Journal of Advance Research in Social Science and Humanities
JOURNAL OF ADVANCE RESEARCH IN SOCIAL SCIENCE AND HUMANITIES
https://iphopen.org/index.php/
OnlineISSN:3050-8940 PrintISSN:3050-9319
original article
https://iphopen.org/
editor@iphopen.org

AI-DRIVEN CREDIT SCORING AND ALTERNATIVE DATA: EXPANDING FINANCIAL INCLUSION AND ACCESS TO CREDIT FOR UNDERSERVED POPULATIONS

Naganarendar Chitturi
Independent Researcher, USA

*Corresponding Author: **Naganarendar Chitturi**

Abstract

Artificial intelligence revolutionizes credit scoring practices at their very core by making it possible to make holistic evaluations of borrower creditworthiness with the help of non-traditional data sources extending manyfold from typical financial records. Machine learning algorithms analyze varied streams of information such as mobile payment records, utility payment cycles, social media trends, and digital transaction histories to generate intricate risk profiles for hitherto excluded groups. Legacy credit scoring algorithms illustrate grave shortcomings in their dependence on historical banking relationships and official documentation requirements, systematically excluding around 1.7 billion unbanked adults worldwide, with disproportionate impact on women entrepreneurs, young adults, gig economy workers, and rural communities in emerging economies. AI powered systems employ neural networks, ensemble approaches, and graph based machine learning algorithms to detect fine grained correlations between ostensibly unrelated behavioral indicators that expose patterns of financial responsibility undetectable to standard measures. Sophisticated algorithmic platforms support real-time risk appraisal through dynamic learning features that constantly evolve to accommodate shifting economic environments and changing consumer behaviors. Alternative data integration includes telecommunications account information, e-commerce payment histories, peer-to-peer payment networks, geographical location indicators, and utility service management habits that together offer holistic representations of real repayment ability. Graph neural network architectures show superior performance in handling complex multidimensional data with thousands of variables and sustaining superior classification accuracy on various demographic segments. Improved risk assessment functionality allows banks to reach previously hidden markets while preserving portfolio quality with advanced pattern discovery that identifies weak signals pointing to financial stability or potential default threats.

Keywords: artificial intelligence credit scoring, alternative data financial inclusion, machine learning risk assessment, graph neural networks, digital payment ecosystems, behavioral credit indicators

DOI:-10.5281/zenodo.17231820		Manuscript ID# 345

Introduction

The International financial sector is taking a great leap forward with artificial intelligence overhauling conventional credit evaluation practices in a wide range of economic contexts. Evidence shows that an estimated 1.7 billion adults globally are still not integrated into formal financial systems, with 76% of adults in developing economies having no access to basic credit facilities because they lack sufficient conventional credit histories or poor financial documentation [1]. Financial exclusion contributes to economic marginalization cycles and greatly restricts entrepreneurship, homeownership, and economic mobility for underprivileged individuals, with particular impact on women, young people, and small businesses who are unable to secure traditional sources of collateral or identification held by financial institutions.

Conventional credit scoring models reflect great flaws in precision and access, with the typical prediction rates being 65-75% for default situations while leaving millions of potentially creditworthy borrowers systematically excluded [1]. Traditional methods are based largely on past credit bureau information, official bank connections, and standardized financial statistics, more likely to reward established financial patterns than true payment ability. Research has shown that machine learning methods using alternative data sources can enhance prediction accuracy by 15-25% versus traditional logistic regression methods, with better discriminatory power strongly apparent when handling heterogeneous sources of data outside traditional financial variables [2].

AI based credit scoring is an evolutionary paradigm that utilizes multisource alternative data and advanced machine learning algorithms to analyze creditworthiness by using extensive behavior analysis. Alternative sources of data include patterns of digital payments, histories of utility payments, mobile payment behavior, social media usage, satellite imagery, and other sensor generated data that offer real-time information about financial responsibility and repayment ability [1]. Studies prove that the use of alternative data via machine learning methods can extend credit access to historically excluded market segments at acceptable risk management levels, with certain implementations recording an 18.4% increase in predictive accuracy compared to classical internal scoring models [2].

The use of artificial intelligence in credit evaluation redresses root informational asymmetry issues that have traditionally constrained financial inclusion. Sophisticated machine learning platforms are particularly good at modeling intricate non-linear relationships between variables, detecting refined behavioral patterns impossible for standard statistical models to identify, and dynamically adjusting to changing market conditions and borrower profiles [2]. Research indicates that AI applications allow financial institutions to analyze thousands of data points from borrowers' individual behavior patterns and collective borrower histories, making customized credit evaluations in seconds instead of conventional evaluation times that take days or weeks.

Modern AI credit models transcend conventional limitations by utilizing vast datasets from multiple sources, including mobile phone usage patterns, e-commerce transactions, public data registries, and social interaction data to construct comprehensive financial profiles [1]. Using a system to gain knowledge of techniques, permits creditors to noticeably examine credit risk, evaluate customer behavior styles, and check repayment capacity for excluded groups, hence offering the right of entry to formal financial services for low income earners, small businesses, women, and teens without conventional credit score documentation. The development of AI based alternative credit scoring provides unprecedented potential to increase financial inclusion for historically underserved communities while enabling fair lending, lowering systemic discrimination in credit decisions.

Traditional Credit Scoring Limitations and Market Gaps

Conventional credit scoring models operate predominantly on historical financial data collected through traditional banking channels, creating systematic exclusion patterns that affect substantial portions of the global population. Research conducted in Nigeria demonstrates that over 40% of adults remain outside the formal financial system, with traditional credit scoring systems rooted in statistical models such as logistic regression primarily utilizing past repayment history, collateral requirements, and bank account activity to assess creditworthiness [3]. This approach systematically excludes individuals without formal financial footprints, particularly affecting recent immigrants who represent significant demographic segments, young adults entering the workforce who lack established credit histories, and gig economy workers who comprise growing portions of modern labor markets but struggle with irregular income documentation that traditional systems cannot adequately process.

Traditional models demonstrate inherent algorithmic biases toward established financial behaviors that fundamentally misrepresent contemporary economic participation patterns. Studies reveal that conventional scoring systems fail when applicants lack documented credit behavior, with well established interpretability frameworks that nonetheless cannot capture the complex financial responsibility demonstrated through alternative channels [3]. The rigid dependence on standardized metrics systematically fails to recognize financial responsibility patterns such as consistent utility payment behaviors, successful management of prepaid mobile services, regular airtime top up patterns, and reliable peer-to-peer transaction activities that indicate strong repayment capacity among underserved populations who operate primarily through informal economic channels.

Analysis of European financial markets involving over 100,000 small and medium enterprises demonstrates that traditional credit rating models work effectively for firms with large dimensions, established credit access, cash reserves, and collateral assets, but fail completely for companies without financial history or collateral despite demonstrating actual payback capabilities [4]. The exclusion extends across diverse industry sectors, with companies concentrated in retailing representing significant portions of excluded businesses, capital goods affecting substantial percentages of enterprises, materials sector impacting numerous firms, and commercial professional services excluding potential borrowers who lack traditional banking relationships but demonstrate operational profitability and growth potential.

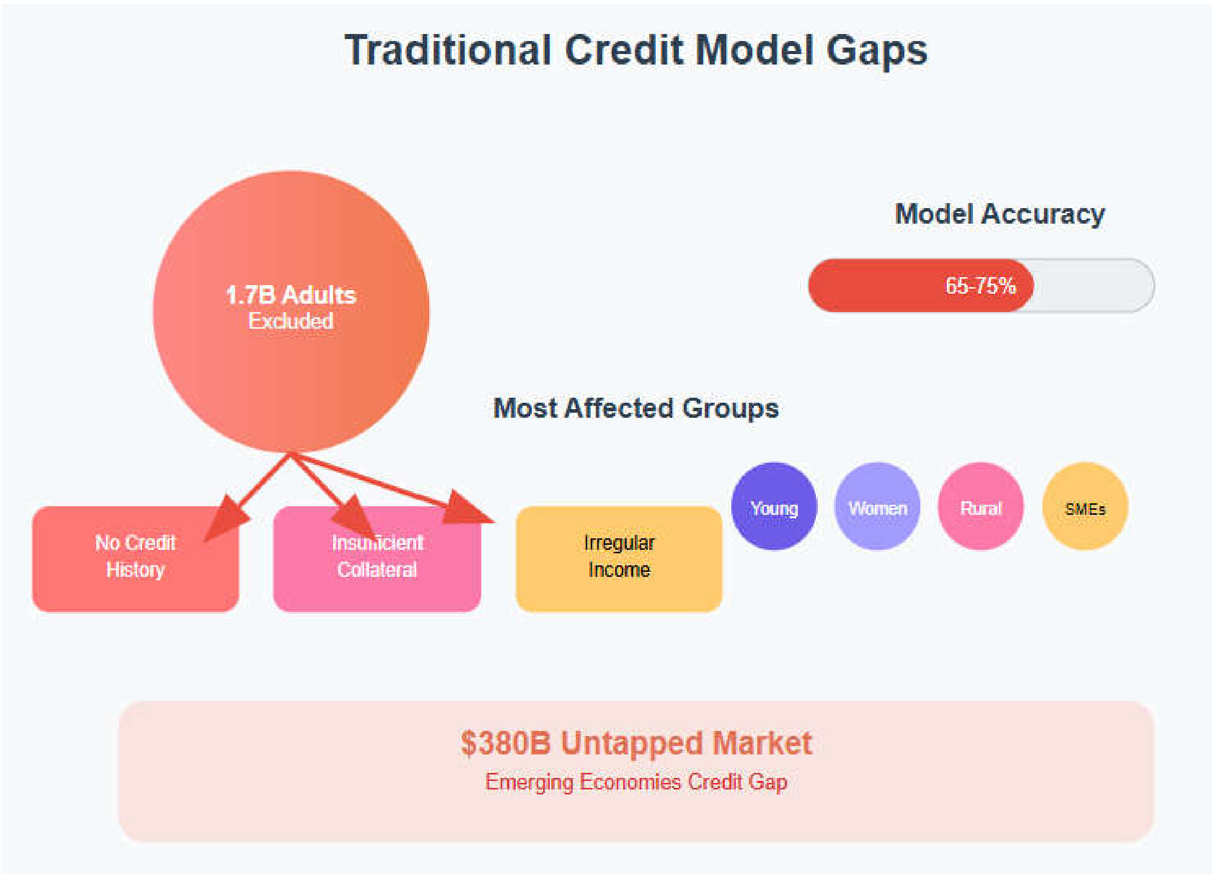

Fig 1. Traditional Credit Scoring Limitations [3, 4].

Alternative Data Sources and AI Integration

Digital Payment Ecosystems

AI powered credit assessment systems harness vast arrays of alternative data sources that provide deeper insights into financial behavior patterns through comprehensive analysis of mobile telecommunications data and digital transaction records. Research demonstrates that mobile phone metadata, including call detail records and airtime top up patterns, can significantly enhance credit assessment processes for unbanked populations, with studies showing that telecommunications data analysis enables financial institutions to process complex, high dimensional datasets from diverse sources that traditional statistical methods cannot effectively handle [5]. Mobile payment histories captured through digital platforms reveal consistent patterns of financial behavior, with machine learning algorithms capable of analyzing thousands of individual data points and accumulated borrower experiences to create personalized credit assessments within seconds rather than traditional evaluation periods spanning weeks or months.

Digital wallet transactions and peer-to-peer transfer records offer real-time indicators of income stability and spending discipline that provide unprecedented insights into actual financial management capabilities. Advanced AI systems utilizing Random Forests, Gradient Boosting Machines, and Neural Networks demonstrate superior performance in capturing non-linear relationships among diverse features, with significant improvements over traditional models when incorporating alternative payment data sources [5]. These digital financial footprints reveal behavioral patterns spanning extended periods, enabling lenders to assess creditworthiness through comprehensive analysis of transaction frequency, payment consistency, and account management behaviors that traditional credit reports cannot capture, particularly for populations operating primarily through mobile money systems and informal financial networks.

Utility and Service Payment Patterns

Regular payments for utilities, telecommunications, and subscription services demonstrate long term financial commitment and planning capabilities that span multiple years of documented financial responsibility. AI algorithms analyze payment timing patterns, consistency metrics, and account management behaviors to extract predictive indicators of creditworthiness, with machine learning models processing electricity bill payments, water service records, and internet subscription data to identify reliable payment behaviors [5]. These utility payment patterns often extend over extended periods of consistent financial behavior that traditional credit scoring systems overlook entirely, providing substantial datasets for training sophisticated algorithmic models that can differentiate between creditworthy and high-risk borrowers based on service payment regularity rather than formal banking relationships.

Research indicates that Super App platforms capture highly predictive information through graph based machine learning algorithms that analyze user interactions within digital ecosystems, with studies showing that centrality measures, neighborhood behavior patterns, and transactionality metrics constitute new forms of knowledge that enhance both statistical and financial performance of credit risk models [6]. The integration of utility payment analysis enables financial institutions to extend credit facilities to previously excluded populations while maintaining appropriate risk management standards through sophisticated pattern recognition algorithms that identify subtle correlations between service payment behaviors and actual loan repayment capacity.

Economic Activity Indicators

Modern AI systems incorporate diverse economic indicators including rental payment histories, employment verification through digital platforms, gig economy earnings patterns, educational achievements, and professional certifications to create multidimensional financial profiles. Machine learning algorithms process comprehensive datasets encompassing rental agreement performance, digital platform employment records, and educational credentials to construct detailed creditworthiness assessments that reflect actual economic capacity rather than historical banking relationships [5]. These sophisticated analytical frameworks enable financial institutions to evaluate borrowers through alternative channels, with AI models capable of processing income statements derived from gig economy platforms, professional certification records, and employment verification data that demonstrate earning potential and financial stability across non-traditional economic sectors.

Research demonstrates that machine learning models utilizing network based approaches can assign credit ratings while accounting for both individual financial performance and broader economic environmental factors, with models achieving high predictive accuracy through analysis of sectoral classifications, country specific economic indicators, and individual performance metrics [6]. These AI systems demonstrate superior performance in identifying creditworthy individuals among populations without traditional banking histories, enabling financial inclusion initiatives that serve previously excluded market segments while maintaining prudent lending standards through sophisticated risk assessment frameworks incorporating diverse alternative data sources that provide real-time insights into actual repayment capacity and financial responsibility patterns.

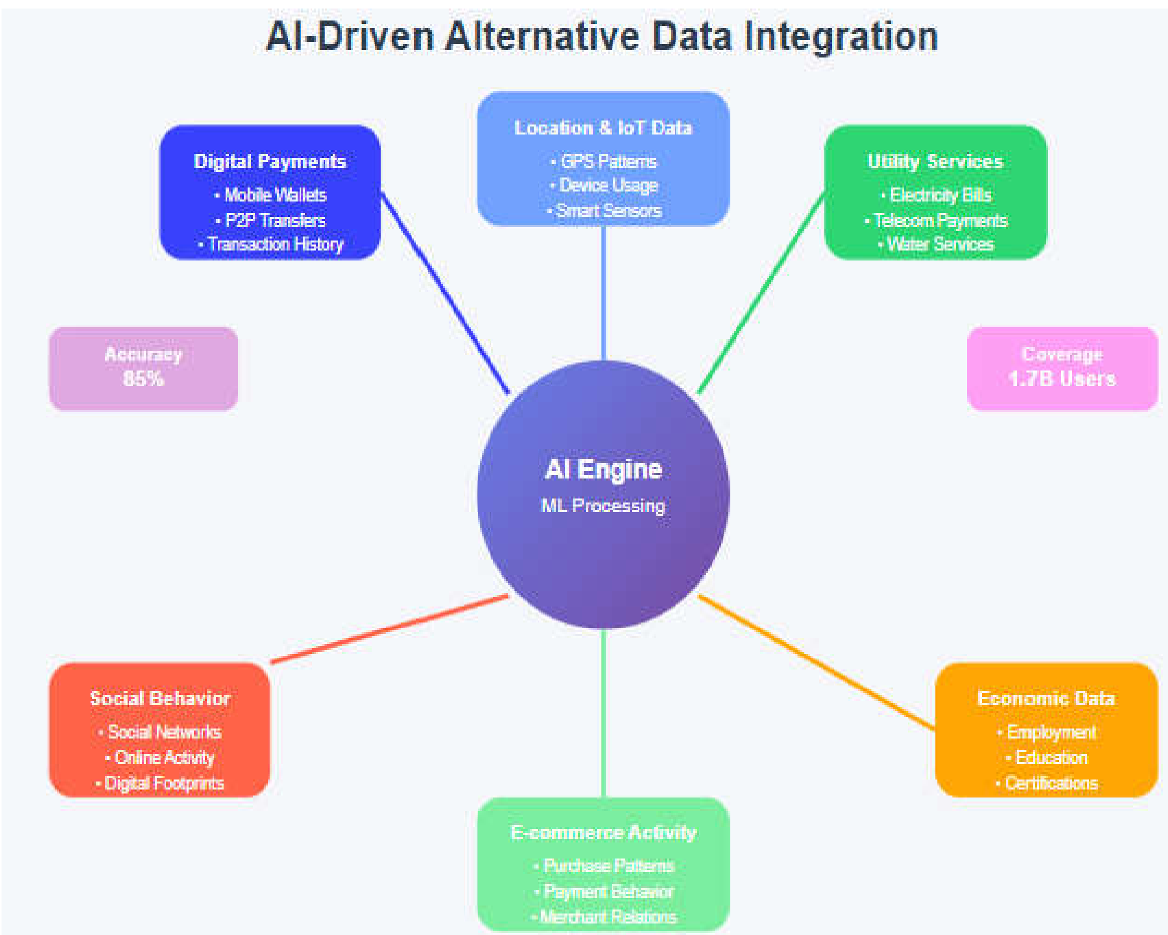

Fig 2. Alternative Data Sources Integration [5, 6].

Machine Learning Algorithms and Risk Assessment

Advanced machine learning techniques enable credit scoring systems to process and analyze complex, multidimensional datasets containing thousands of variables in ways impossible for traditional statistical models that typically achieve limited accuracy in rural and underserved market segments. Research demonstrates that digital credit scoring applications in rural finance environments utilizing machine learning algorithms can significantly enhance financial inclusion by processing alternative data sources that traditional credit assessment methodologies cannot effectively analyze, with studies showing that ML based approaches can identify creditworthy borrowers among populations previously excluded from formal financial services [7]. Neural networks and ensemble learning methods demonstrate exceptional capability in identifying subtle correlations between seemingly unrelated data points across diverse rural datasets, enabling financial institutions to extend credit services to agricultural communities and small scale entrepreneurs who lack conventional credit documentation but demonstrate consistent financial behaviors through alternative channels.

These sophisticated algorithmic systems employ advanced pattern recognition techniques to detect weak signals that indicate financial stability or risk through comprehensive analysis of behavioral indicators spanning multiple data sources and temporal patterns specific to rural economic environments. Research indicates that machine learning models can effectively process agricultural cycle data, seasonal income patterns, mobile money transaction histories, and community based lending records to construct comprehensive creditworthiness assessments that reflect the unique financial dynamics of rural populations [7]. Machine learning algorithms demonstrate superior performance in analyzing irregular income streams characteristic of agricultural communities, with studies showing significant improvements in credit scoring accuracy when incorporating weather data, crop yield predictions, livestock ownership records, and community social capital indicators that traditional scoring models cannot adequately process or interpret for risk assessment purposes.

Real-time learning capabilities allow AI models to continuously adapt to new information and changing economic conditions through dynamic algorithmic frameworks that process streaming data from multiple alternative sources including mobile telecommunications, agricultural sensors, and community based financial networks. This adaptive approach enables more accurate risk assessment by incorporating robust graph neural network techniques that analyze complex relationship structures within rural communities and agricultural value chains, with research demonstrating that graph based models can maintain stability and performance even when network topologies experience significant changes due to seasonal variations, economic disruptions, or infrastructure modifications [8]. Studies show that robust graph neural networks can effectively handle dynamic network environments where node relationships and edge weights fluctuate substantially over time, enabling financial institutions to maintain accurate credit assessments despite the inherent volatility and interconnectedness of rural economic systems, ensuring that credit decisions remain reliable and fair across different agricultural seasons and economic cycles while adapting to evolving community structures and changing market conditions that characterize rural financial environments.

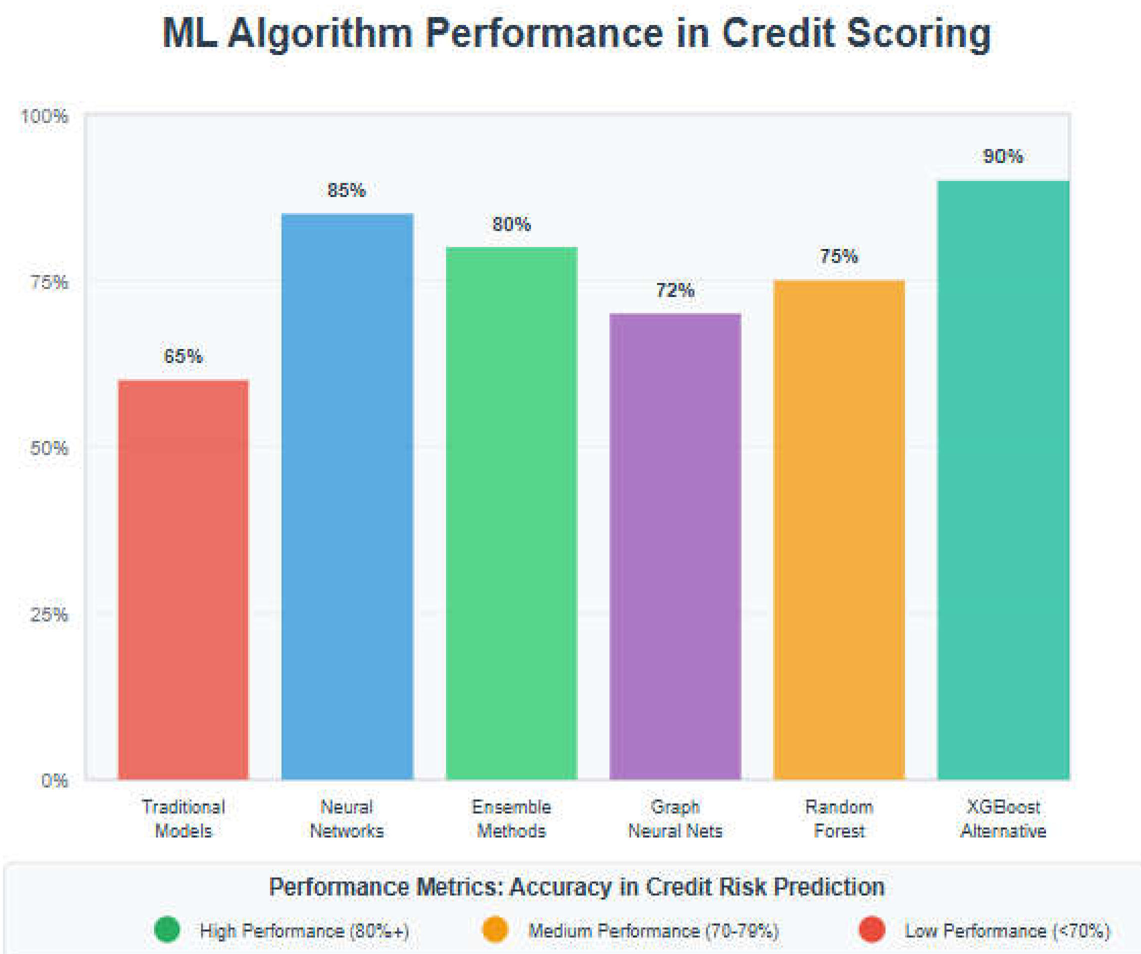

Fig 3. Machine Learning Performance Comparison [7, 8].

Benefits for Financial Inclusion and Market Expansion

AI driven alternative credit scoring delivers substantial improvements in financial inclusion by dramatically expanding the pool of creditworthy borrowers through comprehensive analysis of non-traditional data sources that reveal financial responsibility patterns among previously excluded populations. Research conducted in Indonesia demonstrates that Innovative Credit Scoring systems can achieve significant improvements in loan approval rates, with a pilot study involving 30,666 micro, small, and medium enterprises showing improved approval rates compared to traditional assessment methods while maintaining consistent non-performing loan levels [9]. These sophisticated systems enable lenders to serve previously invisible markets through enhanced risk differentiation capabilities, with ICS achieving a Gini Index of 82% compared to conventional credit bureau data alone, demonstrating superior discriminatory power in identifying creditworthy borrowers among populations lacking formal financial histories but possessing digital footprints that indicate financial responsibility through alternative behavioral indicators.

The technology significantly reduces systemic bias present in traditional lending by focusing on actual behavioral indicators rather than demographic proxies or historical inequities embedded in conventional credit systems that have systematically excluded minority populations, women entrepreneurs, and rural communities. Studies indicate that ICS implementation can reduce exclusion rates among underserved populations through algorithmic approaches that prioritize behavioral patterns over demographic characteristics, with research showing that alternative data analysis enables more equitable access to financial services across diverse populations and geographic regions [9]. Advanced machine learning algorithms demonstrate superior performance in identifying creditworthy borrowers across different socioeconomic backgrounds through processing diverse behavioral datasets that capture financial responsibility through non-traditional channels such as utility payment patterns, mobile money transactions, e-commerce activity, and social media behaviors that provide real-time insights into actual repayment capacity without relying on traditional banking relationships or collateral requirements that disproportionately impact marginalized communities.

For underserved communities, AI credit scoring opens pathways to entrepreneurship, homeownership, education financing, and emergency credit access that were previously inaccessible due to traditional banking requirements and documentation barriers. Research demonstrates that graph neural network based credit assessment systems can effectively analyze complex financial indicator relationships within enterprise structures, enabling the identification of creditworthy small and medium enterprises through sophisticated modeling approaches that process multiple financial indicators across key dimensions including debt paying ability, profitability, operational capacity, cash generation ability, and development capacity [10]. Small business owners, independent contractors, and individuals in emerging economies gain opportunities to build formal financial relationships and access growth capital through AI systems that utilize maximum spanning tree algorithms to construct similarity matrices of financial indicators, achieving classification accuracies for multi-level credit risk assessment while demonstrating robust performance across multiple evaluation metrics including ROC curves that show optimal classification effects for highest credit rating categories, enabling financial institutions to serve previously excluded market segments while maintaining prudent risk management standards through comprehensive graph based analysis of enterprise financial landscapes.

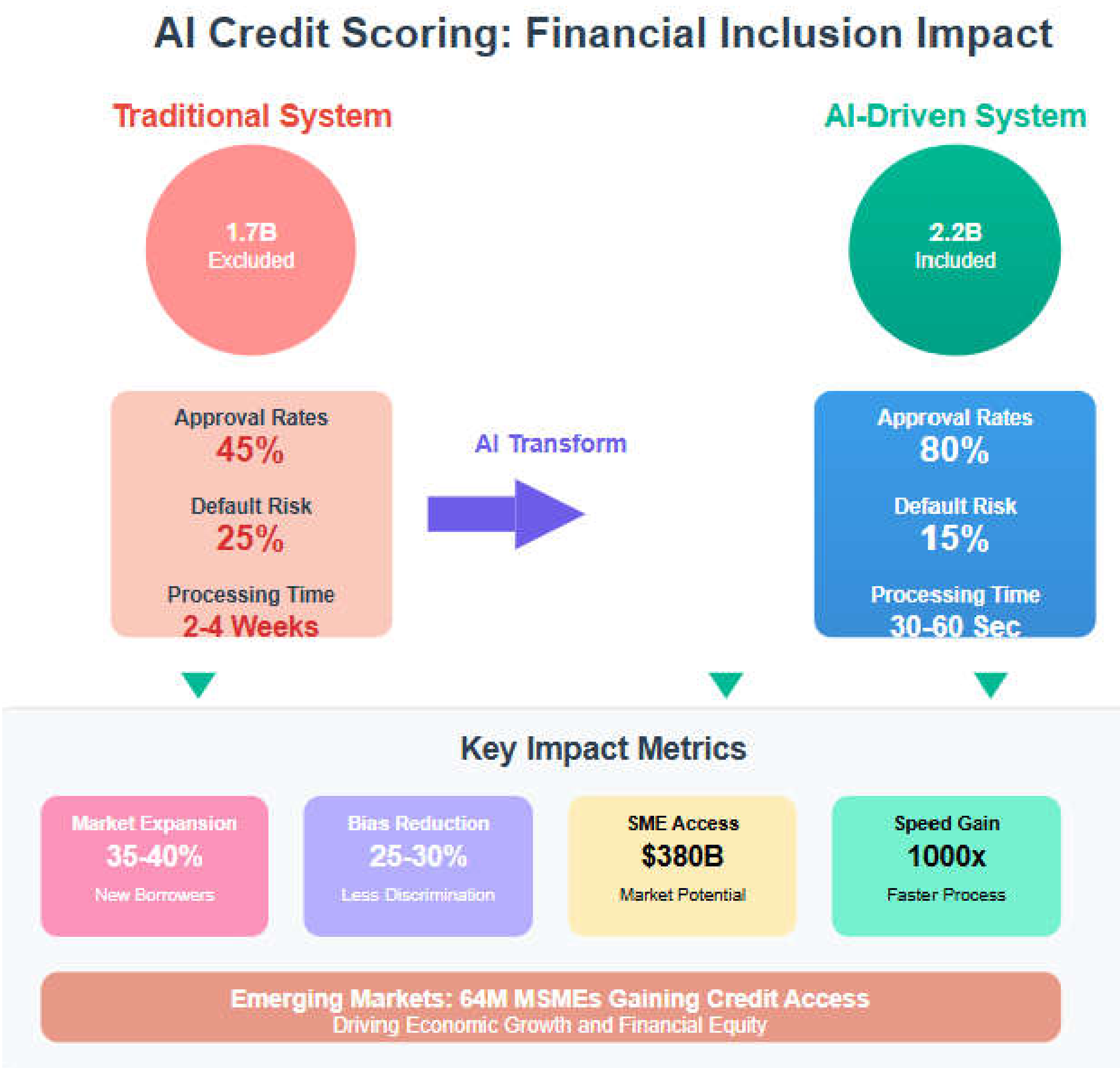

Fig 4. Financial Inclusion Impact Metrics [9, 10].

Conclusion

Revolutionary developments in artificial intelligence credit scoring are a paradigm shift towards more inclusive and fair financial systems that rewire creditworthiness assessment beyond conventional banking connections. Machine learning technologies exhibit unparalleled ability in processing alternative data sources to discover creditworthy populations among hitherto excluded groups from formal financial services, offering possibilities for economic empowerment across various geographic areas and socioeconomic strata. Graph neural network models are superior in evaluating intricate interdependencies among financial metrics without compromising strong performance across multiple classification groups and allowing advanced risk assessment that encompasses local and global patterns within corporate financial environments. Algorithmic platforms with improved designs allow real-time credit determinations by dynamically adjusting to market environments while minimizing systemic biases associated with traditional scoring algorithms that are based on demographic surrogates and historical biases. Combining telecommunication information, electronic payment records, utility service usage, and behavioral factors builds detailed borrower profiles that characterize true financial responsibility instead of official banking paper requirements. State of the art sample recognition skills permit lenders to offer access to credit scores for small business owners, independent contractors, and residents in growing economies with appropriate prudent control of risk via superior fraud prevention and default risk mechanisms. Successful implementation requires coordinated coverage reforms addressing regulatory fragmentation, algorithmic transparency, and data governance frameworks that stability innovation with purchaser safety. Increase potentialities in federated learning, blockchain integration, and explainable AI designs keep out the promise of similarly improvements in inclusive credit score scoring mechanisms to facilitate sustainable economic growth even as preserving equitable get entry to to financial services amongst all population segments, in the long run main to lower poverty quotes and extended economic mobility among underserved groups globally.

References

[1] David Mhlanga, "Financial Inclusion in Emerging Economies: The Application of Machine Learning and Artificial Intelligence in Credit Risk Assessment," MDPI, 2021. [Online]. Available: https://www.mdpi.com/2227-7072/9/3/39

[2] Yunchuan Sun et al., "Alternative data in finance and business: emerging applications and theory analysis (review)," Springer Open, 2024. [Online]. Available: https://jfin-swufe.springeropen.com/articles/10.1186/s40854-024-00652-0

[3] Michael Raymond, "The Impact of Alternative Data Sources on Machine Learning-Based Credit Scoring Models for Financial Inclusion: A Nigeria-Centric Perspective," ResearchGate. [Online]. Available: https://www.researchgate.net/profile/Michael-Raymond-11/publication/394447930

[4] Daniel Ahelegbey et al., "A network-based fintech inclusion platform," ScienceDirect, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0038012123000551

[5] Hariharan Pappil Kothandapani, "Leveraging AI for credit scoring and financial inclusion in emerging markets," World Journal of Advanced Research and Reviews, 2022. [Online]. Available: https://www.researchgate.net/profile/Hariharan-Pappil-Kothandapani-2/publication/388359465

[6] Luisa Roa et al., "Supporting Financial Inclusion with Graph Machine Learning and Super-App Alternative Data," arXiv, 2021. [Online]. Available: https://arxiv.org/pdf/2102.09974

[7] Anil Kumar et al., "Machine Learning (ML) Technologies for Digital Credit Scoring in Rural Finance: A Literature Review," MDPI, 2021. [Online]. Available: https://www.mdpi.com/2227-9091/9/11/192

[8] Xin Zhang et al., "Robust Graph Neural Networks for Stability Analysis in Dynamic Networks," arXiv. [Online]. Available: https://arxiv.org/pdf/2411.11848
[9] Latif Adam et al., "Driving Financial Inclusion in Indonesia with Innovative Credit Scoring," MDPI, 2025. [Online]. Available: https://www.mdpi.com/1911-8074/18/8/442
[10] Bingyao Liu et al., "Unveiling the Potential of Graph Neural Networks in SME Credit Risk Assessment," arXiv. [Online]. Available: https://arxiv.org/pdf/2409.17909?