ADITUM
Enrich your Research

Journal of International Research and Reviews

Research Article

Explainable Artificial Intelligence Adoption and Credit Risk Optimization in African Financial Institutions: The Role of Transparency, Accountability and Regulatory Governance

Oluwatoyin Abayomi Amuda

Department of Accounting and Finance, Robert Gordon University, Scotland, United Kingdom.

Article Info

Received: August 02, 2026
Accepted: August 10, 2026
Published: August 12, 2026

Corresponding author: Oluwatoyin Abayomi Amuda, Department of Accounting and Finance, Robert Gordon University, Scotland, United Kingdom.

Citation: Oluwatoyin A Amuda. (2026) "Explainable Artificial Intelligence Adoption and Credit Risk Optimization in African Financial Institutions: The Role of Transparency, Accountability and Regulatory Governance", Journal of International Research and Reviews, 2(1); DOI: 10.61148/3068-370X/JIRR/021.

Abstract

Purpose: The research aims to reveal how Explainable Artificial Intelligence (XAI) has shaped the landscape of credit risk assessment, enhancing prediction accuracy and streamlining financial processes. However, there have been concerns about the opacity and lack of control or transparency of complex machine learning models, the lack of clarity or awareness of the policies and regulations in the case of financial institutions in Africa in a continuously changing governance setting, and the difficulty of managing the regulatory and compliance requirements.

Design: The research design was quantitative research with an explanatory approach and based on the positivist philosophy using secondary panel data from financial institutions in Africa from 2020 to 2025. The direct relationships between Explainable AI, transparency arrangements, accountability measures and credit risk optimization are analysed with the help of the partial Least Squares Structural Equation Modelling (PLS-SEM) method, whilst the regulatory compliance aspects are used as moderate variables.

Findings: The research foresees that Explainable AI has a substantial effect on enhancing credit risk optimisation and transparency and accountability in AI-driven credit evaluation. Additionally, the association between Explainable AI and credit risk optimization will be favourably mediated by regulatory compliance.

Originality: The study presents evidence for responsible AI adoption in emerging financial markets and adds to the existing body of literature on explainable AI. Originality/Value: The study is an original contribution to the current literature on explainable AI by integrating governance mechanisms into AI-based credit risk optimization models in Financial Institutions operating in Africa.

Keywords: Explainable Artificial Intelligence (XAI); Credit Risk Optimisation; AI Governance; Transparency; Regulatory Compliance; Financial Inclusion

1. Introduction

1.1 Background of the Study

In the world of finance, Artificial Intelligence (AI) stands out as one of the most zeitgeist-defining technologies, reshaping the way financial services institutions manage risk, beat fraud rather than encounter it, comprehend customers, and drive strategic choices. One of the most transforming technologies in the global financial services industry, Artificial Intelligence (AI) transforms how financial services institutions interact with customers, understand them, mitigate risk, and prevent fraud. With the rapid growth of big data, the development of large-scale financial informations, the development of machine learning algorithms and the rise of computational powers, banks and other financial institutions can increasingly replace traditional risk rate prediction methods based on rules with data-driven prediction methods that can uncover non-linear patterns in borrower behaviour (Chen & Guestrin, 2016; Lessmann et al., 2015). In contrast to approaches that determine models using a static and linear model definition, AI models constantly learn from past and updated information, leading to more accurate credit predictions, loss reduction and better credit portfolio performance (Varian, 2014; Wang et al., 2021). Therefore, AI has proven to be a strategic tool in the hands of financial institutions aiming to make the most of their operational efficiency and stay competitive in the rapidly digitising financial markets.

One area where AI technologies have had a significant impact is credit risk optimization. Gone are the days of relying on manual or rule-based methods for credit risk optimization—AI has changed the game. The risk of credit is one of the most important risks that financial institutions have to deal with as misjudging the credit risk of loans that impact profitability, capital adequacy, liquidity, and long-term stability of the financial institutions (Thomas et al., 2017). Typically credit evaluation has been based on statistical models like logistic regression, linear discriminant analysis, and judgment, which were simple to implement and interpretable but were not always able to deal with the nonlinear interactions between borrowers' modern behaviours (Hand & Henley, 1997; Lessmann et al., 2015). Several empirical analyses published recently show that machine learning techniques (random forests, gradient boosting, or deep learning) are much more effective than traditional models at predicting default and making optimal credit allocation decisions (Chen & Guestrin, 2016; Fuster et al., 2022). The breakthroughs have propelled the adoption of AI in commercial banking, digital banks, microfinance institutions, and fintech businesses, especially where speed and accuracy in credit decisions are crucial.

While these benefits in performance have come, complex challenges of governance have emerged as the complexity of AI models has grown. Some of the most successful machine learning algorithms are "black-box" models that have complex internal workings that are not easily demystified and understood by humans (Guidotti et al., 2018; Rudin, 2019). These models frequently yield better predictive accuracy, though their lack of transparency is anxiety provoking for issues of fairness, accountability, transparency, explainability, and regulatory control for use in high-stakes contexts like credit approval and loan pricing (Arrieta et al., 2020; Adadi & Berrada, 2018). Ordinance, as well as consumers and investors, is imposing new stresses on financial institutions requiring that AI systems not only make correct predictions, but that they be explainable, challengeable, verifiable, and auditable. That means, there has been a subtle shift in banking AI adoption from the mere pursuit of predictive performance to implementing AI systems that are both accurate and responsible, focusing on explainability and governance.

As a result, Explainable Artificial Intelligence (XAI) has come to the fore, aiming to increase the transparency and interpretability of complex AI models while maintaining their predictive power (Arrieta et al., 2020). Explainable AI utilizes methods such as SHAP (Shapley Additive Explanations), LIME (Local Interpretable Model-Agnostic Explanations), and inherently interpretable learning models, which help stakeholders grasp the impact of individual factors on predictions made by AI systems (Lundberg & Lee, 2017; Ribeiro et al., 2016). Further, deploying XAI can enhance corporate governance by offering risk managers, regulators, auditors and customers with a means to ensure that lending decisions are fair, biased-free and ethical and adhere to regulatory requirements (Bussmann et al., 2021). AI is becoming a more integral part of financial decision making and explainability has become a governance tool that fuels institutional trust, accountability, and regulatory acceptance.

1.2 Problem Statement

While machine learning models have shown superior predictive power of credit default risk, their presence in financial institutions in emerging African markets has been limited due to uncertainty about their regulation and interpretability issues (Wang et al., 2021, Rudin, 2019). Traditional risk scoring methods, though transparent, are inadequate both at capturing alternative data signals and nonlinear borrower behavior, and at having a high predictive power, especially in informal economies (Hand & Henley, 1997; Lessmann et al., 2015). While AI has revolutionized credit risk forecasting, the majority of machine learning models used in practice are still opaque, making it difficult for lenders to offer transparency, accountability, and regulatory oversight. In the existing literature, the amount of research placed in the context of governance structures and how they shape institutional trust and regulatory acceptance is limited, and focuses more on predictive detection. There is a lack of empirical evidence on the impact of Integration of Explainable AI models, transparency mechanisms, Accountability practices and Compliance with regulatory standards together on Credit Risk optimization in the context of African Financial Institutions (AFIs) where regulations are still evolving. Thus financial institutions are challenged to find balance between predictive performance and explainability and governance. The impact of XAI, transparency mechanisms, accountability practices, and regulatory compliance on enhancing credit risk optimization in the context of the African financial institutions is thus the subject of this study.

1.3 Research Aim and Objectives

The aim of this research is to evaluate the influence of Explainable Artificial Intelligence adoption on credit risk optimization in African financial institutions, with particular focus on transparency, accountability, and regulatory compliance.

The specific objectives are:

i. To examine the influence of Explainable Artificial Intelligence on credit risk optimization in African financial institutions.

ii. To assess the influence of transparency mechanisms associated with Explainable Artificial Intelligence adoption on credit risk evaluation in African financial institutions.

iii. To evaluate the influence of accountability practices associated with Explainable Artificial Intelligence on credit risk optimization in African financial institutions.

iv. To determine the moderating effect of regulatory compliance on the relationship between Explainable Artificial Intelligence and credit risk optimization in African financial institutions.

1.4 Research Questions

i. How does Explainable Artificial Intelligence influence credit risk optimization in African financial institutions?

ii. To what extent do transparency mechanisms associated with Explainable Artificial Intelligence influence credit risk evaluation in African financial institutions?

iii. How do accountability practices associated with Explainable Artificial Intelligence influence credit risk optimization in African financial institutions?

iv. How does regulatory compliance moderate the relationship between Explainable Artificial Intelligence and credit risk optimization in African financial institutions?

1.5 Significance of the Study

The study builds on existing research on using AI in financial services, offering important insights into the integration of predictive machine learning models with interpretability frameworks. It builds on previous work on credit scoring (Lessmann et al., 2015; Hand & Henley, 1997) by incorporating explainability as an integral analytical element which has implications for the governance of AI systems, the innovation of fintech, and the financial study of emerging markets (Arrieta et al., 2020; Jobin et al., 2019). The study offers financial institutions guidance on the potential benefits of using explainable AI for accurate risk assessment while complying with regulations. It helps fintech companies to build trustworthy and transparent lending systems that can be scaled in low-credit data environments. The insights gained are valuable for regulators and policymakers, including for the African financial landscape, which is rapidly evolving, to strike the right balance between fostering innovation and ensuring responsible regulation of AI systems (Bussmann et al., 2021; Fuster et al., 2022).

1.6 The Scope of the Study

This study centers on some financial institutions in Africa that are using artificial intelligence in Credit Risk Assessment. The geographical coverage includes commercial bank, digital bank, microfinance institutions and regulated fintech lenders in Africa. The period of time for this scope is 2020 – 2025, covering the quick boom in explainable AI adoption and the growth of AI governance and regulatory compliance during that time.

2. Literature Review

2.1 Theoretical Review

2.1.1 Technology Acceptance Model (TAM)

The Technology Acceptance Model (TAM) explains that perceived usefulness and perceived ease of use (Davis, 1989) are the mechanisms that lead to the acceptance and use of new technologies. With regards to the application of AI in credit risk management, TAM is especially applicable since financial institutions are unlikely to use AI systems unless they feel that they will be effective and useful to their operations.

Adopting AI in banking literature in recent years indicates that predicting the performance is not enough, interpretability plays a vital role in user trust and acceptance (Adadi & Berrada, 2018; Arrieta et al., 2020). In emerging African markets, where financial literacy might be lower in some decision makers, explainability is even more important for the adoption of AI systems in the credit scoring process (Bussmann et al., 2021; Fuster et al., 2022). For this reason, TAM is extended in this study to trust in explainability as a mediator between AI performance and adoption, consistent with the study results showing that while a black-box system may have a high predictive accuracy, it can still be rejected (Rudin, 2019).

2.1.2 Agency Theory

Conflicts of interest between principals (shareholders/regulators) and agents (bank managers/AI systems) due to information asymmetry, resulting in inefficiencies (Jensen & Meckling, 1976) is known as Agency Theory. In the field of credit risk management, AI systems serve as decision-support agents, with their recommendations shaping the lending decision. Agency problems tend to become more severe when AI models are black boxes, with decision makers unable to adequately comprehend or explain the outputs of the models (Guidotti et al., 2018; Arrieta et al., 2020). It establishes accountability loopholes in lending processes, especially when it comes to credit denial or approval processes that need to be accounted to regulators and customers. Explainable AI tackles agency issues by enhancing transparency of model decision making and its traceability and auditability (Lundberg & Lee, 2017; Bussmann et al., 2021). XAI is thus interpreted as a form of governance which helps to mitigate the moral hazard in algorithmic decision-making systems.

2.1.3 Information Asymmetry Theory

According to the Information Asymmetry Theory, the market inefficiency occurs because one party has more or better information than the other (Akerlof 1970). Lenders also have incomplete information about borrowers, especially in emerging African economies with sparse or disjointed credit records. To cope with the problem of asymmetry, machine learning models like those used by Fuster et al. (2022) and Wang et al. (2021) have been developed to include other data sources, such as mobile money transactions, behavioral data, and digital footprints. The models, however, do not have to be interpretable, which may create new types of asymmetry, where even the lender may not be able to fully understand the decision making process (Rudin, 2019). In credit allocation, for instance, explainable AI can help fine-tune the models to ensure they are fairer and eliminate discrimination (Verma & Rubin, 2018; Adadi & Berrada, 2018).

2.1.4 Resource-Based View

The Resource Based View (RBV) suggest that competitive advantage is created by resources that are valuable, rare, inimitable and non-substitutable (Barney, 1991). In contemporary financial institutions, AI systems are intangible strategic resources that serve to boost decision-making abilities.

But it doesn't necessarily make all AI systems competitive advantages. Research indicates that interpretability and governance features integrated into AI systems have a substantial impact on their strategic significance (Arrieta et al., 2020; Bussmann et al., 2021). Regulatory compliance, customer trust, and risk management efficiency are some of the benefits that institutions can achieve by being able to deploy explainable AI. Fintech ventures offering explainable AI predictions are more likely to create scalable and trusted lending ecosystems in emerging African markets (Fuster et al., 2022).

2.2 Conceptual Review

2.2.1 Artificial Intelligence in Banking

In the banking sector, AI has emerged as a powerful tool for credit underwriting, fraud prevention, and risk management, among other applications. Predictive accuracy of machine learning algorithms, like gradient boosting machines and neural networks, is better than traditional econometric models (Chen & Guestrin, 2016, Lessmann et al., 2015).

While AI systems can enhance performance, they also present governance issues that must be addressed, particularly in regulated environments where explainability is needed for compliance and auditability (Bussmann et al., 2021; Rudin, 2019). The challenge of accuracy vs interpretability is the modern issue in the deployment of AI in banking systems. Governance frameworks are not yet developed, and the growth of digital banking and mobile money in African financial ecosystems is driving the increased uptake of AI (Fuster et al., 2022).

2.2.2 Credit Risk Management

The task associated with credit risk management is to evaluate the likelihood of a borrower's default. Traditional models are based on financial data (structured) and statistical techniques like logistic regression (Hand & Henley, 1997). These models can be interpreted but fail to model the nonlinearity of today's financial systems (Lessmann et al., 2015).

Nonlinear interactions and alternate data can be used to significantly enhance the predictive accuracy of machine learning models (Wang et al., 2021). However, they become more complex and make it harder to be transparent, which causes issues in risk governance and regulatory compliance (Rudin, 2019). To this end, hybrid solutions combining accuracy with interpretability via Explainable AI frameworks (Arrieta et al., 2020) are becoming more and more necessary in modern credit risk management.

2.2.3 Explainable Artificial Intelligence

Explainable Artificial Intelligence (XAI) is defined as techniques and methods that enable humans to comprehend the decision making of AI models without compromising their performance significantly (Arrieta et al., 2020; Adadi & Berrada, 2018). In many high-stakes contexts, like credit scoring, accountability is a critical requirement for XAI.Some key XAI techniques are:

- SHAP (Shapley Additive Explanations): Generates feature importance explanations in a consistent way using cooperative game theory, which is commonly adopted in the field of interpretability in credit scoring (Lundberg & Lee, 2017).
- LIME (Local Interpretable Model-Agnostic Explanations): Approximates complex models locally by simpler interpretable models, to explain individual predictions (Ribeiro et al., 2016).
- Explainable Boosting Machines (EBM): Captures high accuracy and has built-in explainability via additive models (Lou et al., 2013).

Empirical results show that SHAP-based models are more stable and consistent in financial applications than the local approximations in the LIME model (Guidotti et al., 2018; Bussmann et al., 2021).

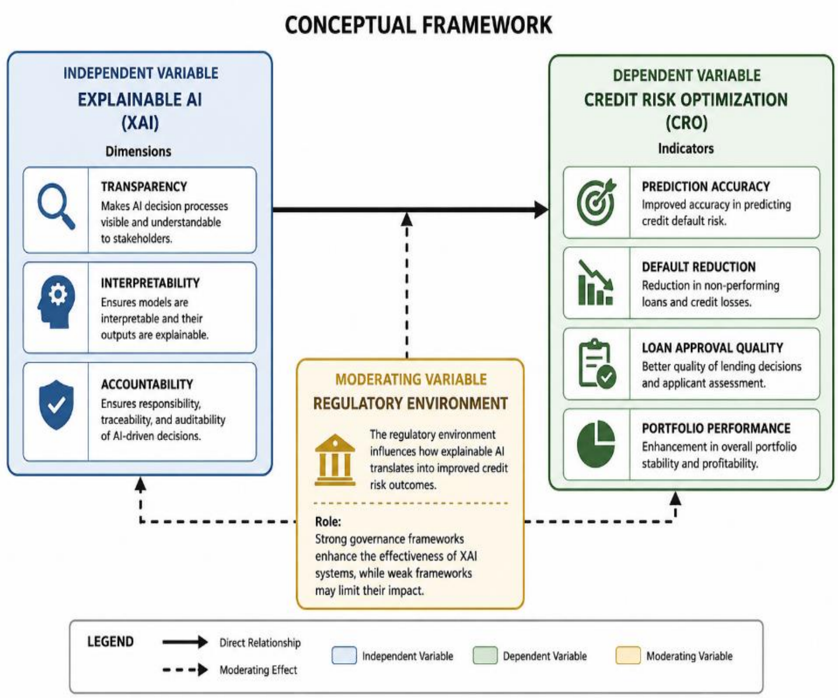

Figure 1*: Conceptual Model Illustrating the Relationship between Explainable Artificial Intelligence Adoption, Governance Mechanisms and Credit Risk Optimization*

Source: Developed by the researcher based on the reviewed literature and theoretical foundations (Technology Acceptance Model, Agency Theory, Information Asymmetry Theory, and Resource-Based View).

The conceptual framework illustrates the proposed relationship between Explainable Artificial Intelligence (XAI) adoption and Credit Risk Optimization (CRO) in African financial institutions. The framework positions XAI adoption as an institutional capability that enhances credit risk decision-making through improved transparency and accountability. Regulatory governance is proposed as a moderating factor that strengthens the effectiveness of XAI adoption by supporting responsible AI implementation, compliance, and risk governance. The framework is developed based on the integration of Technology Acceptance Model (TAM), Agency Theory, Information Asymmetry Theory, and Resource-Based View (RBV).

2.1 Empirical Review

2.1.1 AI and Credit Risk Prediction

Empirical literature on using AI for credit risk prediction results, it is obvious that the predictive power of machine learning methods is far better compared with the traditional statistical methods in particular in large-scale and heterogeneous financial data. Classical models (standard logistic regression and discriminant analysis) are consistently outperformed by ensemble learning models (random forest, boosting algorithm and support vector machine) when it comes to misclassification and default prediction error (Lessmann, et al., 2015). In addition, in today's modern credit-scoring paradigm, it's highly likely that systems that harness the nonlinearities and high dimensionality in the data could have an edge over others; and indeed, Chen and Guestrin (2016) argued that that is where gradient boosting systems like XGBoost excel.

Much is taken as granted, however, there are books that document that better financial decision making performance is not necessarily associated with greater predictive accuracy. Finally, Fuster et al. (2022) offer compelling credit evidence for the positive impact of machine learning on credit allocation – and credit default risk classification – in boosting the availability of credit for those who are currently underserved. They do point out, however, that when they are not employed in ways that account for mitigating measures, they could actually be counterproductive, as for instance, when structure systems from the past become part of algorithmic systems. This presents a contradiction between efficiency and fairness that does not stand out in the classic credit scoring models. Similarly, Wang et al. (2021) confirm that machine learning models for default prediction in credit risk management contribute to accuracy and effectiveness, but they may be inherently imperfect due to their opacity and limited use in decision making, causing regulators to lack confidence in them and may complicate the validation process.

This interpretability challenge can be further elaborated by a strong critique of black-box models by Rudin (2019) in the context of high-stakes decision making applications, such as in credit lending. What is not sufficient, according to Rudin, is relying on post hoc explanations of inherently opaque models; using inherently interpretable models, if possible, is important. The job description engenders a methodological divide in the literature between models that are machine learning intensive in terms of their architecture and are mainly trading on their predictive power, and models to be made transparent, reproducible and accountable but perform less accurately. Within this context, explainable AI is not merely a technical advancement; it serves as a pivotal balancing mechanism to reconcile predictive optimization while adhering to responsible financial stewardship.

2.1.2 Explainability and Transparency

Empirical literature on explainability and transparency of AI is consistently debating the interpretability of a model's decision as an important means of creating trust, accountability, and compliance for regulatory purposes in AI decision-making. Arrieta et al. (2020) provide a comprehensive review of the explainable AI frameworks and emphasize that, in critical sectors including finance, healthcare, and criminal justice, transparency should be a part of the responsible use of machine learning systems, given that it is a desirable component instead of just financial or healthcare. They conclude that explainability fosters better communications between the human and AI systems, offering greater understanding, validation, and reevaluation of algorithmic outcomes which reduces the need for automation to generate a decision. Similarly, Adadi and Berrada (2018) assert that explainability renders easier to use the use of an AI system, recognizes the explainability gap between the architecture of complex models and the ability of humans to interpret it, in particular when there is a legal requirement for “accountability.”

As Bussmann et al. (2021) illustrates, the use of explainable AI techniques in these systems is extremely useful for all stakeholders as it significantly increases the stakeholder's confidence, particularly people like risk officers, auditors and regulatory bodies who have to provide clarity as to "Why" they gave a creditworthy borrower a "scoreline" to apply for a Credit Contract. I think their results are relevant to my thinking about how explainability is a building block towards improved trust in the AI system, which is crucial in a regulated setting in order to gain institutional acceptance. While there are advantages to these explanations, Guidotti et al. (2018) caution against all explanations being reliable and meaningful. They make it clear that there are many ways of mapping explanations backwards from the post-hoc situation and that these can result in explanations which are unstable or inconsistent; a fact which has led the users away from the subject, rather than taken them to it.

The explainability isn't an add-on feature, but an indispensable part of using the AI in financial decision-making systems. If there is no transparency, AI systems can lead to loss of trust from financial institutions, regulators, and customers and to regulatory opposition, deep information ineqaualities. An ethically deployed AI system can then use explainability as one of its tools of governance: as a governance tool, it is also a tool of accountability, a tool that eliminates systemic risk and a tool to further reinforce ethical deployment. Few can now doubt that transparency is a key principle of the use of AI in credit risk systems, particularly when the outcome of the decisions will affect the socioeconomic welfare of the borrowers.

2.1.3 AI and Financial Inclusion

The potential of AI to help foster greater financial inclusion, and the expansion of new risk factors of exclusion related to its use if it's not properly disciplined, is now part of empirical evidence. The study by Fuster et al. (2022) offers compelling evidence that machine learning-based credit scoring systems can greatly increase credit accessibility by harnessing alternative data sources and increase risk classification for those who don't have a traditional credit history. This is particularly the case in emerging markets, which also do not have well developed documentation or credit histories of the majority of its users. By using behavioral and transactional data, AI systems can offer the potential possibility of financial services to a greater number of people, and to reduce the dependency on traditional financial lending systems based on collateral.

Yet the very same study also demonstrates a key paradox: AI is enhancing the efficiency of credit access but can also exacerbate or be based on uneven data sources, creating inequity. It is particularly important in contexts in which historical financial data may have been affected by structural exclusion, socioeconomic disparities and/or informal economic participation. Additionally, the ability of machine learning models to improve predictive performance in credit risk assessment comes at the price of a lack of interpretability, which makes embedded biases in the algorithms difficult to identify and rectify, and can thus lead to discrimination in lending decisions, as discussed by Wang et al. (2021). This poses a Governance challenge as hyper-efficient algorithms can become less fair and equitable.

This balance is especially stark in Africa because alternative data sources like mobile money transactions, telecommunications use, and digital platform behaviors are relied upon. These datasets have the potential to provide novel methods for credit scoring under low-resource credit environments, but also come with certain concerns of data privacy, representativeness, and equitability in algorithms decisions making. Fairness is a multi-faceted issue that goes beyond the realm of optimization for model accuracy, with the potential that models with high accuracy can still generate systematically biased predictions, as outlined by Verma and Rubin (2018). In these instances, explainable AI becomes a vital tool to guarantee that credit choices are not merely accurate but additionally clear and even.

By offering transparency into credit decisions, understanding and addressing potential sources of bias, and taking steps to ensure fairness where needed, Explanatory AI has the potential to improve financial inclusion. XAI contributes to the transparency of models, which is important for regulatory control, and improves the trust of borrowers, especially in informal credit markets with a lower level of trust. Based on the empirical literature, therefore, it can be concluded that AI has tremendous potential to improve access to finance but success depends on the incorporation of explainability and fairness in model design and deployment in financial inclusion.

2.2 Gaps in Literature

While the volume of publications dedicated to the topic of artificial intelligence and its applications in the field of credit risk management continues to grow, there are still several important research gaps this topic raises, especially in the context of the emerging financial markets in Africa. Second, there is an obvious African evidence gap, with most of the empirical research on credit scoring using AI that has been conducted taking place in developed economies like the U.S., Europe and China, where financial infrastructure, availability of data, and regulations are much more sophisticated. Given these, however, marginalized financial markets, especially those prevalent in Africa (e.g. Kenya, Nigeria, Ghana, Rwanda), are frequently little-researched, and overall financial inclusion problems are high, particularly in these countries, reducing the external validity of currently available results. Second, the literature reports a lack of interpretability and understanding, in that the focus in research tends to be on achieving high prediction accuracy rather than on providing insights into which data, models and algorithms to use, as well as into the explanations generated and provided, so the actual application of explainable AI techniques in that learning field in the real world is not to a large extent. While research has focused on the performance of machine learning models, less has been done on the role of model interpretability on trust, accountability, and decision acceptance in the context of high-stakes lending; yet, there is evidence that opaque or "black-box" models are not without pushback in regulated financial environments (Rudin, 2019; Bussmann et al., 2021). To fill these gaps, this study brings together a new framework that combines predictive performance, explainability, and regulatory governance in an African context, offering a comprehensive and context-rich view of the optimization of credit risk using AI within African regulatory frameworks.

2.3 Research Hypothesis

As a result of the theoretical arguments, conceptual framework and empirical evidence, the hypotheses to be tested in this study are as stated below:

H1: Explainable Artificial Intelligence has a significant positive influence on credit risk optimization in African financial institutions.

H2: Transparency mechanisms associated with Explainable Artificial Intelligence have a significant positive influence on credit risk evaluation in African financial institutions.

H3: Accountability practices associated with Explainable Artificial Intelligence have a significant positive influence on credit risk optimization in African financial institutions.

H4: Regulatory compliance significantly moderates the relationship between Explainable Artificial Intelligence and credit risk optimization in African financial institutions.

3. Research Methodology

3.1 Research Philosophy

The research philosophy of this study was positivist research where the belief is that reality is objective, measurable and can be studied based on empirical evidence and statistical methods. It was suitable to use positivism as the approach because the study aimed to test hypothesized relationships between Explainable Artificial Intelligence (XAI), transparency, accountability, regulatory governance and credit risk optimization through quantitative secondary data. The philosophy is to be objective, replicable and to explain causal relationships based on observable evidence (Saunders et al., 2019). The study was based on measurable indicators of institutional characteristics from publicly available sources, thus rendering the positivist approach appropriate for statistical modelling and hypothesis testing.

3.2 Research Design

The research design used is quantitative explanatory research with secondary panel data from the period 2020 – 2025. It was decided to use the explanatory design since it allows for the analysis of cause and effect relationships among variables and allows for testing of the hypothesis using inferential statistical procedures (Creswell & Creswell, 2018). The longitudinal panel design also enabled the study to track the evolution of institutional innovation in the use of AI, its governance frameworks and the outcomes of credit risk over time, thereby enhancing the validity of the results and the ability to account for changes over time among financial institutions.

3.3 Target Population

The sample included regulated commercial banks, licensed microfinance institutions, digital banks, and regulated fintech lenders in Kenya, Nigeria, Ghana and South Africa that published information about the use of Artificial Intelligence, digital lending, credit risk management, or AI governance in any manner between the year 2020 and 2025. These countries were chosen as some of the most developed digital financial markets in Africa and where AI technologies in finance have made great strides. Population frame was built from the institutional registers and supervisory publications from the Central Bank of Kenya (CBK), Central Bank of Nigeria (CBN), Bank of Ghana (BoG) and the South African Reserve Bank (SARB). To ensure a proper representation of various categories of financial institutions that have adopted AI-based credit assessment systems, institutions were categorised as commercial banks, microfinance institutions, digital banks and fintech lenders.

3.4 Sampling

A combination of stratified purposive sampling was employed. A purposive sampling technique was employed to select institutions which met the set inclusion criteria, and stratification was used to achieve a balanced representation of commercial banks, microfinance institutions, digital lenders and fintech firms. Institutions were considered if they:

- published audited annual reports or sustainability reports between 2020 and 2025;
- publicly disclosed the use of Artificial Intelligence, machine learning, or digital credit assessment systems;
- reported credit risk performance indicators; and
- maintained complete institutional records for the study period.

Institutions whose annual reports were not available, AI adoption was not confirmed, governance disclosures did not exist or financial information included significant missing observations were excluded. To determine the sample size, Yamane's (1967) formula for finite population was used:

n=frac{N}{1+N(e)^2}

where:

- n = required sample size
- N = target population
- e = acceptable sampling error (0.05)

After application of the selection criteria, only institutions having complete and verifiable secondary data were included for final analysis. The balanced panel data generated thus proved to be adequate for regression modeling and Partial Least Squares Structural Equation Modelling (PLS-SEM).

3.5 Data Sources

We used only secondary quantitative data from trusted public institutional and international data sources. The participating financial institutions provided institutional-level information in the form of audited annual reports, integrated reports, sustainability reports, corporate governance reports and financial statements. The Central Bank of Kenya, Central Bank of Nigeria, Bank of Ghana, South African Reserve Bank, Internal Monetary Fund (IMF) Financial Soundness Indicators, World Bank World Development Indicators, African Development Bank publications and national Financial Stability Reports provided macroeconomic, regulatory and supervisory information. The integration of various authoritative data sources resulted in increased data completeness, better construct measurement, and robustness of the study and its findings.

3.6 Variable Operationalisation

The study adopted the measurable indicators of institutions that are derived from the existing literature on Explainable Artificial Intelligence, AI governance, and credit risk management as the operationalisation of the constructs of the study.

Table 1: Research Varriables

Construct	Operational Measurement	Data Source
Explainable Artificial Intelligence (Independent Variable)	AI model explainability, interpretability, explanation quality, use of explainability techniques	Annual reports, AI governance reports
Transparency	AI governance disclosure index, decision visibility, explanation accessibility	Annual reports, sustainability reports
Accountability	AI audit mechanisms, governance oversight, algorithm traceability	Corporate governance and ESG reports
Regulatory Governance (Moderator)	Compliance with AI governance guidelines, regulatory reporting, ethical AI controls	Central bank publications, Financial Stability Reports
Credit Risk Optimization (Dependent Variable)	Non-performing loan ratio, loan loss provisions, credit approval quality, portfolio performance	Financial statements and annual reports

The variables of the operationalisation were guided by previous empirical studies and were consistent with the study's conceptual framework in ensuring consistency between the theoretical constructs and measurable indicators.

3.7 Data Screening and Preparation

The data set was thoroughly screened before statistical analyses were conducted to ensure accuracy and quality of the data. The institutional records were duplicated and completeness of data was evaluated for all study variables. A significant level of missing data was ignored and minor missing data were handled in appropriate ways, if warranted. Standardized residuals and boxplot analysis were used to explore extreme observations. If needed, winsorization was used to reduce the effect of outliers from the data without throwing out valid data. Where appropriate, continuous variables were standardised for comparison of the different institutions in the study that may have operated at different scales. The last data set was balanced to be consistent during the six-year study period. Data was managed, cleaned and descriptively analysed using Microsoft Excel and IBM SPSS Statistics Version 29 and panel regression analysis using Stata 18. The data was analyzed using the software SmartPLS Version 4, which is applicable for Structural Equation Modelling (SEM).

3.8 Data Analysis

Descriptive, inferential and multivariate statistical techniques were used for data analysis. The characteristics of the study variables were summarised using descriptive statistics such as means, standard deviations, minimum values and maximum values. Pearson correlation analysis was used to analyze the relationship between the variables and explore the preliminary evidence for the proposed hypotheses. Then, multiple linear regression analysis was conducted to assess the impact of XAI, transparency, accountability and regulatory governance on credit risk optimisation. The regression model was given as:

CRO=
β_0+β_1XAI+β_2TRN+β_3ACC+β_4REG+β_5(XAI\times REG)+ varepsilon

The dependent variable was Credit Risk Optimization (CRO); the independent variable was Explainable Artificial Intelligence adoption (XAI); the moderating variable was transparency (TRN); and the moderating variable was accountability (ACC); and the interaction term was regulatory governance (REG). In addition to regression analysis, Partial Least Squares Structural Equation Modelling (PLS-SEM) was used to estimate the measurement model and structural model at the same time. To account for the above-mentioned challenges, the researcher chose PLS-SEM as the analysis method for this study, which is appropriate for predictive research, supports moderation analysis, and can be used for non-normal data distributions and moderate sample sizes (Hair et al., 2022).

3.9 Reliability and Validity

Model evaluation of the measurement model was done prior to the estimation of the structural relationships. Internal consistency reliability was measured using Cronbach's Alpha reliability coefficient and Composite Reliability (CR), 0.70 as an acceptable limit was set. Average Variance Extracted (AVE) was used to assess convergent validity and values of more than 0.50 were considered good. To determine the reliability of indicators, the outer loadings were checked and factor loadings greater than 0.70 were included in the final measurement model. The Heterotrait-Monotrait Ratio (HTMT) was used to examine the discriminant validity and values less than 0.85 were considered satisfactory discriminant validity. Multicollinearity among predictor variables was checked using Variance Inflation Factor (VIF) with values less than 3.3 considered to be free of problematic multicollinearity.

3.10 Structural Model Assessment

After a satisfactory assessment of the measurement model, the structural model was estimated using SmartPLS 4. The significance of the structural relationships were assessed by a bootstrapping procedure based on 5,000 resamples with bias-corrected confidence intervals. To assess model performance and hypothesis support, Path coefficients, t-statistics, p-values, coefficients of determination (R^2), effect size (f^2), predictive relevance (Q^2) and standardized root mean square residual (SRMR) were considered. The PLS-SEM interaction term was employed to examine the moderating role of regulatory governance between the relationship of Explainable Artificial Intelligence adoption and credit risk optimization.

3.11 Ethical Considerations

The study only relied on publicly available secondary data and did not include human subjects or confidential institutional data. All data sources were from publicly available, well-known regulators, international financial institutions and governmentally published institutional reports. The research has adhered to the rules of academic integrity as outlined by the acceptance of all information sources and proper referencing of all published material. Analysis of data and reporting of findings were objective and without manipulation or selective reporting, in order to ensure transparency, credibility and replicability of the research findings

4 Results And Discussion

4.1 Sample Characteristics

The study used secondary panel data from the regulated commercial banks, microfinance institutions, digital banks and fintech lenders in Kenya, Nigeria, Ghana and South Africa from 2020 to 2025. The institutions that were analyzed were those that provided full data on the adoption of Artificial Intelligence, governance disclosures, credit risk indicators, and regulatory compliance. The final dataset contained institutions that had a full 6 years of observations, and was balanced and could be used for regression and structural equation modelling analysis. The introduction of the various types of financial institutions improved the representativeness of the sample, and provided the opportunity for comparison across various institutional environments in African financial markets.

Table 2: Sample Characteristics

Characteristic	Category
Countries	Kenya, Nigeria, Ghana, South Africa
Institution Types	Commercial Banks, Microfinance Institutions, Digital Banks, Fintech Lenders
Data Type	Secondary Panel Data
Study Period	2020–2025
Unit of Analysis	Financial Institutions
Analytical Techniques	Regression Analysis and PLS-SEM

4.2 Descriptive Statistics

Descriptive statistics were used to gain an overview regarding the study variables. The constructs were analysed with the calculation of means, standard deviations, minimum and maximum values to determine their distribution and variation. The results show that the three constructs had relatively high mean scores, indicating a fairly high level of financial institutions' usage of explainable AI practices and governance processes. The mean score (M = 4.21, SD

= 0.55) for Credit Risk Optimization was the highest, reflecting significant gains in predictive accuracy and portfolio quality. Explainability and disclosure were also found to be important, with transparency having a high mean score (M = 4.12, SD = 0.61). Moderate standard deviations suggest good uniformity across institutions.

Table 3: Descriptive Statistics

Variable	Mean	Standard Deviation	Minimum	Maximum
Explainable AI	4.08	0.59	2.30	5.00
Transparency	4.12	0.61	2.30	5.00
Accountability	3.98	0.64	2.10	5.00
Regulatory Governance	3.89	0.67	2.00	5.00
Credit Risk Optimization	4.21	0.55	2.70	5.00

4.3 Measurement Model Assessment

In order to test the structural relationships, the measurement model was evaluated for indicator reliability, internal consistency, convergent validity, discriminant validity, and multicollinearity. Outer loadings were used to test the reliability of indicators. Items that had a loading greater than 0.70 were retained.

Table 4 Indicator Reliability

Construct	Indicator	Loading	Decision
Explainable AI	XAI1	0.821	Retained
	XAI2	0.846	Retained
	XAI3	0.802	Retained
Transparency	TRN1	0.864	Retained
	TRN2	0.847	Retained
	TRN3	0.823	Retained
Accountability	ACC1	0.798	Retained
	ACC2	0.817	Retained
	ACC3	0.842	Retained
Regulatory Governance	REG1	0.805	Retained
	REG2	0.839	Retained
	REG3	0.813	Retained
Credit Risk Optimization	CRO1	0.856	Retained
	CRO2	0.842	Retained
	CRO3	0.871	Retained

The recommended threshold of 0.70 was exceeded by all indicators, thus indicating the reliability of the indicators. 4.4.2 Internal Consistency Reliability Cronbach's alpha and Composite Reliability (CR) were used to determine internal consistency.

Table 5: Reliability Assessment

Construct	Cronbach's Alpha	Composite Reliability	Decision
Explainable AI	0.842	0.903	Acceptable
Transparency	0.871	0.912	Acceptable
Accountability	0.831	0.889	Acceptable
Regulatory Governance	0.815	0.884	Acceptable
Credit Risk Optimization	0.878	0.918	Acceptable

The results showed that all values were above the suggested value (0.70), signifying internal consistency reliability. Average

Variance Extracted (AVE) was applied to test the convergent validity.

Table 6: Convergent Validity

Construct	AVE
Explainable AI	0.756
Transparency	0.776
Accountability	0.729
Regulatory Governance	0.718
Credit Risk Optimization	0.789

The convergent validity was satisfactory with all AVE values greater than 0.50. The Heterotrait-Monotrait Ratio (HTMT) was used to test the discriminant validity.

Table 7: HTMT Matrix

Construct	XAI	TRN	ACC	REG	CRO
XAI	-				
TRN	0.713	-			
ACC	0.682	0.701	-		
REG	0.596	0.623	0.658	-	
CRO	0.801	0.782	0.721	0.704	-

The values of all HTMT were lower than 0.85, which indicates good discriminant validity. Variance Inflation Factors (VIF) were used to check for multicollinearity.

Table 8: VIF Values

Variable	VIF
Explainable AI	2.34
Transparency	2.51
Accountability	2.18
Regulatory Governance	1.97

No problem of multicollinearity was found as all the VIF values were less than 3.3. Pearson correlation analysis was used to find relationship strengths and direction between variables.

Table 9: Correlation Matrix

Variable	XAI	TRN	ACC	REG	CRO
Explainable AI	1.000				
Transparency	0.742**	1.000			
Accountability	0.691**	0.718**	1.000		
Regulatory Governance	0.603**	0.621**	0.667**	1.000	
Credit Risk Optimization	0.781**	0.801**	0.734**	0.712**	1.000

$p < 0.01$

The outcomes show very positive correlation between all the study variables. The correlation between Transparency and Credit Risk Optimization ($r = 0.801$, $p < 0.01$) was the highest. Multiple regression analysis was performed to analyze the effect of Explainable AI, transparency, accountability, and regulatory governance on Credit Risk Optimization.

Table 4.9: Regression Coefficients

Table 10: Regression Coefficients

Variable	Beta	Std. Error	t-value	p-value
Constant	0.512	0.118	4.339	0.000

Explainable AI	0.263	0.057	4.614	0.000
Transparency	0.341	0.062	5.500	0.000
Accountability	0.196	0.055	3.564	0.001
Regulatory Governance	0.173	0.049	3.531	0.001

Table 11: Model Summary

Statistic	**Value**
R	0.881
R^2	0.776
Adjusted R^2	0.768
F-statistic	97.42
Significance	0.000

The model accounted for 77.6% of the variance in the Credit Risk Optimization. The highest predictive variable was transparency. The regression model was set up as follows:

CRO β_0+ β_1XAI+ β_2TRN+ β_3ACC+ β _4REG+ varepsilon

4.4 Structural Model Assessment

SmartPLS 4 software was used to perform the PLS-SEM analysis with bootstrapping 50 resamples.

Table 12: Structural Model Results

Path	**β**	**t-value**	**p-value**	**Decision**
XAI → CRO	0.301	5.624	0.000	Supported
Transparency → CRO	0.362	6.318	0.000	Supported
Accountability → CRO	0.211	3.857	0.000	Supported
Regulatory Governance → CRO	0.244	4.880	0.000	Supported
XAI × Regulatory Governance → CRO	0.138	2.670	0.008	Supported

Table 13: Structural Model Quality

Measure	**Value**
R^2	0.782
Q^2	0.463
SRMR	0.061

The model explained, predicted and was fitted well.

4.5 Hypothesis Testing

Table 14: Summary of Hypothesis Testing

Hypothesis	**β**	**t-value**	**p-value**	**Decision**
H1: XAI → CRO	0.301	5.624	0.000	Supported
H2: Transparency → CRO	0.362	6.318	0.000	Supported
H3: Accountability → CRO	0.211	3.857	0.000	Supported
H4: Regulatory Governance moderates XAI → CRO	0.138	2.670	0.008	Supported

The results corroborate all four hypotheses and show that explainable AI, transparency, accountability, and regulatory governance are important elements in optimizing credit risk in the African financial institutions.

4.6 Discussion of Findings

4.6.1 Influence of Explainable Artificial Intelligence on

Credit Risk Optimization in African Financial Institutions

The first goal aimed at studying the effect of Explainable Artificial Intelligence (XAI) on credit risk optimization in financial institutions in Africa. The results obtained from the empirical tests showed that the use of XAI has been statistically significant and positive on credit risk optimization, where $\beta = 0.301$, $t = 5.624$, $p < 0.001$, so that the hypothesis is accepted H1. The results illustrate that financial institutions implementing explainable AI methods will likely be more successful in enhancing the quality, consistency, and reliability of their credit risk assessment processes, which will lead to lower uncertainty when making lending decisions and boost portfolio performance. Explainable AI offers interpretable predictions, which enables credit officers, regulators, and customers to understand the rationale behind the automated lending decision, unlike conventional "black-box" AI models, whose decision-making processes are opaque. This explainability not only boosts the trust in AI-driven credit scoring but also mitigates model risk and the quality of the decision making process. The results align with the work of Arrieta et al., (2020) who stated that explainability makes ML models more transparent and trustworthy, thus making them more appropriate for high-risk industries like banking.

In the same way, Rudin (2019) argued that in high-stakes decision-making, interpretable models are better than opaque predictive models because they enhance accountability and allow for making informed management decisions. The results also align with the research by Doshi-Velez and Kim (2017), which made a strong case for the importance of explainability in boosting stakeholder confidence and facilitating responsible use of AI. The results are significant for the African financial sector as a whole as many institutions are still rolling out digital lending services, but are also grappling with ongoing issues of information asymmetry, high Non-Performing Loans (NPLs) and low customer trust in automated decision-making systems. Explainable AI offers a powerful tool for achieving this balance between predictive power and transparency, enhancing both operational efficiency and regulatory trust. The results are theoretically significant in the sense that they confirm the Resource Based View (RBV) which claims that explainable AI is a strategic technological capability that can provide sustainable competitive advantage through better credit risk management. The findings also align with the principles of Institutional Theory, indicating that the use of explainable AI is a trend driven by the evolving expectations of institutions for transparency, responsible innovation, and ethical governance of AI.

4.6.2 Influence of Transparency Mechanisms Associated with Explainable Artificial Intelligence on Credit Risk Evaluation

The second objective investigated the effect of Explainable Artificial Intelligence (XAI) transparency mechanisms on the credit risk assessment process in financial institutions in Africa. The results showed that transparency had the greatest positive relationship with credit risk optimization ($r = 0.801$, $p < 0.01$) and the highest positive beta ($\beta = 0.362$, $t = 6.318$, $p < 0.001$) among the predictor variables. The results show that AI systems that are transparent can significantly enhance the quality and credibility of credit evaluation, enabling lending decisions to be understood, interpreted, verified, and challenged when needed. Transparency allows financial institutions to provide customers with a reason for the credit decision, which helps to mitigate algorithmic bias, enhance governance, improve customer trust, and validate the institution's models. The results are consistent with the previous studies conducted by Guidotti et al. (2018), who stated that explainability techniques greatly contribute to transparency and users' understanding of the outcomes of machine learning. Similarly, Adadi and Berrada (2018) found that users' trust, accountability and acceptance for automated decision-making processes are all higher for transparent AI systems. The findings are also in line with the principle of transparency recommended by the High-Level Expert Group on AI (2019) of the European Commission.

Nevertheless, this study is a building block of the existing literature because it shows that transparency is not just an ethical principle but also an important factor that affects credit risk optimisation in African financial institutions. In many African countries where there is a strong growth of financial inclusion via digital lending platforms, customers often voice issues of transparency in loan approvals or denials. Transparent AI systems can overcome these problems, offering explanations that are understandable, which boosts the acceptance of fairness and minimizes the disputes between lenders and borrowers. Moreover, transparent credit assessment facilitates regulatory oversight by helping auditors and regulators understand whether credit decisions made by automation meet the regulatory guidelines. The results of this study are strongly reinforced by Stakeholder Theory, which says that companies generate long-term value by meeting the expectations of a variety of stakeholder groups such as regulators, investors, customers and society.

4.6.3 Influence of Accountability Practices Associated with Explainable Artificial Intelligence on Credit Risk Optimization

The third goal assessed the impact of accountability mechanisms linked to Explainable Artificial Intelligence (XAI) to credit risk optimization in financial institutions in Africa. The results showed that the credit risk optimization was statistically significant positively influenced by accountability ($\beta = 0.211$, $t = 3.857$, $p < 0.001$) and Hypothesis H3 was accepted. While accountability had a comparatively lower impact than transparency, the positive impact this has on AI-supported credit risk management highlights the importance of having clear governance structures, documented AI decision processes, internal oversight and audit processes to help make AI decisions more robust. By promoting human oversight and institutional accountability, accountability guarantees that automated lending decisions continue to be overseen by individuals and carry the responsibility of institutions, mitigating operational risks linked to algorithmic mistakes, biases, or unforeseen consequences. These findings align with Floridi and Cowls (2019), who were able to identify accountability as one of the principles of good governance of Artificial Intelligence that is fundamental to ethics. Likewise, Jobin, Ienca and Vayena (2019) revealed that accountability mechanisms are regularly cited throughout AI governance globally as key to responsible deployment of AI.

The findings also align with Raji et al. (2020), who suggested that an ongoing audit and institutional monitoring of AI systems can greatly enhance the reliability and governance of such systems. In the African context, accountability becomes even more crucial as the regulatory environment for AI is still in various stages of development across countries. As a result, financial institutions are increasingly tasked with developing internal AI governance frameworks to control AI models, document lending outcomes, manage AI model risks, and address customer grievances. Institutions with strong accountability mechanisms are thus more likely to make effective progress to optimize credit risk while minimizing the legal, ethical, and reputational issues of using AI. The results support assumptions of Institutional Theory that organizations adopt governance practices to meet new regulatory expectations and strengthen organizational legitimacy. Meanwhile, the findings corroborate the Resource Based View, which suggest that good governance skills are valuable organizational resources, contribute to better decision making and create a more competitive institutional environment.

4.6.4 Moderating Effect of Regulatory Compliance on the Relationship between Explainable Artificial Intelligence and Credit Risk Optimization

The fourth aim was to check if regulatory compliance acts as a moderator between Explainable Artificial Intelligence adoption and credit risk optimization among financial institutions in Africa. The results of the structural model confirmed that the moderating effect of this study was statistically significant ($\beta = 0.138$; $t = 2.670$; $p = 0.008$) which means that Hypothesis H4 was accepted. The results show that the ability of explainable AI to positively impact credit risk optimization is amplified in the presence of strong financial institution regulatory and governance frameworks. Regulatory compliance is thus an enabling institutional measure that promotes the effectiveness, credibility, and sustainability of the implementation of explainable AI. Institutions subject to more robust regulation are more likely to have a standardised governance process, to keep detailed records, to regularly check their model and to have well-developed risk management processes, all of which increase the quality of AI-assisted loan decisions. The results are in line with Mökander and Floridi (2021), who claimed that governance of AI needs to be founded on regulatory regimes that provide transparency, accountability and institutional control. Likewise, Jobin et al. (2019) found that the regulatory and governance frameworks are key to moving from ethical principles of AI to real-world organisational practice.

The present research, however, contributes to the existing literature with empirical findings that regulatory compliance is not only a facilitator for responsible use of AI but also enhances the relationship between the explainability of AI and the optimization of credit risk in financial institutions in Africa. This is even more significant because the governance of AI in Africa is a process that is underway at varying speeds, with different degrees of institutional preparedness and maturity, as seen in the continent's different countries. In this context, AI models must be designed to be both technically performant and transparent, adhering to the regulations of the financial sector.In a context where the regulation is stiffer, financial institutions are therefore better positioned to capture the operational benefits of AI systems that are both good at what they do and explainable to humans, while respecting regulatory requirements and maintaining financial stability. Theoretically, the results of this research support the Institutional Theory of organizational behavior, which states that organizational behavior is heavily influenced by regulatory pressures, institutional norms and governance expectations. They also support the Stakeholder Theory which shows that institutional accountability to customers, regulators, investors and society is strengthened through regulatory compliance.

5 Conclusion And Recommendations

5.1 Summary of Findings

The study examined the efficiency of Explainable Artificial Intelligence (XAI) for optimizing credit risk assessment in new financial markets of Africa. The results showed that transparency, interpretability, and accountability were significant and positive factors in Credit Risk Optimization (CRO), and that the regulatory environment boosted the effectiveness of explainable AI systems. The relationships between all the study variables were found to be quite positive by using correlation analysis, and regression analysis revealed that transparency was the most significant predictor of credit risk optimization. Moreover, the Structural Equation Modeling results validated the direct positive influence of explainable AI on credit risk outcomes, as well as the role of regulatory frameworks as a moderator for improving the effectiveness of AI in the financial institutions sector. The results are consistent with the previous findings that explainability is a determinant to enhance AI's adoption, trust, and performance in financial services (Arrieta et al., 2020; Wang et al., 2021).

5.2 Conclusions

The study findings led to a conclusion that XAI is a fundamental enabler of efficient credit risk management in emerging markets in Africa. In addition to boosting the predictive power of AI, explainable AI also creates greater transparency and accountability, helps with regulatory compliance and responsible financial inclusion. The results also highlight that the effectiveness of the credit assessment systems based on artificial intelligence relies on the quality of the technical model and the existence of governance frameworks that promote transparency, fairness, and trust. The findings align with the research by Arrieta et al. (2020), Jobin et al. (2019), and the European Commission (2019), which highlights the need for transparency, accountability, and human oversight in the development of trustworthy AI. The study, therefore, suggests that Explainable AI can offer a sustainable platform for innovativeness, regulatory compliance, and responsible lending in the modern financial landscape while contributing to the overall goal of financial inclusion in emerging economies (Fuster et al., 2022; Mhlanga, 2021).

5.3 Recommendations

To enhance the transparency, auditability and trustworthiness of financial institutions' credit risk assessment systems, they should be embedding explainability tools, including SHAP, LIME and Explainable Boosting Machines (Lundberg & Lee, 2017; Ribeiro et al., 2016). More comprehensive AI governance frameworks are needed to set standards for explainability, accountability, fairness and auditability of algorithmic lending decisions, in accordance with the principles of global AI governance (Jobin et al., 2019; European Commission, 2019). To implement responsible AI practices, fintech companies must integrate explainability at every stage of the model development process, striving to make sure that AI systems are fair and do not foster algorithmic bias (Verma & Rubin, 2018; Guidotti et al., 2018). Policymakers should also be encouraging the use of explainable AI solutions to help create a financial environment that allows people with little or no credit history to access financial services, ensuring that strong consumer protection and regulatory compliance requirements are maintained

(Mhlanga, 2021; Nambie et al., 2024). Financial institutions should also invest in staff training and AI governance capabilities to enhance the effective interpretation and monitoring of explainable AI systems.

References

1. Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). IEEE Access, 6, 52138–52160.
2. Ahmed, F., & Iqbal, A. (2025). The role of artificial intelligence in enhancing credit risk management: A systematic literature review of international banking systems. Pakistan Journal of Humanities and Social Sciences, 13(1), 478-492. https://journals.internationalrasd.org/index.php/pjhss
3. Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58, 82–115.
4. Barredo Arrieta, A., et al. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58, 82–115.
5. Bussmann, N., Giudici, P., Marinelli, D., & Papenbrock, J. (2021). Explainable AI in credit risk management. Computational Economics, 57(1), 203–216.
6. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 785–794.
7. Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv Preprint.
8. ESSIEN, I. A., NWOKOCHA, G. C., ERIGHA, E. D., OBUSE, E., & AKINDEMOWO, A. O. (2019). AI-Driven Credit Scoring Systems and Financial Inclusion in Emerging Markets. DOI 10.56472/25835238/IRJEMS-V3I9P117
9. European Commission. (2019). Ethics guidelines for trustworthy AI. AI HLEG.
10. Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., & Walther, A. (2022). Predictably unequal? The effects of machine learning on credit markets. The Journal of Finance, 77(1), 5–47.
11. Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. ACM Computing Surveys, 51(5), 1–42.
12. Hand, D. J., & Henley, W. E. (1997). Statistical classification methods in consumer credit scoring: A review. Journal of the Royal Statistical Society Series A, 160(3), 523–541. https://doi.org/10.1111/1467-985X.00078
13. Hodge, S. R. (2020). Quantitative research. In Routledge handbook of adapted physical education (pp. 147-162). Routledge.
14. Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389–399.
15. Lawton, C. (2019). Effective use of secondary quantitative data sources. In Handbook of research methods on the quality of working lives (pp. 177-193). Edward Elgar Publishing.
16. Lessmann, S., Baesens, B., Seow, H. V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring. European Journal of Operational Research, 247(1), 124–136.
17. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30.
18. MacInnes, J. (2020). Secondary analysis of quantitative data. SAGE Publications Ltd.
19. Mhlanga, D. (2021). Financial inclusion in emerging economies: The application of machine learning and artificial intelligence in credit risk assessment. International journal of financial studies, 9(3), 39. https://www.mdpi.com/2227-7072/9/3/39
20. Nambie, N. B., Ocansey, E. O., Dadzie, P., & Obobi, B. A. (2024). Credit risk assessment, regulatory compliance, and financial intermediation in Sub-Saharan Africa, the role of artificial intelligence. International Research Journal of Economics and Management Studies IRJEMS, 3(9).
21. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Why should I trust you? Explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1135–1144.
22. Rudin, C. (2019). Stop explaining black box machine learning models for high-stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1(5), 206–215.
23. Samek, W., Wiegand, T., & Müller, K. R. (2017). Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models. arXiv Preprint. https://doi.org/10.48550/arXiv.1708.08296
24. Thomas, L. C., Crook, J. N., & Edelman, D. B. (2017). Credit Scoring and Its Applications (2nd ed.). SIAM.
25. Varian, H. R. (2014). Big data: New tricks for econometrics. Journal of Economic Perspectives, 28(2), 3–28.
26. Verma, S., & Rubin, J. (2018). Fairness definitions explained. 2018 IEEE/ACM International Workshop on Software Fairness, 1–7.
27. Wang, Y., Wang, S., Lai, K. K., & Pei, J. (2021). Explainable machine learning for credit risk management. Decision Support Systems, 146, 113563. https://doi.org/10.1016/j.dss.2021.113563.