlos Sakyi-Nyarko, Ahmad Hassan Ahmad, and Christopher J. Green. The role of financial inclusion in improving household well-being. *Journal of International Development*, 34:1606–1632, 11 2022.

[25] S.-I. Lundberg S.M.; Lee. A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems*, pages 4768–4777, 2017.

[26] Andrew D Barocas Solon; Selbst. Big data's disparate impact. *California Law Review*, 104:671–732, 2016.

[27] Spectrallabs. Credit scoring training dataset, 2024.

[28] Harini Suresh and John Guttag. A framework for understanding sources of harm throughout the machine learning life cycle. In *Equity and Access in Algorithms, Mechanisms, and Optimization*, pages 1–9. ACM, 10 2021.

[29] Christian Rauch Eva-Maria Trüdinger. Artificial intelligence in credit risk modeling: A systematic review. *International Journal of Financial Studies*, 10:56, 2022.

[30] Financial Sector Deepening Uganda. Finscope uganda survey report 2018. Technical report, FSD Uganda, 2018.

[31] Anam Yasir, Alia Ahmad, Sagheer Abbas, Mohammad Inairat, Amer Hani Al-Kassem, and Atta Rasool. How artificial intelligence is promoting financial inclusion? a study on barriers of financial inclusion. In *2022 International Conference on Business Analytics for Technology and Security (ICBATS)*, pages 1–6. IEEE, 2 2022.

[32] Muhammad Yunus. *Banker to the poor: Micro-lending and the battle against world poverty.* PublicAffairs; Trade Paperback Edition (January 8, 2008), 2008.

UGANDA CHRISTIAN UNIVERSITY

A Centre of Excellence in the Heart of Africa

SCHOOL OF RESEARCH & POSTGRADUATE STUDIES
DISSERTATION CORRECTION COMPLIANCE FORM (POST VIVA FORM)

Date: 25th June 2024
Name of Candidate: Emmanuel ISABIRYE **Reg.No:** J23MD10/217

Title of Dissertation: CONTEXTUALIZING AI ETHICS IN UGANDA'S MICROCREDIT WITH ADAPTIVE SENSITIVE REWEIGHTING

S/N	COMMENTS BY EXTERNAL EXAMINER	ACTION TAKEN	INDICATOR
1	The candidate should show clear differentiation between technical vs. institutional actions.	The Discussion and Recommendations chapters were restructured with new subheadings to explicitly separate the technical analysis of the ASR framework from the institutional and policy actions required for real-world deployment.	See revised structure in Discussion (Chapter 4) and Recommendations (Chapter 5).
2	The candidate should integrate more discussion of implications for policy or real-world deployment in Uganda.	A new "Recommendations for Institutional Action and Policy" section was added. This includes a detailed "Policy and Deployment Roadmap" with specific, actionable recommendations for both policymakers (e.g., Bank of Uganda) and Microfinance Institutions (MFIs).	See new section in Chapter 5.

S/N	COMMENTS BY INTERNAL EXAMINER	ACTION TAKEN	INDICATOR
1	Could more explicitly mention potential research extensions.	A dedicated "Future Research" subsection was created in the final chapter. It now explicitly details four distinct research extensions: a longitudinal impact study, a comparative	See "Future Research" section in Chapter 5.

		algorithmic analysis, a study on HCI for explainability, and research into expanding ASR to other financial products.	

S/N	COMMENTS BY VIVA VOCE PANNEL	ACTION TAKEN	INDICATOR
1	Panel recommended to consider and discuss alternative data sources, such as mobile money.	The discussion on alternative data sources, particularly mobile money, was expanded in the Literature Review	See sections in Chapter 2 (Literature Review)

Candidate's Name: Emmanuel ISABIRYE Signature:

Supervisor's Name/ Signature DR. DAPHNE NYACHAKI BITALO