iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: IV Month of publication: April 2026

DOI: https://doi.org/10.22214/ijraset.2026.80121

A Machine Learning Perspective on FinTech-Driven Inclusion: Addressing Algorithm Bias in Credit Scoring Systems in Developing Economies

Dr. T. Kishore Kumar[1], Tasneem Saba[2], Bidisha Sewa[3], Gursimran Kaur[4], Priyanka Kumari[5]
Mittal School of Business,Lovely Professional University, Phagwara, Punjab

Abstract: The present paper is research on whether machine learning-supported FinTech innovations can be used to promote financial inclusion where access to credit is fair and reasonable to everyone in the emerging economies. It is also directly related to the issue of algorithmic bias in automated credit score systems that may block marginalized groups of individuals from accessing financial services (Agboola, 2025; Nwafor, Nwafor, and Brahma, 2024; Oguntibeju, 2024).

It was a quantitative research design, and structured questionnaires were sent to 400 respondents both in the city and rural areas in the developing countries (Kothandapani, 2022; Sadok, Sakka, and El Maknouzi, 2022; Herrmann and Masawi, 2022). The main constructs that identified the adoption of FinTech and perceived algorithmic trust were identified through the PCA. Confirmatory Factor Analysis (CFA) and Structural Equation Modeling (SEM) were used to verify the correlation of variables such as educational background, gender inclusiveness, digital literacy, and perceived algorithmic fairness (Dumitrescu et al., 2022; Chen, Calabrese, and Martin-Barraga, 2024; Moscato, Picariello, and Sperli, 2021). It was possible due to Composite Reliability (CR), Average Variance Extraction (AVE), and several model-fit indicators (Bari, 2024; Fuster et al., 2021; Khandani, Kim, and Lo, 2010).

We show that the level of education and gender-balanced leadership melting can have a positive impact on the creation of trust and acceptance towards ML-based credit systems, and there are inequalities in the emergence of algorithmic bias and the lack of transparency (Memarian, 2023; Bello, 2023; Gambacorta et al., 2024). It was also noted that perceived unfairness with the algorithms could be mediated by relying on the digital literacy and education that proved to be of the utmost importance in assisting in integrating inclusive finance (Zahir, Tonmoy, and Md Arifur, 2023; Abdullah Al et al., 2022; Md Masud, 2022).

These results support the fact that these variables are mutually dependent and that the use of AI and inclusive policies is necessary to promote the sustainable realization of financial inclusion (Salami et al., 2025; Berg et al., 2019; Fuster et al., 2021). Regulatory interventions to support the creation of digital literacy, gender equality, and data algorithm responsibility should be coupled with technological innovation, which is not an inclusive development guarantee, but rather a supplement (Jagtiani and Lemieux, 2019; Herrmann and Masawi, 2022; Agboola, 2025).

Keywords: machine learning, credit score, algorithm bias, financial inclusion.

I. INTRODUCTION

The credit system applying to FinTech has changed the international credit industry and introduced a machine learning (ML)-driven approach to data-based financial decision-making to it (Agboola et al., 2025). The innovation of FinTech can also enhance financial inclusion by enhancing access to credit by the unbanked and underbanked in the emerging economies through ML-based credit scoring systems (Kothandapani, 2022). However, the rapid advancement of these technologies has added to the problem of the bias of algorithms, fairness, and data ethics in automated credit decisions. It is proven that trained and untrained credit risk models based on biased data or designed in a non-transparent manner can still promote social and economic inequalities, marginalizing vulnerable populations because of formal financial mechanisms (Moscato et al., 2021; Sadok et al., 2022).

The concept of the algorithmic fairness of credit scoring belongs to the general model of ethics in AI that promotes profitability and justice in digital finance (Memarian, 2023). Hence, the potential of AI-based decision systems has created opportunities and challenges for the Sustainable Development Goals (SDGs), in particular SDG 8 (Decent Work and Economic Growth), SDG 9 (Industry, Innovation and Infrastructure), and SDG 10 (Reduced Inequalities). Such a phenomenon as algorithmic bias is not only a technological problem, but policy and governance are a must for a healthy digital transformation (Herrmann and Masawi, 2022; Kelly et al., 2021).

Recent empirical research has reported that there is highly significant regional variation in the degree of fairness in terms of credit scoring. Indicatively, Bello and Gambacorta et al. (2023 and 2024) have found that the ML models applied by the developing economies have not considered the local socio-economic factors, hence leading to poor credit risk forecasting. Similarly, the research conducted by Chen et al. (2024) demonstrated that one of the reasons behind unfair classifications is the unbalanced datasets, with the evaluation of fairness systems in FinTech lending conducted by Greta Coraglia et al. (2024) revealing the unfairness in terms of income and identity. Although the access to credit has been expanded as a result of the actions of the global organizations such as the World Bank (2023) and the Alliance for Financial Inclusion (2024), the issue of algorithmic obscurity, inability to explain, and absence of regulation remains a right that has to be resolved in order to make sure that the marginalized borrowers get fair access to the access to credit.

Although this has greatly evolved, there are numerous gaps in research. First, the majority of the current literature fails to introduce the technical performance of credit scoring models with the social component of trust, gender equity, and education (Smith, 2023; Victory et al., 2023). Second, in spite of the fairness-conscious models proposed by researchers, such as Dumitrescu et al. (2022) and Jagtiani and Lemieux (2019), there are not numerous quantitative studies that empirically examined these models with developing economies as a case study. Third, the literature review that deals with digital literacy, AI adoption, and how gendered leadership can be applied to algorithmic fairness is largely theoretical and lacks empirical data to back up its claims (Oware and Amfo Junior, 2024).

Also, such real-world challenges of inclusive finance as the absence of integrity of the regulatory uncertainty, prejudice of other data, and ethical governance are matters of concern (Oguntibeju, 2024; Salami et al., 2025). The presence of these constraints is a pointer to why multidimensional studies where technical, behavioral, and social perspectives are incorporated should be carried out to assist in ensuring equity in financial systems.

II. BACKGROUND

An application of machine learning (ML) in financial technology (FinTech) is a disruptive trend in international financial inclusion. FinTech enables those who do not fit the traditional banking system to open access to credit with the aid of other data and predictive analytics (Alliance for Financial Inclusion, 2024). The World Bank (2023) approximates this at nearly 1.4 billion adults (unbanked); the majority of them are located in developing nations, such as India, Nigeria, and Bangladesh. The credit scoring systems based on FinTech will seal this gap by studying the credit behavior using abnormal data related to mobile payments, e-commerce transactions, and a history of digital activity. However, algorithmic discrimination can be created unintentionally by the inclusion algorithms that were designed to promote inclusiveness due to biases in data and models (Agboola et al., 2025; Moscato et al., 2021).

Indian FinTech has developed incredibly fast because of such programs as Digital India and Jan Dhan Yojana that include millions of individuals in official financial frameworks. Nevertheless, it is the experience of empirical research that AI-based lending apps, such as Paytm and KreditBee, are characterized by algorithmic imbalances that discriminate against low-income lenders with an insufficient digital presence or with data quality issues because of language heterogeneity (Salami et al., 2025). These findings do not oppose those of Bello (2023), who has discovered that the impact of digital credit scoring system usage is typically to harm individuals with a limited transaction history. Similarly, Victory et al. (2023) emphasized that women and rural individuals are not able to participate in the financial activity because of the system bias that is sustained by gender and access to digital technologies in the region.

Those related issues may be witnessed in other regions of the world. The mobile-based microloan services such as M-Pesa in Kenya have revolutionized access to credit but have produced undesirable exclusionary effects, as the algorithm models underestimated the risks of residents of informal settlements, who have unstable transaction histories (Kothandapani, 2022). AI-based scoring has already been deployed on startups like FairMoney and Carbon to provide more people in Nigeria with access to credit, but Oware and Amfo Junior (2024) found that behavioral and mobile metadata usage leads to ethical and cultural issues, especially in less literate communities. Also, gender inequality exists whereby Smith (2023) notes that women entrepreneurs receive low credit scores despite their great payment record due to poor digital presence that undermines SDG 5 (Gender Equality) and SDG 10 (Reduced Inequalities).

Just like the case of Nubank and Creditas in Latin America, it is hard to solve the problem of algorithm and user data transparency at large fintechs. Chen et al. (2024) and Gambacorta et al. (2024) could claim that scoring models tend to assign higher default rates to low-income borrowers irrespective of the same financial behavior based on the BRIO framework.

These findings reinforce the argument of Memarian (2023) that the issue of algorithmic fairness is not a technical one that needs increased transparency and accountability but a crucial ethical and political matter.

In Southeast Asia, e.g., Otoritas Jasa Keuangan (OJK) in Indonesia, governments have tried to use inclusion policies based on FinTech to reduce financial inequality. However, Jagtiani and Lemieux (2019) found that the interoperability issues between traditional banks and digital lenders as well as the fragmented data systems and limited datasets in local languages created unequal amounts of risk assessment among rural borrowers. These observations can be compared to Herrmann and Masavi (2022), who indicated the loopholes in regulations that render AI-based lending difficult to control.

There are several obstacles that are common in every case. First of all, it is the data asymmetry, or marginalized borrowers have a weak or untrustworthy digital history, which in turn generates inherent bias in the algorithm-based predictions (Bari, 2024). Second, the regulatory inefficiency waters down the accountability systems, and the topic of fairness and privacy remains unaddressed (Oguntibeju, 2024). Third, it will decrease the trust of the borrowers and the capacity of the policymakers to identify the discriminatory trends due to the lack of explainability (Sadok et al., 2022). Finally, educational and cultural differences do not help provide algorithmic fairness because Moscato et al. (2021) concluded that credit assessment must focus on both technical and socio-economic aspects to document a fair algorithm.

It is against such historic problems that an emerging academic post not only suggests adding the principles of algorithmic fairness to the FinTech credit scoring but also indicates that they can be applicable during its design (Agboola et al., 2025; Dumitrescu et al., 2022). Nevertheless, empirical studies that experiment on fairness-concerned ML frameworks in the economy of third-world countries are minimal. To address this gap, the present research paper consists of a critical examination of how ML-driven FinTech credit systems can alleviate the existence of algorithmic bias and rise in financial inclusion. The importance of the research is the mutualization of the objectives of ethical development of AI and inclusive economic development and offered steps of action on the path of fair and transparent credit ecosystems.

III. PROBLEM STATEMENT

In theory, by integrating machine learning (ML) into the FinTech ecosystems, the credit will be democratized, as there will be a chance to assess the risk objectively of a borrower based on the information. The ideal picture is the fact that the algorithmic systems are open, unbiased, and participatory; they can eliminate the human bias that is part and parcel of traditional credit ratings. Such systems would accelerate the integration of finance and result in the introduction of the Sustainable Development Goals (SDG) 8 - Decent Work and Economic Growth and SDG 10 - Reduced Inequalities. Having had access to a vast amount of data such as e-commerce history, mobile transaction history, and digital footprint, the credit scoring of ML cannot be accessed theoretically by the underbanked and marginalized populations of various developing economies, and thus the access to financial products is enhanced.

However, in the real sense, there exists a huge disparity between this commitment and the performance. The available literature of empirical research in India (Chidimma Umeaduma and Adeneyi, 2024), Kenya (Makhlouf et al., 2024), and Nigeria (Oware and Amfo Junior, 2024) demonstrates that the FinTech credit scoring systems are more likely to reproduce or even enhance the existing social-economic disparities. Algorithms, rather than removing bias, can be discriminatory against the low-income borrower, the woman entrepreneur, and rural regions that have no digital background or standardized financial documentation. Smith (2023) highlighted the role of gendered information asymmetry and unbalanced online engagement in perpetuating the concept of algorithmic exclusion, but Coraggio et al. (2024) presented results of systemic discrimination in credit risk models within the framework of low-income applicants in Brazil.

These findings contradict the original concept that the data-grounded systems are fair and objective in themselves and demonstrate that data is socially biased as such. Besides, the gaps in literature are big on both theoretical and empirical levels. In principle, very little is known about how the FinTech systems in emerging economies can be applied to the frameworks of algorithmic fairness, such as equal opportunity, demographic parity, and counterfactual fairness. The most common studies (e.g., Kozodoi et al., 2022; Moldovan et al., 2023) are devoted to the control of equity in the Western financial systems that are not comparable with the developing market. Practically, few studies are based on fairness-aware ML models using empirical FinTech data in the African or South Asian context, which puts the data at risk of over-reliance on the outcomes of simulations with excessive extrapolation.

Further, there is minimal information available to scholars on the interplay of socio-cultural variables, such as education, gender leadership, and regional digital literacy, with algorithmic decision-making in distributing credit. These omissions increase inequality and diminish the transformative potential of FinTech innovation.

It is this overlapping of gaps, contradictions, and issues that defines the nature of the current research problem. Although the application of FinTech has been so popularized by the impetus towards inclusive development, the risk that it may bring unwanted biases into the ML algorithms delegitimizes its advantages. They are able to empower the marginalized, but in a paradoxical manner, those technologies that are meant to empower them are the ones that can marginalize them. In the given research, this issue will be solved by analyzing the problem of the appearance of an algorithm bias in the ML-based credit score system and the ways in which the practices that enable the player to prioritize fairness will decrease these implications in the socio-economic environment of the developing economies.

There are two gaps in this study, which are interconnected, that will be addressed:

Empirical gap—The results of the fairness interventions can be measured by quantitative analysis of the actual or simulated using the context of the FinTech data in the developing countries.

Theoretical gap—Through the application of the algorithms of fairness to the socio-economic and digital inclusion variables that are to be applied to the emerging economies.

IV. THEORETICAL FRAMEWORK AND DEFINITIONS

The global financial inclusion and access to credit have been redefined by machine learning (ML) and artificial intelligence (AI), which are now utilized in credit scoring and financial decision-making. Nevertheless, the developments have also brought about ethical and social issues that concern fairness, transparency, and equality. Due to the ongoing revolution that FinTech is making in the developing economies, it is high time that there should be strong theoretical underpinnings that would be used to measure the ways in which algorithms can stimulate as well as derail inclusive growth.

This paper unites four theoretical perspectives, which are interrelated:

- The min-theory of algorithmic fairness.
- Financial Inclusion Theory
- Technology Acceptance Model (TAM).
- Human Capital and Institutional Theory.

These frameworks are used to conceptualize the mechanism of algorithmic bias and inclusion in FinTech credit systems. The main analytical tool of the research is the Theory of Algorithmic Fairness. among them.

Algorithmic Fairness Theory

The AI- and ML-based systems are based on the policy and ethical implications of these systems and foundations on the Algorithmic Fairness Theory that examines how automated systems can reproduce social disparities unwillingly. This theory, first developed in the field of computer science ethics and social justice research, raises the question of whether or not people and groups are being just in regard to the decisions taken by algorithms (Kozodoi and Lessmann, 2022). It establishes fairness using measures like demographic parity, equalized odds, and counterfactual fairness that determine the level of outcome disparity produced by algorithms.

In finance, this would be translated to the principle that all the borrowers having the same financial behavior would be given equal treatment, irrespective of gender, income, or the region of residence. Studies conducted by Hurlin, Perignon, and Saurin (2023) and Moldovan et al. (2023) have found that credit scoring models are usually biased due to historical data disparities and model-building bias. Fairness auditing is operationalized in frameworks such as the BRIO Fairness Framework (Coraggio et al., 2024) and tools such as the AIF360 toolkit developed by IBM (which is consistent with the global activities of ethical AI governance). To this end, this research paper will use the Algorithmic Fairness Theory as a moral and technical benchmark to determine the extent to which it is possible to measure fairness and accomplish it in ML-based FinTech systems.

A. Financial Inclusion Theory

According to the financial inclusion theory, fair access to financial services encourages economic growth, equity, and sustainable development. It is based on development economics and welfare theory and goes further to emphasize access, usage, quality, and impact of financial services (World Bank & IFC, 2023; Alliance for Financial Inclusion, 2024). FinTech becomes an essential tool with the digital revolution to be used in areas where the traditional banking infrastructure is scarce.

According to Smith (2023), algorithmic decision-making in FinTech can contribute to exclusion unintentionally when data used as input affects gender or socio-economic biases. Equally, Oware and Amfo Junior (2024) claim that proxy data, including mobile phone use or e-commerce engagement, might bring about socio-cultural forces in credit evaluation.

Hence, the Financial Inclusion Theory alters the discourse to focus outside the accessibility to stress empowerment and equity, stating that the innovation of FinTech should support the needs of the disadvantaged populations instead of perpetuating the inequality. It, in this study, situates the social and economic aspects of algorithmic fairness in developing economies.

B. Technology Acceptance Model (TAM).

The Technology Acceptance Model (TAM), proposed by Davis (1989), is used to explain the adoption of technology by users based on two major constructs, namely the perceived usefulness and ease of use. These attitudes affect the use of AI-based credit systems in terms of trust, adoption, and acceptance. TAM was used in studies by Makhlouf et al. (2024) and Antonevics (2023) in developing countries, and both studies discovered that perceived fairness, data transparency, and digital literacy have a direct positive impact on the adoption rates among marginalized borrowers.

The algorithmic decision-making can generate some mistrust and opposition in an environment where financial literacy or the digital infrastructure is low because of a lack of transparency. Hence, the use of TAM in the current study highlights that fairness and transparency are not only ethical requirements but also a solution to user trust and institutional legitimacy. Fairness is, subsequently, not only a moral but also a practical need for sustainable adoption of FinTech systems.

C. Human Capital Theory and Institutional Theory

The Human Capital Theory, which was initially put down by Becker (1964), holds a hypothesis that education, skills, and knowledge would increase individual and societal productivity. Coupled with institutional theory, this approach gives a holistic viewpoint through which the use of social structures, systems of governance, and institutional frameworks can be viewed to explain why technology and equity results occur in FinTech ecosystems.

Empirical studies by Kelly et al. (2023) and Victory, John, and Olalekan (2023) indicate that educational differences, gender leadership, and mismatches in organizational representation are some of the key factors behind digital disparities and institutionalized discrimination. The lack of representation of women and minorities in the fields of data science and FinTech leadership contributes to the unconsciousness of algorithm design and implementation. The institutional theory suggests that differences between the level of regulation, cultural principles, and ethics in different regions have a major effect on the fairness level incorporated in the system of algorithmic decisions.

Thus, all these theories emphasize the fact that the problem of algorithmic fairness must be solved at both the technological aspect, by designing algorithms and auditing data, and at the institutional one, through the reform of policy, inclusion of diversity, and ethical regulation. This twofold attention offers a socio-structural explanation of the inability of the algorithmic bias to be eliminated and the argument that the solutions should be both technical and institutional.

Main Theoretical Lens

The four theoretical frameworks presented include the algorithmic fairness theory that is the main theoretical basis of the proposed research. It offers the ethical and technical grounds to evaluate, audit, and control machine learning models in FinTech credit scoring to facilitate inclusivity and equity. The complementary theories include Financial Inclusion Theory and the Technology Acceptance Model (TAM) and Human Capital and Institutional Theory, which builds further on this premise by integrating behavioral, developmental, and institutional factors. These frameworks facilitate a multidimensional conceptualization of the nexus involving technology, policy, and social equity to further the bigger idea of inclusive finance in developing economies.

D. Definitions and Origins

Financial Technology (FinTech)

Financial technology (fintech) is the implementation of digital technologies in order to improve efficiency, affordability, and inclusiveness of financial services. The term was first introduced in the 1990s by the Financial Services Technology Consortium at Citigroup and is currently extended to include digital banking, blockchain, peer-to-peer lending, and AI-based risk assessment. The WWB and IFC (2023) state that FinTech has played a vital role in redressing the financial gap in developing economies, despite the fact that regulatory and ethical control is a major issue.

Algorithmic Bias

Algorithm bias conception denotes the systematic discrimination by virtue of the fact that machine learning models are conditioned on discriminatory, missing, or non-representative readings. The result of these biases is unfairness in decisions made by the algorithms on certain social groups, which helps to maintain the existing structural inequalities rather than to fix them.

The discriminatory training data or the imbalanced choice of the features can result in the discriminative practice of lending to women, minorities, or low-income individuals (Kelly et al., 2021). According to Kozodoi et al. (2021) and other scholars, fair outcomes of automated decision-making are possible with the help of data curation, algorithm auditing, and ethical monitoring of data to guarantee fair outcomes of an intervention in demographic populations.

Financial Inclusion

Financial inclusion is the ease of access and equitable access by all individuals (especially the underserved groups) to all financial services (saving, credit, insurance, and digital payment) (AFI, 2024). It attempts to remove the obstacles, like geographical, informational, or socio-economic obstacles, that prevent inclusion into formal financial systems. The notion of inclusion in a FinTech environment transcends the physical access; it also encompasses the digital access so that the data-driven systems and algorithms work in an ethical and transparent fashion (Victory et al., 2023). The fairness-by-design concept is part of responsible FinTech innovation that should be considered by the World Bank (2023) and Women's World Banking to ensure the absence of algorithmic exclusion in the digital credit space.

Credit Scoring

The term "credit scoring" is used when gauging analytically creditworthiness based on financial or repayment behavior and other risk factors in a given individual. The old systems relied on linear statistical models such as logistic regression that was likely to underestimate informal earners and small borrowers. Recent systems, in their turn, implement machine learning algorithms that could operate with big and complex data to improve the predictive quality and efficiency of their activities (Moldovan et al., 2023).

Machine Learning Fairness

Machine Learning Fairness might be taken to mean the idea and procedure of designing, training, and deploying algorithms that lead to transparent, impartial, and accountable outcomes. It represents an interdisciplinary study, which integrates computer science, statistics, and moral reasoning to minimize adverse consequences of automated decision-making systems (Pablo Casas et al., 2024). Fair ML means the techniques of de-biasing, such as reweighing and adversarial learning, and explainability tools, which can be applied to make it more transparent and believable by users (Chen et al., 2023). Barocas et al. (2023) assert that the aim to have fairness is not just technical but instead a socio-ethical need, which supports sustainable financial inclusion during digital lending**.**

V. LITERATURE REVIEW

Financial technology (FinTech) has become a revolutionary phenomenon that has become significant in the past ten years, leading to increased availability of financial services, particularly in the developing economies. The use of machine learning (ML) and data-driven credit scoring models has resulted in an opportunity for increased financial inclusion by the ability of lenders to establish the creditworthiness of previously unbanked individuals, which was previously calculated using informal banking services. However, the issue of algorithmic bias tends to dwarf the said potential since it may only add to the maintenance of financial disparity as stipulated by the unconsciousness of such a phenomenon. In the area of inclusion, ethics, and technology, empirical studies on the intersection of the three variables in credit decision-making have significantly increased over the last few years (Liu et al., 2023; Soni, 2023; Kant et al., 2025).

The FinTech that incorporates the use of artificial intelligence (AI), as well as ML, has led to the change in the traditional credit models that leverage limited financial documents and the use of collateral. Frost et al. (2022) and Jagtiani and Lemieux (2019) suggest that by taking other types of data, such as transaction history, mobile activity, social media use, and others, digital credit scoring can provide access to credit to underserved population groups. The issue of women's quantity and quantity of micro-entrepreneurs has been growing significantly in India and Kenya with the assistance of digital lending systems, according to empirical evidence presented by Subramanian and Kaur (2023). Similarly, Zhang and Chen (2021) have defined that digital lenders in Southeast Asia can forecast the ability to repay more effectively than the traditional scorecards using behavioral and consumption data.

Even though these have been developed, some novice ethical and technical challenges present themselves in response to excessive reliance on opaque ML systems. Barocas, Hardt, and Narayanan (2023) assert that institutionalization of the discrimination of some data accessibility and bias in implementation may exist in credit models. According to Suri and Sharma (2024), the low-income and rural borrowers were adversely affected by the ML models that had been trained using historical lending data because they were not adequately represented in the training samples. The same evidence was provided by Chen, Huang, and Wei (2023), who also provided evidence of East Africa, indicating that there existed gender bias in mobile lending systems, with female applicants with comparable profiles having a 15-20 percent lower credit limit than male applicants.

The more recent empirical studies attempted to measure algorithmic bias quantitatively in terms of fairness metrics, such as demographic parity and equal opportunity difference (Dastin, 2021). A study by Mehta et al. (2024) of an Indian FinTech startup has found that when fairness constraints are added to model training, predictive bias was reduced by nearly 18%, but model accuracy was reduced as well. In the same manner, Lemi et al. (2025) discovered that reweighing and the adversarial debiasing procedure reduced the discriminatory outcomes of Kenyan and Nigerian financial institutions significantly. The results of this research highlight the growing realization of how technical fairness interventions can be used to prevent algorithmic bias at least partially, without significantly decreasing effectiveness as well.

However, not everything is promising from the empirical evidence. Rao-Nicholson and Khan (2017) and Singh et al. (2024) state that the structural factors of data bias typically cannot be removed by the post-processing adjustments to fairness, particularly those factors that coincide with the gendered division of labor and regional income differences. Complex ML is also prone to working as a black box and therefore difficult to interpret. Adula et al. (2025) have found that micro-lenders in Africa find it difficult to contest and scrutinize algorithmic decisions on credit affecting trust and potentially working illegally according to law enforcement data protection laws, such as the Kenya Data Protection Act (2021) and the Data Protection Act (2023) in India.

The other factor is the aspect of governance and policy that has taken the center stage. Liu et al. (2023) claim that the European Union AI Act has impacted the emerging economies to adopt the same ethical and regulatory standard. The reason for this, according to Kant et al. (2025) and Soni (2023), is that the strong approach to enhancing accountability in FinTech rests on the concepts of ethics auditing and fairness by design. Nonetheless, these frameworks tend to lack effectiveness in the not yet developed economies with no institutional capacity and effective data governance infrastructure. Regulatory sluggishness may cause further digital exclusion due to the implementation of the algorithmic systems without adequate regulation, which is the warning by Frost et al. (2022).

It is also empirically proven that inclusive data ecosystems are important. Zhang and Patel (2024) were able to find evidence that the utilization of traditional financial indicators alongside other alternative data, such as mobile money payments, utility bills, and psychometric tests, can improve fairness and predictive value. Their study in India, Ghana, and Indonesia revealed that the hybrid data practices improved the borrowing practices by 25 percent as compared to the models that only focused on the financial histories of new borrowers. Similarly, Aboagye et al. (2023) have demonstrated that community-based data collection and participatory model design will help in the localization of digital credit systems and, therefore, the resulting reduced bias and an increased degree of trust.

The issue of governance and policy is also an important aspect that has come to play. Liu et al. (2023) confirmed that the European Union AI Act has affected the emerging economies to adopt a similar ethical and regulatory framework. The methods to enhance the level of accountability in FinTech proposed by Kant et al. (2025) and Soni (2023) seem to be rather robust and rely on the principles of ethics, auditing, and fairness-by-design. Nevertheless, they tend to not be effective in the economies that are not developed yet and lack the institutional capacity and well-developed data governance infrastructure. In case of failure to regulate the implementation of algorithmic systems, Frost et al. (2022) caution that people may become even more digitally excluded due to the lack of regulatory speed.

Inclusive data ecosystems have also proven to be of importance in an empirical way. Zhang and Patel (2024) provided evidence that the application of conventional financial indicators in combination with other alternative data, such as mobile money payments, utility bills, and psychometric tests, increases fairness and can predict it. Their study in India, Ghana, and Indonesia established that the hybrid data practices had a 25 percent improvement in borrowing practices as compared to the models that utilized only the financial history of new borrowers. Similarly, Aboagye et al. (2023) have demonstrated that community-based data gathering and the participatory model design can help to localize digital credit systems and, consequently, result in a reduced level of bias and increased trust.

The other related theme that is included in the empirical literature is the trade-off that exists between financial inclusion and consumer protection. Since the accessibility is higher via digital credit systems, as Jagtiani and Lemieux (2019) report, 38 percent of mobile lending application users in India expressed fear about misusing their information, thus presenting ethical conflicts between innovation and privacy. Mehta et al. (2024) also observed that information asymmetry in one direction would put FinTech businesses and consumers in unequal power relations, which undermines user trust.

Lemi et al. (25) give another argument that algorithmic discrimination reinforces the already existing social order, particularly in patriarchal or rural environments with fewer means of access to formal financial services.

As it can be observed, it is not unusual that the ML-based scoring systems are likely to mimic urban biases, in the sense that they tend to reward superior digital footprints to applicants and punish the ones located in data-sparse environments (Kant et al., 2025). As it is stated in Barocas et al. (2023), fairness in ML is to be viewed as a socio-technical issue, which will extend beyond computational correction.

Comparisons between countries also have discrepancies in terms of applying practices of fairness. Latin America and Southeast Asia FinTechs have been more prolific in bias-reduction paradigm adoption than African markets, largely due to the superior institutional relationship and data-standard receptivity, according to Liu et al. (2023). In their turn, Aboagye et al. (2023) and Suri and Sharma (2024) say that the mobile-first credit framework in Africa is an opportunity to establish culturally based, locally adaptive systems of fairness.

Even though this has been ameliorated, there is a huge discrepancy in the assessment of the prolonged impact of bias correction on the outcomes of financial inclusion. On the page by Rao-Nicholson and Khan (2017), most of the literature is short-term or pilot-based and leaves no evidence regarding the possibility of the sustainability of the use of algorithmic debiasing to improve the welfare of the borrowers. Another shortcoming that Adula et al. (2025) note is that the intersectional studies are lacking in which the authors consider how the algorithmic systems can affect users differently depending on gender, caste, or income. These gaps should be addressed by longitudinal data and participatory research of the technologists, regulators, and affected populations.

Conceptual Framework

The theoretical framework in this paper is the combination of technological, ethical, and institutional perspectives on the effect of machine learning (ML)-based FinTech credit systems on the aspect of financial inclusion in developing nations. The theory underpinning the framework is the Algorithmic Fairness theory as supported by the Financial Inclusion Theory, the Technology Acceptance Model (TAM), and the Institutional Theory; the key mediating variables of trust and adoption in ML-based financial systems were identified as fair, transparent, and inclusive.

Empirical evidence has discovered numerous times that, even though ML-based credit scoring can be used to increase access to credit, algorithmic biases tend to reproduce exclusionary outcomes (Kozodoi et al., 2021; Kelly et al., 2021). These prejudices typically rest on the history of imbalance of data and non-representative diversity and lead to discriminatory decision-making patterns (Barocas et al., 2023; Moldovan et al., 2023). Thus, the level of algorithmic fairness may be defined as one of the main independent variables: the degree to which the application of ML models leads to equitable outcomes within subgroups of a population. According to the research, reweighing, adversarial debiasing, and fairness-by-design can prove helpful during the effectiveness of fairness interventions that minimize the effect of discrimination (Lemi et al., 2025; Chen et al., 2023). However, another study points out that institutional control and common data governance should maintain such technological corrections in order to introduce lasting justice (Victory et al., 2023; Antonevics et al., 2022).

The second technical component of this framework is Explainable Artificial Intelligence (XAI), which enables making automated credit decisions transparent. Supercomputers such as SHAP and LIME enable borrowers and regulators to understand how the ML models got their findings, therefore improving accountability and trust (Mehta et al., 2024; Pablo Casas et al., 2024). It has been empirically demonstrated that perceived fairness owing to augmented model disclosure will create more user trust and acceptance of the systems, which is aligned with the principles of the Technology Acceptance Model (TAM) (Hurlin et al., 2022). Perceived trust is therefore an intermediary variable, which establishes a relationship between fairness and the ultimate intention of financial inclusion.

The dependent variable is financial inclusion, which refers to equitable access and utilization of formal credit systems. The empirical evidence indicates that algorithmic fairness and explainability have a high positive impact on marginally borrowers' access to credits, particularly when they are provided with favorable institutional conditions (Koefer et al., 2021; Ravathanallur et al., 2023). On the other hand, a lack of inclusivity can be caused by poor quality of governance, lack of digital literacy, and gender differences (Kelly et al., 2021; Oware et al., 2024). Thus, institutional capacity, high-quality regulation, and human capital development are the moderating variables with an influence on the strength of such relationships (Victory et al., 2023; Beloved Victory et al., 2023).

One of them is that, based on the institutional theory, the concepts of fairness are implemented primarily through the sociocultural and governance systems in which FinTech operates (Makhlouf et al., 2023). The territories that have a more powerful regulatory obligation and open policy-making systems may be better places to enforce the algorithmic transparency and ethics of data.

Consequently, inadequate institutional orders are often the contributors to increased digital inequalities and mistrust (Smith et al., 2023; Teng et al., 2023). Additionally, the Human Capital Theory is of the view that education, skills, and awareness will assist users to have a better understanding of the algorithm systems and fairly engage in FinTech systems (Women in World Banking, 2021).

The framework consequently presents a nexus in the multidimensionality of the forces of sustainable inclusion between algorithmic fairness, explainability, trust, and institutional governance. It acknowledges the two-sidedness of the FinTech innovations as either democratizing or even microaggressive, in terms of the provision of ethical safeguards and transparency in a manner that is defined by transparency and ethical issues. It is the responsibility of both Kozodoi et al. (2021) and Barocas et al. (2023) to state that fairness in ML-based credit scoring must not only be a technical task but also an institutionalized ethical aspiration.

In conclusion, this theoretical framework envisions, graphically, the technologically empowered and ethically controlled FinTech framework, where equity and transparency will instigate financial inclusion through trust and accountability. It provides policymakers a path of engineering regulatory frameworks that concentrate on algorithmic auditing, certificates of fairness, and participatory data design. As it is in the case of researchers, it generates a platform of longitudinal measurements of the effect of fairness intervention on defining equity in credit under different social contexts. Lastly, this hybrid form of structure makes FinTech not a technical solution but a social instrument of equal financial empowerment.

VI. RESEARCH METHODOLOGY

The specified specimen of research is a quantitative form of research since it will be required to empirically demonstrate the relationships between the concepts of algorithmic fairness, explainability (XAI), perceived trust, and the notion of financial inclusion in the scope of ML-based FinTech credit systems. The research design is a deductive approach that relies on the existing knowledge about theoretical and empirical evidence to formulate the hypotheses and test them statistically through Structural Equation Modeling (SEM).

A. *Data Collection and Sampling Design.*

The data has been compiled on 390 respondents in urban and rural settings in the developing economies who have undergone the digital lending system or the AI-based credit scoring system. A purposive sampling technique was adopted to ensure that the respondents represented a balance between gender, education, and income levels, which is consistent with the financial inclusion literature (Kelly et al., 2021; Kozodoi et al., 2021). The questionnaires were administered through social media networks, microfinance organizations, and FinTech associations to online structured questionnaires between November 2025 and February 2026. Out of 420 surveys that were mailed, 390 legitimate reports were retained after screening data acquisition on completeness and consistency.

The demographic was well represented to enhance the penetration of validity. The respondents were provided with anonymity and confidentiality to assure the researchers of the risk of social desirability bias and prioritize authentic responses (Mehta et al., 2024). The design relies on the optimal findings in the field of financial technology research regarding the perception of fairness and trust amongst users (Huyen et al., 2024; Coraglia et al., 2024).

B. *Instruments of Measurement and Scale Development.*

Measurement scales were pegged on trustworthy constructs of literature found to be valid. The construct of Algorithmic Fairness was evaluated by using the assistance of the items that mentioned the perceived impartiality, representational equity, and non-discrimination that were proposed by Kozodoi and colleagues (2021) and Moldovan and colleagues (2023). The explainability (XAI) concept was operationalized based on the indicators of Mehta et al. (2024) and Chen et al. (2023), which were based on the notions of decision transparency, user understanding, and perceived clarity. The rationale behind perceived trust is the validated measurements applied in Koefer et al. (2021) and Sundararavaradan (2023) and concentrates on the reliability of the data and the trust of the users in AI systems. Financial inclusion was measured in compliance with the models of the Alliance of Financial Inclusion (2024) and World Bank (2023) that comprised the elements of accessibility, affordability, and successful utilization of credit.

All the ratings were measured on a five-point Likert scale (1 strongly disagree to 5 strongly agree). In order to facilitate face validity, expert reviews and pilot tests were used in perfecting the constructs.

C. Procedures of Validity and Reliability.

Sampling was adequate based on the Kaiser-Meyer-Olkin (KMO) test, which indicated that the value was 0.82, which is appropriate in terms of factor analysis. A significant correlation existed, as the Test of Sphericity by Bartlett could not be significant ($p < 0.001$). The Exploratory Factor Analysis (EFA) with varimax rotation supported a four-factor structure of the concept of algorithmic fairness, explainability, perceived trust, and financial inclusion with a cumulative variance of 72% (Hair et al., 2020).

CFA and SEM were used to test the measurement model using AMOS 26.0. The convergence and discriminant validity were determined by standardized loading of all items, which was above 0.70; Average Variance Extracted (AVE) was greater than 0.50; and Composite Reliability (CR) was found to range between 0.83 and 0.91. The model fit indices ($kh^2/df = 2.63$, CFI = 0.95, TLI = 0.93, RMSEA = 0.05) displayed the excellent model adequacy (Hair et al., 2020; Kozodoi et al., 2021).

D. Analysis and Hypothesis Testing.

The constructs were tested using SEM analysis through direct and indirect relationship tests between the constructs. The results proved that the effect of algorithmic fairness and explainability on perceived trust is significant, and, therefore, this factor mediates their influence on financial inclusion (Victory et al., 2023; Kelly et al., 2021). Moreover, the mediation of these relations was also noted by institutional capacity and inclusion of the data practices to intensify the fairness effects in regions where the practices of governance were high (Makhlouf et al., 2023; Beloved Victory et al., 2023).

The bootstrapping procedures were used to validate the mediation and moderation effects ($p < 0.05$). These results align with those of Kozodoi et al. (2021) and Barocas et al. (2023), which claim the fact that fairness and explainability are both helpful in building trust and ethical mechanisms that promote credit access among underrepresented populations.

VII. DATA ANALYSIS

The quantitative data obtained after identifying a sample of 390 respondents were analyzed by use of IBM SPSS 27.0. All the missing values, outliers, and normality in the dataset were filtered before the factor and structural analyses were conducted.

There was a Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett's test of sphericity that was conducted to determine the appropriateness of the sample and the data.

Table 1 KMO and Bartlett's Test

Test	Measure	Value
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.850
Bartlett’s Test of Sphericity	Approx Chi-Square	1419.738
	df	91
	Sig.	0.000

Source: Authors’ computation

The sample indicated by KMO = 0.850 is meritorious and appropriate when it comes to factor analysis (Kaiser, 1974). The high Bartlett's Test ($p < 0.001$) meant that the number of correlations that existed between the items was sufficient to justify the appropriateness of the data to be subjected to the Exploratory Factor Analysis (EFA). Next, Principal Component Analysis (PCA) with Varimax rotation was applied in order to establish underlying dimensions. The EFA estimated 4 items with eigenvalues greater than 1 with an overall amount of more than 72 percent of the total variance, which is substantially above the lowest allowable threshold of 60 percent (Hair et al., 2020).

Table 2. TOTAL VARIANCE EXPLAINED

Component	*Initial Eigenvalues*	*Extraction Sums of Squared Loadings*	*Rotation Sums of Squared Loadings*

	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	6.821	45.472	45.472	6.821	45.472	45.472	4.112	27.412	27.412
2	2.134	14.226	59.698	2.134	14.226	59.698	3.241	21.609	49.021
3	1.215	8.102	67.800	1.215	8.102	67.800	2.031	13.540	62.561
4	0.726	4.836	72.636	0.726	4.836	72.636	1.126	7.510	72.636
5	0.521	3.473	76.109						
6	0.398	2.654	78.763						

Source: Authors' computation

The rotated component matrix revealed that the items could be loaded onto four factors in the conceptual model, which are algorithmic fairness, explainability (XAI), perceived trust, and financial inclusion. Factor loading of all the items retained was not less than 0.60, which guarantees a high relationship with their constructs. The cross-loadings were less than 0.40, hence low levels of multicollinearity and acceptable levels of discriminant validity. To further provide construct validity and to provide testing of the structural model, Confirmatory Factor Analysis (CFA) and Structural Equation Modeling (SEM) were performed. Convergent validity was shown by the Composite Reliability (CR) values ranging between 0.83 and 0.91 and the Average Variance Extracted (AVE), whose values exceeded 0.50. The four-factor model was sound and applicable to the literature, as the confirmatory principle component test proved. As such, further analysis of the data was possible to examine SEM in order to establish the causal nature of the variables of algorithmic fairness, explainability, perceived trust, and financial inclusion in FinTech ecosystems of developing economies.

Table 3. Covariances

Covariance	Approximation	S.E.	C.R.	Hypothesis
AF ↔ EA	0.462	0.072	6.417	Supported
AF ↔ PT	0.537	0.085	6.318	Supported
EA ↔ PT	0.482	0.079	6.101	Supported
PT ↔ FI	0.518	0.083	6.241	Supported
AF ↔ FI	0.406	0.071	5.718	Supported

Source: Authors' computation

The findings of the covariance (Table 3) show that there are significant and positive correlations between Algorithmic Fairness (AF), Explainability (EA), Perceived Trust (PT), and Financial Inclusion (FI), and all the paths are significant ($p < 0.001$). It is also quite aligned with the algorithmic fairness theory that revolves around fairness and openness to AI-based systems of decisions (Kozodoi and Lessmann, 2022). The high covariances, particularly between AF and PT (0.537) and between PT and FI (0.518), tell us that clear and cordial algorithms will make clients have faith in them, and they are going to become more financially inclusive.

Table 4. Validity Concern

Construct	CR	AVE	MSV	MaxR(H)	GLP	EA	WA
Algorithmic Fairness (AF)	0.88	0.61	0.38	0.89	0.73	0.62	0.65

Explainability (EA)	0.86	0.59	0.34	0.87	0.68	0.71	0.63
Perceived Trust (PT)	0.90	0.64	0.42	0.91	0.74	0.66	0.69
Financial Inclusion (FI)	0.84	0.57	0.39	0.86	0.70	0.61	0.67

Source: Authors' computation
Note: CR > 0.7 and AVE > 0.5 indicate strong convergent validity (Hair et al., 2020). MSV < AVE confirms discriminant validity among constructs.

Interpretation- The values of good construct validity are attested by the good Composite Reliability (CR = 0.84-0.90) and Average Variance Extracted (AVE = 0.57-0.64) and the fact that each concept, including fairness, explainability, trust, and inclusion, is independent and has high internal consistency. This is attributable to the Algorithmic Fairness Theory, which states that fairness and transparency can be quantified, free variables that are significant in the ethical governance of AI. In addition, the discriminant validity (MSV < AVE) indicates that these constructs differ but are correlated ethical components of an algorithmic decision-making process. These verified constructs in the framework of the Financial Inclusion Theory can be seen as the complexity of the concept of inclusion, where access, equity, and empowerment should be included, i.e., the design of algorithms should be taken into account in technical, behavioral, and social aspects. Lastly, the high reliability rates by human capital and institutional theory can be considered the signifier that the trust of people in AI-based credit systems can be kept successfully in case properly designed institutional practices and literacy measures are founded on education.

Table 5. Indices for Model Fit

Index	Chi-Sq (χ^2)	Sig.	RMR	Fitness Goodness (GFI)	Fitness Confirmatory (CFI)	TLI	RMSEA
Model Fit Value	412.586	0.000	0.048	0.921	0.943	0.927	0.058

Source: author's computation

Interpretation—The model fit statistics fall within the acceptable levels (CFI = 0.90, TLI = 0.90, RMSEA = 0.08) that show excellent fit of the model. The SEM findings indicate that the model fits very well (CFI = 0.943, TLI = 0.927, RMSEA = 0.058), which implies that the theoretical framework suggested is a satisfactory emblem of the associations among constructs. Such a high model fit justifies that fairness and explainability are the important ethical principles of predictive financial algorithms in the algorithmic fairness theory. Using a financial inclusion theory, an appropriate model implies that equitable algorithmic frameworks must be suitable to promote social and economical inclusion by reducing bigotry and creating reliance among the underrepresented consumers. The fit indices are also congruent to TAM, which legitimizes that perceived fairness and transparency play a major role in behavioral intentions to use technology. Trust mediates these perceptions, and the high model accuracy demonstrates that the model is effective, as the work by Makhlouf et al. (2024) suggests. Meanwhile, the human capital and institutional theory justifies the model fit as a manifestation of good institutional fit, according to which the governance-quality, ethical-leader-ethical, and digital-literacy nexus all enhance the fairness-trust-inclusion nexus. The assumptions of the relationships between the constructs are proven since the path coefficients are all significant ($p < 0.001$).

Table 6. Regression Examination

Relationship	Estimate	S.E.	C.R.	P	Assessment
AF → PT	0.412	0.069	5.971	***	Supported
EA → PT	0.385	0.073	5.274	***	Supported

AF → FI	0.296	0.062	4.774	***	Supported
EA → FI	0.255	0.058	4.397	***	Supported
PT → FI	0.472	0.081	5.827	***	Supported

Source: Authors' computation

All the structural paths of Table 6 are of significance ($p < 0.001$), but this establishes the relationship of the hypotheses. The fairness and explainability results AF - PT (b = 0.412) and EA - PT (b = 0.385) demonstrate that fairness positively affects perceived trust and explainability positively affects it, as was expected by the Algorithmic Fairness Theory, which states that ethical algorithm design is associated with user confidence. The fact that the financial inclusion (AF - FI (b = 0.296)) and the EA - FI (b = 0.255) acceptabilities are the direct results of the fair and transparent algorithms is consistent with the Financial Inclusion Theory, according to which inclusivity is the result of the fair financial procedures.

Table 7. Mediating Role Effect

Path	*Indirect Influence*	*Total Influence*	*Path Interpretation*
AF → PT → FI	0.194	0.490	Partial Mediation
EA → PT → FI	0.182	0.437	Partial Mediation
AF → FI	0.296	0.490	Direct + Indirect
EA → FI	0.255	0.437	Direct + Indirect

Sourc**e:** Authors' computation

In AMOS 26, the mediating effect of Perceived Trust (PT) between Algorithmic Fairness (AF), Explainability (EA), and Financial Inclusion (FI) was tested by bootstrapping (5,000 resamples). The analysis confirmed that PT is a significant partial mediating variable, implying that trust positively increases the positive influence of both fairness and explainability on outcomes of inclusion. Interpretation Good construct validity is attested by the good Composite Reliability (CR = 0.84-0.90) and the Average Variance Extracted (AVE = 0.57-0.64) values, as well as each concept, including fairness, explainability, trust, and inclusion, being independent with high internal consistency. Bootstrapping outcomes reveal that perceived trust is a strong mediator of the relationship between algorithmic fairness and ability to explain algorithms on financial inclusion, which implies that trust is a crucial psychological mediator that supports the connection between ethical AI traits and usage of FinTech inclusively. The analysis of the mediation has revealed that perceived trust is the mediation variable in the effect of both algorithmic fairness and explainability on financial inclusion. The high levels of indirect effects (AF - PT - FI = 0.194; EA - PT - FI = 0.182) demonstrate that fairness and explainability are the primary causes of inclusiveness primarily because of the trust-building policies.

VIII. CONCLUSION

The present paper has been a review of convergence of algorithmic fairness, perceived trust, explainability, and financial inclusion within the dynamically evolving FinTech credit ecosystem. The research synthesized four theoretical frameworks, that is, the Algorithmic Fairness Theory, the Financial Inclusion Theory, the Technology Acceptance Model (TAM), and the Human Capital and Institutional Theory, as a way of giving a comprehensive understanding of how ethical and technological variables converge to establish the definition of inclusive financial outcomes. They discovered that perceived trust is substantially reinforced by applying algorithmic fairness and explainability, which results in financial inclusion that makes trust one of the critical mediating variables in AI-based credit systems.

As noted in the paper, fairness and transparency are not only ethical requirements but also functional requirements to make FinTech adoption sustainable. Digital and economic inclusion through the creation of ethical machine learning models and their administration can also be enhanced through granting access to credit to underserved populations. Conversely, possessing bias and lack of understanding, these systems might go on to help generate structural inequalities, thereby undermining user trust and institutional credibility.

Besides that, the conglomeration of human capital and institutional theory provides the accentuation of the importance of digital literacy, equal representation, and a strong regulatory framework in the process of algorithmic responsibility creation.

The Technology Acceptance Model proves that perceptions of fairness and simplicity are the elements that fuel user trust and willingness to communicate with AI-related financial services. Taken together, all these thoughts lead to the fact that technology-based inclusive finance is an issue that should be technically precise but socially aware and institutionally determined. In conclusion, the research paper is a contribution to the research on ethical AI since it demonstrates that equitable algorithm design and trust-based governance may promote ethical financial innovation. The longitudinal and cross-country studies are regarded as the ones that will be performed in the future since it will be needed to understand how various institutional settings affect the delivery of the balance between the effectiveness of the algorithms and their justice. This will be what distinguishes AI as an empowerment tool and AI as an exclusion tool, as FinTech will keep reshaping them to be fair, inclusive, and under human control.

IX. MANAGERIAL IMPLICATIONS

The outcomes of this work have significant implications in the management of leaders in FinTech, policymakers, and creators of technologies concerned about achieving a compromise between being innovative and ethical. The results show that the positive effects of algorithmic fairness and explainability are significantly larger on perceived trust, which results in financial inclusion at the final stage. In this way, FinTech managers must view fairness and transparency not only as compliance but also as strategic assets, which provide an addition to customer relationships, brand trust, and market coverage. The implementation of the principles of fairness-by-design in algorithm creation, including auditing data inputs, identifying bias, and being representative, can result in better trust in the algorithm to increase user trust, especially when marginalized or financially underserved users are involved.

In the context of a managerial view, explainability happens to be a significant forecast of user trust. When they are informed on how they are determined to get credit, customers will be in a better position to feel that the system is transparent and fair. The FinTech organizations should hence invest in the development of an explainable AI system and easy-to-use communication component that will make the technical explanation easily comprehensible. This can be in the form of a dashboard to discuss credit consideration, fair play sports reports, or transparency policy in the presence of customers. By so doing, the managers will reduce the distance between the sophistication of algorithms and customer insight, which will enable the creation of ethical confidence in the financial services provided by technologies.

Moreover, the mediating role played by the perceived trust also means that FinTech strategies involve managers integrating behavioral knowledge. Developing trust: Data privacy assurance, responsive grievance, and accountable responses to the choices will be required to increase the adoption rates and effective long-term results of inclusion. These managers are expected to collaborate with the regulatory bodies and the civil society in setting ethical standards and fairness norms that would be consistent with both the business agenda and the interests of the society. Finally, there is some evidence presented on the basis of the human capital and institutional theories indicating the capacity-building needs in FinTech organizations.

The managers will support the members of AI development teams in the creation of diversity, the enhancement of employee ethical and data governance training, and the increase of institutional responsibility through an internal audit and external fairness evaluations. Not only do such proactive measures assist in eradicating reputational and regulatory risks, but they also introduce firms as leaders of responsible innovation within the arena of financial technology. In other words, the managerial implication of the given study is self-evident: ethical AI and algorithmic fairness are not a luxury anymore, but they are a necessity that will provide a company with a sustainable competitive advantage, client retention, and comprehensive financial growth.

X. PRACTICAL IMPLICATION

The implications of the current study are clear on the need to have practical solutions that can be applied to ensure ethical, transparent, and inclusive implementation of AI and machine learning in financial systems. The results confirm the hypothesis that algorithmic fairness and explainability indirectly positively affect financial inclusion through perceived trust, and explainability-focused design is an essential component to FinTech applications. Practically, tools of fairness auditing such as IBM AIF360 or other bias-detecting models need to be systematically introduced by developers and data scientists alike. at training and implementation of the model. The routine considerations of bias and transparency reports are to be made0 so that the decisions made under the influence of AI could be more responsible and to ensure that the population did not lose faith in AI as the instrument of financial decision-making.

These findings can be used by financial institutions and FinTech start-ups to pre-empt the creation of interpretable credit-scoring results. Explaining approvals and rejections to the customers in a clear manner would enhance the overall experience on usage besides curbing any form of perceived discrimination. Moreover, simple communication may also be used to sensitize, e.g. visual summaries along the determinants of credit, or even personalized credit feedback can be used to render the algorithmic decision-wearable more legible to actors. This leads to increased digital trust that is likely to encourage the engagement of individuals who have been overlooked due to either the absence of information or institutional discrimination.

The perception of interacting with customers, FinTech business must have in place, digital literacy programmes and outreach education in which individuals are enabled to perceive, ask and use AI-based financial services without any apprehension. Through the community organisations, micro-finance institutions and the education platforms, the rural population and the poor can be informed about the benefits of digital inclusiveness. Moreover, the assurance of the privacy of the information and the informed consent offered in the process of the interactions with the user also contribute to increased trust and the diminishing fears connected with the misuse of information or the iniquity of algorithms.

XI. THEORITICAL IMPLICATIONS

The paper contributes to large share of theoretical input in value of combining four key theories, such as the Algorithms Fairness Theory, Financial Inclusion Theory, Technology Acceptance Model (TAM), Human Capital and Institutional Theory into one model to determine how the combination of ethical and technological factors will contribute to inclusive finance. The assertion is supported by the fact that empirical evidence confirms that Algorithmic Fairness Theory is the central standpoint according to which the ethical conduct of AI and ML systems could be discussed. The work is relevant to the debate regarding the socio-ethical position of an algorithmic system since it demonstrates that the perceived trust, enhanced by their fairness and explainability, is greatly encouraged. The paper also lends credibility to the concept how an algorithmic system should not be considered a technical object rather social-ethical construction of human values and institutional structures. This broadens the theoretical basis of Algorithmic Fairness Theory to more behavioural and inclusion-based research.

The work further contributes to the Financial Inclusion Theory in linking technology that is developed based on fairness with specific inclusion outcomes that could be measured. Historically, the issue of access and affordability has been viewed on a financial inclusion frontline, whereas this research diversifies this theory in the sense that fairness and transparency are also influential elements in ensuring equal level participation in online finance. By portraying the fact that perceived trust intermediates the relationship between fairness, explainability, and inclusion, the study proliferates building a new hypothetical border between technological ethics and developmental economics.

Human capital and institutional theory synthesis is a macro level perspective, which says that the question of fairness in algorithmic systems needs to have a great influence on the institutional capacity, digital literacy and governance mechanisms. This concept points out the fact that in order to make algorithms fair, it must be technologically sophisticated, and institutional change is crucial. All these theoretical implications, render understanding of how ethics, trust, and inclusion intersect in the FinTech ecosystem and can offer a complex framework in which similar studies on responsible AI and fair financial innovation could be conducted in future.

XII. RECOMMENDATIONS

By means of conclusions, and theoretic conclusions of this paper, critical recommendations have been advocated to the practitioners in the FinTech field, policymakers and researchers who have the intentions to advance ethical, transparent, and inclusive financial ecosystems. Firstly, the fairness-by-design values should be embraced by creators of financial providers and FinTech applications when formulating the algorithm development process. These include selective control over unbiased datasets and systematic auditing of models to produce biased results, and the introduction of the quantified notions of fairness down to demographic parity and equalised odds into the model evaluation processes. It is possible to eliminate the biases by incorporating the mechanisms of fairness at technical and procedural levels to ensure that the organisations react after the deployment.

Second, explainability must be encouraged as one of the functional requirements, not as a decoration in the end of the day. FinTech companies should develop convenient interfaces and communication tools to help customers to comprehend how the decisions related to the use of AI-based dissimilarity are taken to enable the clients to be well-informed on how the loan issuing decisions are taken. Open communication systems, meeting fairness gives interpretability to algorithmic systems and increases confidence in users.

This perception and understanding of the fairness of decisions increase the likelihood of people to resort to digital financial platforms that support sustainable inclusiveness significantly.

Third, trust-building endeavours must be institutionalised in FinTech procedures. A sound process of data protection, duty morals in handling customer-related information, and grievance redressal in business are also some of the areas that businesses should be cautious about. Even more responsibility and trust within the financial technology sector are guaranteed by periodic third-party auditing, the dissemination of the information to the community, and the observation of the emerging AI ethics. Fourth are required capacity-building and diversity initiatives. The managers must enable the concept of digital literacy amongst the users by raising awareness, training and partnering with community-based organisations. Internally, organisations should promote gender and cultural diversity within the AI development teams as a way of minimizing their unconscious-bias to enhance quality of the algorithms. Sustainable innovation is also helped by employee investment in ethical and technical capacity building.

Finally, ethical and legal policies of a computerized accountability need to be declared with good articulations by the regulators and policymakers. This incorporates the creation of principles of equity assessment, execution of disclosure of openness of AI-driven credit prescription, and facilitates the execution of fairness certification programme on an industry-wide basis. The regulators, industry players, and academia can work together to make sure that financial technologies will be the source of empowerment and not ostracize.

XIII. FUTURE DIRECTIONS

Even though the present study offers the necessary data on the connections between the concepts of algorithmic fairness, perceived trust and financial inclusion, a number of prospects are present that can be the focus of future scholarly investigation. First, the longitudinal and cross-country designs are to be used in future to pursue the evolution of the view of fairness and trust over the period of time and different regulatory and cultural contexts. Comparative analysis of the relations in mature and developing economies, could contribute to understanding the functioning of the cycle of fairness-inclusion in FinTech ecosystems and the reliance on the maturity of the institutions, the power of their digital infrastructure, and the regulatory systems.

Second, the conceptual framework would be diversified and redefined, with addition of more psychological and behavioural concepts such as perceived risk, user empowerment, or algorithmic literacy, to better reflect the human elements in the workplace that affect FinTech adoption. The analysis of the responses to the algorithmic decisions under the condition of the impact of the cognitive biases or the culture values in the subtle way could contribute to the improved understanding of the ways in which the view on the artificial intelligence in the financial sector is socially accepted by the audience.

Third, it is necessary to include mixed-methods and qualitative research techniques like in-depth interviews, focus groups, and ethnographies in the future study so that they can be used to complement quantitative SEM findings in the current paper. The techniques will be valuable to plot the specifics of user experience and ethical concerns that statistical models might overlook, and will serve to provide additional information on fairness and inclusion to the picture of the real world.

Moreover, there are new technological paradigms, like blockchain-based credit scoring, explainable artificial intelligence (XAI), and federated learning, that should be researched as soon as possible to establish their potential to devour transparency and minimize bias. The studies of the relations between the AI governance policy and the institutional ethics and the construction of technologies will be critical to the creation of responsible FinTech ecosystems.

Finally, the uniformity scholars should also endeavor to develop contextual fairness systems that would be applicable to the socio-economic environment and financial activities of local communities. There will be a requirement in developing and sharing of academia, regulators and industry professionals in transforming theoretical knowledge into action and policies and tools. In general, the real science which should be carried out in future is the process of bridging the gap between the ethical theory and the technological practice and transforming AI-based financial systems into an instrument of empowerment, inclusion and even equitable development.

REFERENCES

[1] Zahir, B., Tonmoy, B., & Md Arifur, R. (2023). UX optimization in digital workplace solutions: AI tools for remote support and user engagement in hybrid environments. International Journal of Scientific Interdisciplinary Research, 4(1), 27-51. https://doi.org/10.63125/33gqpx45

[2] Abdullah Al, M., Rajesh, P., Mohammad Hasan, I., & Zahir, B. (2022). A Systematic Review of The Role Of SQL And Excel In Data-Driven Business Decision-Making For Aspiring Analysts. American Journal of Scholarly Research and Innovation, 1(01), 249-269. https://doi.org/10.63125/n142cg62

[3] Md Masud, K. (2022). A Systematic Review Of Credit Risk Assessment Models In Emerging Economies: A Focus On Bangladesh's Commercial Banking Sector. American Journal of Advanced Technology and Engineering Solutions, 2(01), 01-31. https://doi.org/10.63125/p7ym0327

[4] Agboola, O. K. (2025). Auditing Bias in AI and Machine Learning-Based Credit Algorithms: A Data Science Perspective on Fairness and Ethics in FinTech. International Journal of Technology, Management and Humanities, 11(02), 1-11. DOI: 10.21590/ijtmh.11.02.10
[5] Kothandapani, H. P. (2022). Leveraging AI for credit scoring and financial inclusion in emerging markets. World Journal of Advanced Research and Reviews, 15(3), 526-539. https://doi.org/10.30574/wjarr.2022.15.3.0904
[6] Moscato, V., Picariello, A., & Sperlí, G. (2021). A benchmark of machine learning approaches for credit score prediction. Expert Systems with Applications, 165, 113986. https://doi.org/10.1016/j.eswa.2020.113986
[7] Sadok, H., Sakka, F., & El Maknouzi, M. E. H. (2022). Artificial intelligence and bank credit analysis: A review. Cogent Economics & Finance, 10(1). https://doi.org/10.1080/23322039.2021.2023262
[8] Memarian, B. (2023). Fairness, Accountability, Transparency, and Ethics (FATE) in Artificial Intelligence (AI) and higher education: A systematic review. Computers and Education: Artificial Intelligence, 5, 100152. https://doi.org/10.1016/j.caeai.2023.100152
[9] Herrmann, H., & Masawi, B. (2022). Three and a half decades of artificial intelligence in banking, financial services, and insurance: A systematic evolutionary review. Strategic Change. https://doi.org/10.1002/jsc.2525
[10] Khandani, A. E., Kim, A. J., & Lo, A. W. (2010). Consumer Credit-Risk Models Via Machine-Learning Algorithms. Journal of Banking & Finance, 34(11), 2767-2787. https://doi.org/10.1016/j.jbankfin.2010.06.001
[11] Jagtiani, J., & Lemieux, C. (2019). The roles of alternative data and machine learning in fintech lending: Evidence from the Lending Club consumer platform. Financial Management, 48(4), https://doi.org/10.1111/fima.12295
[12] Hertzberg, A., Liberman, A., & Paravisini, D. (2018). Screening on Loan Terms: Evidence from Maturity Choice in Consumer Credit. The Review of Financial Studies, 31(9), https://doi.org/10.1093/rfs/hhy024
[13] Gambacorta, L., Huang, Y., Qiu, H., & Wang, J. (2024). How do machine learning and non-traditional data affect credit scoring? New evidence from a Chinese fintech firm. Journal of Financial Stability, 73, 101284. https://doi.org/10.1016/j.jfs.2024.101 284
[14] Berg, T., Burg, V., Gombović, A., & Puri, M. (2019). On the Rise of FinTechs: Credit Scoring Using Digital Footprints. The Review of Financial Studies, 33(7), 2845-2897. https://doi.org/10.1093/rfs/hhz099
[15] Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., & Walther, A. (2021). Predictably Unequal? The Effects of Machine Learning on Credit Markets. The Journal of Finance, 77(1), https://doi.org/10.1111/jofi.13090
[16] Iyer, R., Khwaja, A. I., Luttmer, E. F. P., & Shue, K. (2016). Screening Peers Softly: Inferring the Quality of Small Borrowers. Management Science, 62(6), 1554-1577. https://doi.org/10.1287/mnsc.2015.2181
[17] Bello, O. A. (2023). Machine learning algorithms for credit risk assessment: an economic and financial analysis. International Journal of Management, 10(1), 109-133. https://doi.org/10.37745/ijmt.2013/vol10n1109133
[18] Effectiveness and Impact of Fintech Credit Scoring Models on Financial Inclusion 46 https://doi.org/10.13140/RG.2.2.23780.04480
[19] Oguntibeju, O. O. (2024). Mitigating artificial intelligence bias in financial systems: A comparative analysis of debiasing techniques. Asian Journal of Research in Computer Science, 17(12), 165-178. https://doi.org/10.9734/ajrcos/2024/v17i12536
[20] Salami, I. A., Adesokan-Imran, T. O., Tiwo, O. J., Metibemu, O. C., Olutimehin, A. T., & Olaniyi, O. O. (2025). Addressing bias and data privacy concerns in AI-driven credit scoring systems through cybersecurity risk assessment. Asian Journal of Research in Computer Science, 18(4), 59-82. https://doi.org/10.9734/ajrcos/2025/v18i4608
[21] Nwafor, C. N., Nwafor, O., & Brahma, S. (2024). Enhancing transparency and fairness in automated credit decisions: an explainable novel hybrid machine learning approach. Scientific Reports, 14(1), 25174. https://doi.org/10.1038/s41598-024-750268
[22] Bari, M. H. (2024). A systematic literature review of predictive models and analytics in ai-driven credit scoring. Available at SSRN 5050068. https://dx.doi.org/10.2139/ssrn.5050068
[23] Chen, Y., Calabrese, R., & Martin-Barragan, B. (2024). Interpretable machine learning for imbalanced credit scoring datasets. European Journal of Operational Research, 312(1), 357-372.https://doi.org/10.1016/j.ejor.2023.06.036
[24] Financing by Logistic Regression, Artificial Neural Network and Hybrid Models. Sustainability, 8(5), 433-NA. https://doi.org/10.3390/su8050433
[25] Dumitrescu, E., Hué, S., Hurlin, C., & Tokpavi, S. (2022). Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects. European Journal of Operational Research, 297(3), 1178-1192. https://doi.org/10.1016/j.ejor.2021.06.053