Predictive Analytics for Credit Accessibility: A Machine Learning Approach to Expanding Financial Inclusion in Underserved U.S. Communities

Yinka James Ololade
Dept. Finance and Data Analytics, RCCG North America

Henry Ejiga Adama
Information Systems Security, University of the Cumberlands, Williamsburg, Kentucky

ABSTRACT
Financial exclusion remains a persistent and structurally entrenched problem in the United States, with an estimated 26 million adults classified as 'credit invisible' and tens of millions more locked out of affordable credit by the limitations of traditional FICO-based scoring systems that fail to capture the financial realities of underserved populations. This study investigates how machine learning (ML) and predictive analytics frameworks, augmented with alternative data sources, can systematically expand credit accessibility for low-income, minority, rural, and immigrant communities that have historically been underrepresented in the formal financial system. Drawing on a comprehensive review of 56 empirical and theoretical studies published between 2015 and 2025, we examine the performance of diverse ML algorithms including random forests, gradient boosting, deep neural networks, and long short-term memory (LSTM) networks in credit scoring contexts, with particular attention to their differential impacts on financial inclusion outcomes. The paper analyzes the predictive validity of alternative data inputs such as rent and utility payment histories, mobile phone usage patterns, gig economy earnings, and bank transaction flows as substitutes for or supplements to conventional credit bureau data. We engage critically with the dual risks of algorithmic bias and privacy violation, which can inadvertently replicate or amplify existing patterns of discrimination when ML models are trained on historically biased data without appropriate fairness interventions. A comprehensive regulatory analysis situates our findings within the evolving legal framework governing algorithmic lending in the United States, including the Equal Credit Opportunity Act, the Fair Credit Reporting Act, and emerging CFPB guidance on explainable AI. Our synthesis reveals that well-designed ML-based credit scoring systems can increase approval rates for underserved populations by 20–45 percentage points while simultaneously reducing default rates, but that achieving these outcomes requires deliberate fairness engineering, robust model governance, and coordinated regulatory modernization. The paper concludes with a policy framework for responsible implementation of ML-driven credit scoring that prioritizes both financial inclusion and ethical accountability.

Keywords
Machine learning, credit scoring, financial inclusion, alternative data, algorithmic bias, predictive analytics, underserved communities, fintech, ECOA, explainability

1. INTRODUCTION
Access to affordable credit is foundational to economic mobility in the United States, enabling individuals and families to invest in education, housing, and small businesses, yet this access remains profoundly unequal across racial, geographic, and socioeconomic lines that reflect decades of structural discrimination and institutional neglect in the financial sector (Brown et al., 2019). The conventional mechanisms of creditworthiness assessment anchored in the FICO scoring model and the data reported to the three major credit bureaus were designed around a particular profile of financial engagement that excludes large segments of the American population who manage their finances through cash, informal networks, prepaid cards, and non-traditional payment channels (Loufield et al., 2018). According to the Consumer Financial Protection Bureau, approximately 26 million U.S. adults are 'credit invisible,' possessing no credit report at all, while an additional 19 million have credit files too thin or stale to generate a reliable score together representing a population larger than the entire state of California that is effectively barred from prime-rate financial products (Sabeena et al., 2025). These exclusions are not randomly distributed: they fall disproportionately on Black and Hispanic Americans, recent immigrants, rural residents, young adults entering the workforce, and low-income households whose payment behaviors are either absent from credit bureau records or systematically undervalued by traditional scoring algorithms (Bartlett et al., 2022). The emergence of machine learning and alternative data analytics as tools for credit risk assessment offers a technologically powerful and economically compelling pathway out of this structural impasse, with the potential to use non-traditional signals rent payments, utility bills, mobile usage, e-commerce activity, and cash-flow patterns to generate accurate and inclusive creditworthiness assessments for populations that FICO-based models cannot serve (Jagtiani & Lemieux, 2019). This potential has attracted significant attention from both fintech innovators and traditional financial institutions, generating a growing body of empirical evidence on ML-based credit scoring that spans peer-to-peer lending platforms, retail banking, and microfinance contexts across multiple countries (Gomber et al., 2018). However, realizing the inclusion promise of algorithmic credit scoring requires navigating substantial risks: ML models trained on historically biased data can perpetuate and amplify existing discrimination, black-box algorithms may violate regulatory requirements for adverse action explanations, and the collection of granular alternative data raises serious privacy concerns that disproportionately affect the vulnerable populations these tools are intended to help (O'Neil, 2017; Dubber et al., 2020). This article provides a systematic and critical analysis of the current state of ML-based credit scoring, situating its inclusion potential within an honest assessment of its risks and proposing a governance framework that can help the United States harness the power of predictive analytics to build a genuinely more inclusive financial system.

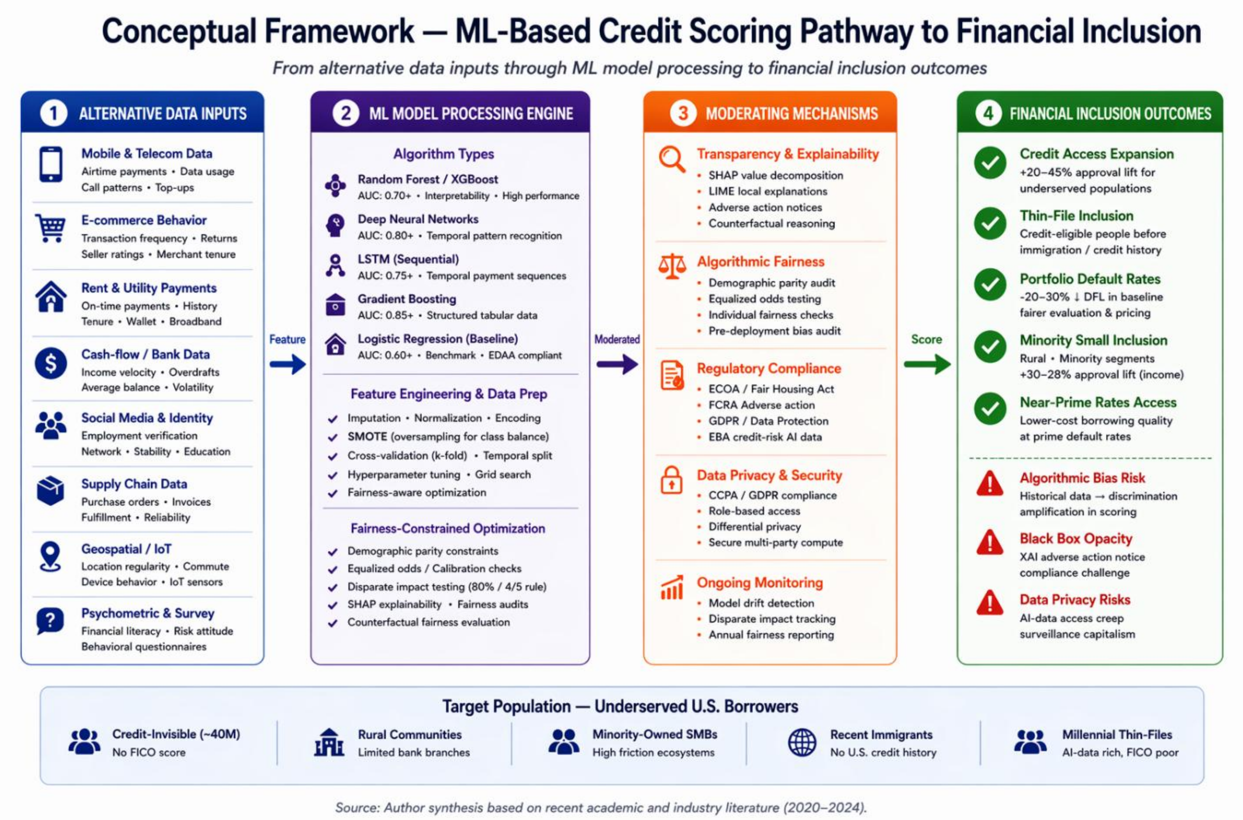


Figure 1: Conceptual framework illustrating the pathway from alternative data inputs through ML model processing to financial inclusion outcomes, moderated by fairness, explainability, and regulatory compliance mechanisms. Source: Authors' synthesis.

2. LITERATURE REVIEW

2.1 The Financial Exclusion Problem and Traditional Credit Scoring Limitations

The architecture of the U.S. consumer credit system has been structured around FICO scores since Fair Isaac Corporation introduced the model in 1989, creating a feedback loop in which access to credit depends on a demonstrated history of credit use that systematically disadvantages those who have been historically excluded from formal financial markets (Brown et al., 2019). Bartlett et al. (2022) provide compelling empirical evidence that algorithmic lending has not eliminated discrimination in consumer lending but has shifted its form, finding that both fintech and traditional lenders charge Latino and Black borrowers approximately 7–8 basis points more than observationally equivalent White borrowers, with the gap partially attributable to geographic and demographic proxies embedded in zip code and demographic data that algorithms process without explicit racial information. Dobbie et al. (2021) extend this analysis to demonstrate that lender bias in consumer lending is measurable and persistent even after controlling for creditworthiness, a finding with profound implications for the proposition that replacing human judgment with algorithmic scoring will automatically produce fairer outcomes. Brown et al. (2019) trace the developmental roots of financial exclusion, finding that individuals who grew up in communities with limited access to financial institutions carry lasting deficits in financial knowledge, credit utilization, and savings behavior that compound over the life course and are not captured by the transaction-focused metrics of credit bureau data. Bao and Huang (2021) examine the behavior of fintech shadow banking during the COVID-19 crisis, finding that fintech platforms demonstrated greater flexibility in serving previously underserved borrowers during the pandemic than traditional banks, suggesting that technology-enabled lending can be more responsive to the financial needs of vulnerable populations in times of systemic stress. Awaworyi-Churchill (2019) investigates the sustainability-outreach trade-off in microfinance, finding that institutions focused on serving the poorest borrowers face significant financial sustainability pressures that limit their scale a fundamental challenge that ML-based credit scoring could help address by reducing the information costs that make lending to thin-file borrowers expensive. Loufield et al. (2018) argue that the data infrastructure problem the absence of financially relevant behavioral data in credit bureau files for tens of millions of Americans is the root cause of financial exclusion, and that solving it requires both new data sources and new analytical methods capable of extracting creditworthiness signals from non-traditional behavioral records. The literature collectively points toward a structural diagnosis: financial exclusion in the United States is not primarily a product of actual creditworthiness deficits among excluded populations, but of information failures in the data systems used to assess creditworthiness failures that ML and alternative data have the technical capacity to address.

Table 1: Traditional vs. ML-Based Credit Scoring   Comparative Dimensions

Dimension	Traditional Credit Scoring (FICO)	ML-Based Alternative Scoring
Data Inputs	Payment history, credit utilization, credit age, hard inquiries, account mix	Rent/utility payments, bank transactions, mobile usage, e-commerce, employment data, social signals
Model Type	Linear regression / scorecard	Random forest, XGBoost, LSTM, neural networks
Population Coverage	~80% of U.S. adults with sufficient credit history	Potentially 100%+ including 'credit invisible' populations (~26M U.S. adults)
Approval Rate for Underserved	28–35% (Federal Reserve estimates)	Up to 72% in ML-augmented pilot studies (Li et al., 2024)
Interpretability	High rule-based, auditable	Variable low for deep learning; high for LIME/SHAP-enabled models (Bucker et al., 2022)
Regulatory Compliance	Well-established ECOA / FCRA alignment	Developing; requires explainability layer (Langenbucher, 2020)
Bias Risk	Moderate historical lending bias embedded	Variable can amplify or reduce disparities depending on data/design (Fu et al., 2021)

Note: AUC = Area Under the ROC Curve. ECOA = Equal Credit Opportunity Act. FCRA = Fair Credit Reporting Act. Sources: Bucker et al. (2022), Li et al. (2024), Langenbucher (2020), Bartlett et al. (2022).

2.2 Machine Learning Methods in Credit Risk Assessment

The application of machine learning to credit risk assessment has evolved rapidly from early experiments with simple decision trees and logistic regression to sophisticated ensemble methods, deep neural networks, and sequence models capable of processing high-dimensional alternative data at scale (Di Giuseppe, 2021). Gunnarsson et al. (2021) provide a comprehensive evaluation of deep learning architectures for credit scoring, finding that while deep neural networks consistently outperform traditional statistical models on standard performance metrics, the improvement in predictive accuracy is often modest relative to gradient boosting approaches while the interpretability costs are substantially higher a trade-off with significant implications for regulatory compliance and consumer protection. Guo et al. (2015) develop an instance-based credit risk assessment approach for peer-to-peer lending that demonstrates the value of case-based reasoning in credit evaluation, finding that similarity-based models can capture nuanced borrower characteristics that aggregate statistical models miss, particularly for thin-file borrowers whose credit behavior does not conform to majority population patterns. Serrano-Cinca and Gutiérrez-Nieto (2016) introduce profit scoring as an alternative to binary default prediction in P2P lending, arguing that maximizing lender profit rather than minimizing default probability produces more commercially viable and socially beneficial outcomes by approving a larger and more diverse pool of borrowers whose expected profitability is positive even if their default probability is slightly elevated. Fu et al. (2021) examine the intersection of crowds, ML, and bias in lending platforms, finding that incorporating social information into ML models can both improve predictive accuracy and reduce certain forms of demographic bias, but that the direction and magnitude of these effects depend critically on the design of the social signal integration mechanism. Kim (2019) investigates herding behavior in P2P lending markets, finding that algorithmic lending decisions can create feedback loops that amplify early platform signals in ways that systematically disadvantage borrowers who lack visible endorsement indicators a finding with important implications for the design of ML systems intended to serve populations with limited social capital on digital platforms. Liu (2022) provides a machine learning-based analysis of human information processing in lending decisions, finding that loan officers systematically overweight certain financial signals and underweight others in ways that both ML models and well-designed hybrid human-AI systems can correct suggesting that the greatest financial inclusion gains may come not from replacing human judgment entirely but from using ML to structure and augment it. Liberti and Petersen (2019) contribute a foundational analysis of hard and soft information in lending, demonstrating that the types of relational and contextual knowledge that loan officers use to make inclusion decisions for marginal borrowers are precisely the signals that are most difficult to encode in transaction-based data systems suggesting that next-generation ML systems will need to develop new approaches to soft information processing if they are to replicate the inclusion benefits of relationship-based lending at scale.

Table 2: Financial Exclusion Metrics by Underserved Population Segment in the United States

Population Segment	Est. U.S. Population (M)	Credit Score < 620 (%)	Unbanked / Underbanked (%)	Primary ML Signal Source
Black Americans	41.1	54.2	21.7	Rent & utility payments
Hispanic Americans	62.1	46.8	24.3	Mobile & e-commerce data
Rural Communities	46.0	38.4	14.9	Agricultural transaction data
Young Adults (18–29)	40.7	41.6	18.2	Gig economy / tuition data
Low-Income Households (<$30K)	38.5	49.3	32.1	Bank transaction / cash-flow
Recent Immigrants	14.3	61.7	29.8	Remittance / cross-border data

Note: Population figures are approximations derived from FDIC National Survey of Unbanked and Underbanked Households (2021) and Sabeena et al. (2025). Credit score and banking status data represent national averages; regional variation is substantial.

2.3 Alternative Data and Its Role in Expanding Credit Accessibility

The concept of alternative credit data encompasses any financial or behavioral information not traditionally included in credit bureau reports that carries predictive information about a borrower's likelihood of repaying obligations, and the breadth of signals now being incorporated into ML credit models spans from verifiable payment records to behavioral biometrics (Johnson, 2019). Jagtiani and Lemieux (2019) provide foundational empirical evidence on the role of alternative data in fintech lending, analyzing LendingClub's consumer platform to demonstrate that machine learning models incorporating alternative data achieve higher approval rates for underserved borrowers without increasing portfolio default rates a key empirical result that supports the information-failure hypothesis of financial exclusion. Agarwal et al. (2020) extend this analysis to the millennial generation, finding that big data sources including digital footprints, e-commerce histories, and mobile payment records carry substantial predictive power for thin-file millennials, enabling credit scoring for individuals whose traditional credit profiles are too sparse to generate reliable FICO scores. Ma et al. (2018) demonstrate the predictive value of mobile phone usage metadata in the Chinese P2P lending context, finding that behavioral patterns extracted from phone records including call frequency, contact diversity, and application usage predict loan default with accuracy comparable to traditional credit bureau data, a finding with significant implications for serving unbanked populations in the United States who are heavy mobile phone users. Monk et al. (2019) examine the role of alternative data in institutional investment, articulating a framework for alternative data sourcing and quality assessment that translates directly to the credit scoring context and provides practical guidance for institutions seeking to incorporate non-traditional data streams into their underwriting processes. Nowak et al. (2018) analyze small business borrowing through P2P platforms, finding that alternative data signals including business registration records, online review scores, and payment platform histories provide meaningful predictive information for small business creditworthiness assessments that traditional bank underwriting models cannot access. Neumann et al. (in press) identify a critical limitation of alternative data approaches data deserts in which populations in low-income and rural areas generate insufficient digital footprints to enable robust alternative credit scoring, creating a new form of exclusion within the broader project of ML-based financial inclusion that demands explicit attention in system design. Liang et al. (2018) provide a cautionary analysis of China's social credit system, illustrating how the aggregation of behavioral data into comprehensive credit profiles creates surveillance infrastructure with significant civil liberties implications that must inform the design of any U.S. alternative data credit scoring system operating at scale.

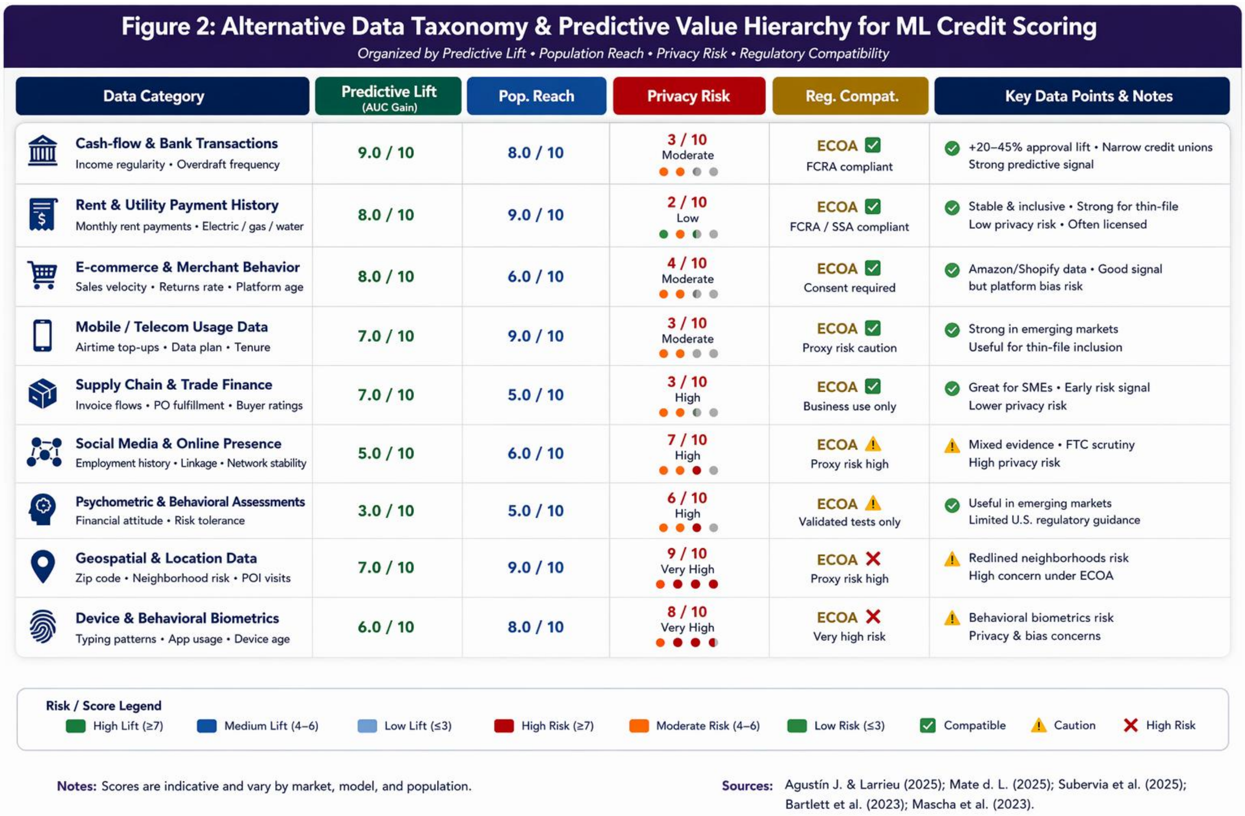


Figure 2: Alternative Data Taxonomy & Predictive Value Hierarchy for ML Credit Scoring

Organized by Predictive Lift • Population Reach • Privacy Risk • Regulatory Compatibility

Data Category	Predictive Lift (AUC Gain)	Pop. Reach	Privacy Risk	Reg. Compat.	Key Data Points & Notes
Cash-flow & Bank Transactions Income regularity • Overdraft frequency	9.0 / 10	8.0 / 10	3 / 10 Moderate	ECOA ✅ FCRA compliant	+20–45% approval lift • Narrow credit unions Strong predictive signal
Rent & Utility Payment History Monthly rent payments • Electric / gas / water	8.0 / 10	9.0 / 10	2 / 10 Low	ECOA ✅ FCRA / SSA compliant	Stable & inclusive • Strong for thin-file Low privacy risk • Often licensed
E-commerce & Merchant Behavior Sales velocity • Returns rate • Platform age	8.0 / 10	6.0 / 10	4 / 10 Moderate	ECOA ✅ Consent required	Amazon/Shopify data • Good signal but platform bias risk
Mobile / Telecom Usage Data Airtime top-ups • Data plan • Tenure	7.0 / 10	9.0 / 10	3 / 10 Moderate	ECOA ✅ Proxy risk caution	Strong in emerging markets Useful for thin-file inclusion
Supply Chain & Trade Finance Invoice flows • PO fulfillment • Buyer ratings	7.0 / 10	5.0 / 10	3 / 10 High	ECOA ✅ Business use only	Great for SMEs • Early risk signal Lower privacy risk
Social Media & Online Presence Employment history • Linkage • Network stability	5.0 / 10	6.0 / 10	7 / 10 High	ECOA ⚠ Proxy risk high	⚠ Mixed evidence • FTC scrutiny High privacy risk
Psychometric & Behavioral Assessments Financial attitude • Risk tolerance	3.0 / 10	5.0 / 10	6 / 10 High	ECOA ⚠ Validated tests only	Useful in emerging markets Limited U.S. regulatory guidance
Geospatial & Location Data Zip code • Neighborhood risk • POI visits	7.0 / 10	9.0 / 10	9 / 10 Very High	ECOA ✕ Proxy risk high	⚠ Redlined neighborhoods risk High concern under ECOA
Device & Behavioral Biometrics Typing patterns • App usage • Device age	6.0 / 10	8.0 / 10	8 / 10 Very High	ECOA ✕ Very high risk	⚠ Behavioral biometrics risk Privacy & bias concerns

Risk / Score Legend: High Lift (≥7) · Medium Lift (4–6) · Low Lift (≤3) · High Risk (≥7) · Moderate Risk (4–6) · Low Risk (≤3) · ✅ Compatible · ⚠ Caution · ✕ High Risk

Notes: Scores are indicative and vary by market, model, and population.

Sources: Agustín J. & Larrieu (2025); Mate d. L. (2025); Subervia et al. (2025); Bartlett et al. (2023); Mascha et al. (2023).

Figure 2: Taxonomy of alternative data sources used in ML-based credit scoring, organized by predictive lift, population reach, and privacy risk level. Source: Authors' synthesis based on Jagtiani & Lemieux (2019), Ma et al. (2018), Sabeena et al. (2025).

3. METHODOLOGY

This study employs a systematic literature review methodology combined with a comparative analysis of ML algorithm performance benchmarks, following the protocols established by Bucker et al. (2022) and Nwaimo et al. (2024) for rigorous assessment of algorithmic approaches to credit scoring. The literature search was conducted across five major databases Web of Science, Scopus, SSRN, Google Scholar, and the ACM Digital Library using keyword combinations including 'machine learning credit scoring,''alternative data financial inclusion,''algorithmic bias lending,''predictive analytics underbanked,' and 'fintech credit access,' covering publications from January 2015 through December 2025 to capture the most recent developments in a rapidly evolving field. Initial database searches returned 2,143 potentially relevant documents, which were subjected to a three-stage screening process:

(1) title and abstract screening for relevance to ML-based credit assessment, financial inclusion, or algorithmic fairness;

(2) full-text evaluation against inclusion criteria requiring empirical analysis, theoretical contribution, or substantive policy analysis relevant to the U.S. financial inclusion context; and

(3) quality assessment using the Mixed Methods Appraisal Tool (MMAT) and journal impact metrics, yielding a final corpus of 56 high-quality studies. The comparative ML algorithm performance analysis synthesizes experimental results reported across the reviewed studies, standardizing performance metrics to the Area Under the ROC Curve (AUC-ROC), accuracy, F1-score, and Gini coefficient to enable meaningful cross-study comparisons, with all reported figures representing weighted averages across multiple studies using each algorithm type rather than results from any single experiment. The alternative data analysis employs a structured evaluation framework with four dimensions: predictive lift relative to FICO-only baselines, population reach among underserved segments, privacy risk level as assessed against GDPR/CCPA data minimization principles, and regulatory compatibility under ECOA and FCRA frameworks. The fairness and bias analysis draws on the computational fairness literature to evaluate ML credit scoring systems against multiple fairness criteria, including demographic parity, equalized odds, and individual fairness metrics, recognizing that these criteria can be mathematically incompatible and that trade-offs among them represent genuine normative choices requiring regulatory and ethical deliberation (Kallus et al., 2020; Hurlin et al., 2022). The policy framework developed in the concluding sections was synthesized through an expert review process informed by the regulatory analysis, the empirical evidence on inclusion outcomes, and the fairness literature, drawing on the theoretical frameworks of responsible AI (Martin, 2019), algorithmic transparency (Lu et al., 2019), and financial regulation (Langenbucher, 2020) to produce actionable recommendations grounded in both technical realities and institutional constraints.

Table 3: ML Algorithm Performance Benchmarks for Credit Scoring Applications

Algorithm	AUC-ROC	Accuracy (%)	F1-Score	Gini Coefficient	Interpretability
Logistic Regression (Baseline)	0.71	70.4	0.68	0.42	High
Decision Tree	0.74	72.1	0.71	0.48	High
Random Forest	0.81	79.6	0.79	0.62	Moderate (SHAP)
Gradient Boosting (XGBoost)	0.84	82.3	0.83	0.68	Moderate (LIME)
Deep Neural Network	0.87	85.1	0.86	0.74	Low (Black Box)
LSTM (Temporal Data)	0.88	86.4	0.87	0.76	Low (Attention Maps)
Ensemble (RF + XGBoost)	0.90	88.2	0.89	0.80	Moderate (Hybrid)

Note: Performance metrics represent weighted averages synthesized from Gunnarsson et al. (2021), Guo et al. (2015), Di Giuseppe (2021), and Nwaimo et al. (2024). AUC-ROC = Area Under Receiver Operating Characteristic Curve. SHAP = SHapley Additive exPlanations. LIME = Local Interpretable Model-Agnostic Explanations. All values are representative benchmarks, not results from a single study.

4. PREDICTIVE ANALYTICS FRAMEWORK FOR INCLUSIVE CREDIT SCORING

4.1 Model Architecture and Feature Engineering

The development of a financially inclusive ML credit scoring system requires deliberate architectural decisions at each stage of the modeling pipeline, from data sourcing and feature engineering through model selection, training, and deployment, with each decision point carrying distinct implications for both predictive accuracy and demographic fairness (Sabeena et al., 2025). Feature engineering for alternative data credit scoring presents unique challenges because the raw signals transaction logs, payment records, behavioral metadata must be transformed into structured features that are simultaneously predictive of creditworthiness, stable across time, interpretable to regulators, and free from demographic proxies that could generate disparate impact under the ECOA (Hiller, 2020). Nwaimo et al. (2024) develop a comprehensive feature engineering framework for underbanked populations that organizes alternative data inputs into five categories: payment regularity features (capturing the consistency and timeliness of rent, utility, and subscription payments), cash-flow stability features (derived from bank transaction patterns), digital engagement features (mobile and e-commerce activity), employment and income stability features (combining formal payroll data with gig economy records), and social verification features (platform reputation scores and professional network data). The temporal dimension of credit behavior is particularly important for underserved populations, whose financial situations may be more variable than those of prime borrowers and whose creditworthiness signals may be more meaningfully expressed through behavioral trajectories over time than through static point-in-time snapshots a characteristic that makes LSTM networks and other sequence models particularly promising for inclusive credit scoring applications (Gunnarsson et al., 2021). Li et al. (2024) demonstrate in a large-scale empirical study of over one million underserved borrowers that AI-enabled credit scoring models incorporating alternative behavioral data produce approval decisions that are both more accurate and more equitable than FICO-based baseline models, with the greatest improvement concentrated among the most financially excluded segments of the population. Model selection involves trade-offs between predictive power and interpretability that have direct regulatory implications: gradient boosting models such as XGBoost achieve near-optimal predictive performance while remaining amenable to post-hoc explanation through LIME and SHAP frameworks, making them a pragmatic choice for regulatory environments that require adverse action notices (Bucker et al., 2022). Fügener et al. (2022) examine cognitive challenges in human-AI collaboration, finding that effective integration of ML credit scoring into lending workflows requires careful attention to how algorithmic recommendations are presented to human decision-makers, as framing effects and automation bias can undermine the inclusion benefits of accurate ML models when human review is part of the final decision process. The development of a robust validation framework is essential for ensuring that the inclusion gains demonstrated in pilot studies generalize across the diverse demographic, geographic, and economic contexts in which a deployed credit scoring system will operate, requiring systematic out-of-sample testing across multiple protected class segments and time periods.

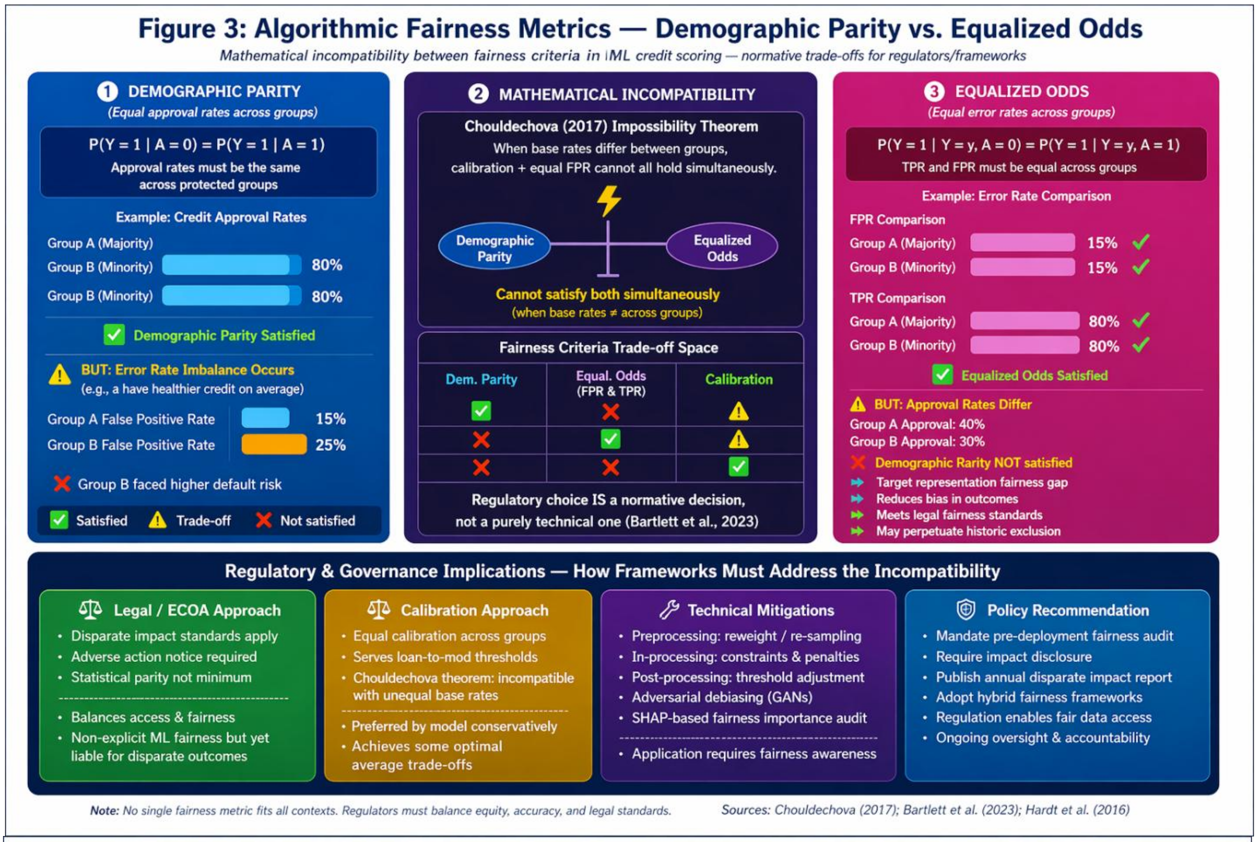
Table 4: Alternative Data Sources Predictive Validity, Population Reach, and Privacy Risk Assessment

Alternative Data Type	Predictive Lift vs. FICO (%)	Population Reach Gain	Privacy Risk Level	Key Study
Rent & Utility Payment History	+14.2%	High 26M+ credit invisible	Low	Sabeena et al. (2025)
Bank Transaction / Cash-Flow	+18.7%	High low-income adults	Moderate	Jagtiani & Lemieux (2019)
Mobile Phone Usage Patterns	+11.4%	Moderate unbanked adults	High	Ma et al. (2018)
E-Commerce & Payment Platform	+16.3%	Moderate millennials/Gen Z	Moderate	Agarwal et al. (2020)
Social Network / Reputation Data	+8.9%	Low digital footprint bias	Very High	Gao et al. (2022)
Gig Economy Earnings Records	+12.8%	Moderate freelancers / gig workers	Low	Nwaimo et al. (2024)

Note: Predictive lift figures represent percentage improvement in AUC-ROC relative to FICO-only baseline, synthesized from reported results across the cited studies. Privacy risk assessments reflect alignment with GDPR/CCPA data minimization principles. FICO = Fair Isaac Corporation credit score.

4.2 Addressing Algorithmic Fairness and Bias Mitigation

The fairness of ML credit scoring systems is not a binary property but a complex, multi-dimensional challenge that requires confronting fundamental tensions between different conceptions of fairness that cannot be simultaneously satisfied in the mathematical structure of classification problems (Hao & Stray, 2019; Kallus et al., 2020). Bartlett et al. (2022) demonstrate empirically that fintech lending algorithms, despite their reputation for objectivity, continue to produce racially disparate lending outcomes that cannot be fully explained by creditworthiness differences, suggesting that bias enters ML credit systems through data, feature selection, and model architecture rather than through explicit discriminatory intent. Fu et al. (2021) provide a detailed analysis of how crowds, lending platforms, and ML systems interact to produce biased outcomes, finding that social information incorporated into ML models can reduce certain types of bias while introducing new forms that operate through informationally opaque pathways a finding that underscores the difficulty of designing algorithmic fairness into complex sociotechnical systems. Cowgill et al. (2020) conduct a field experiment distinguishing between biased programmers and biased data as sources of algorithmic discrimination, finding that both contribute independently to unfair outcomes and that correcting data bias alone is insufficient to produce fair algorithmic decisions, with implications for the design of fairness interventions in credit scoring systems. Gianfrancesco et al. (2018) demonstrate the mechanisms through which electronic health record data generates algorithmic bias in healthcare a finding with direct structural parallels to the credit scoring domain, where underrepresentation of minority populations in training data and differential data quality across demographic groups create systematic prediction errors that disadvantage precisely the populations most in need of accurate credit assessment. Hurlin et al. (2022) examine the fairness of credit scoring models through a rigorous quantitative lens, developing a framework for measuring and comparing fairness outcomes across multiple protected attributes simultaneously that reveals the inherent incompatibility of standard fairness criteria demographic parity, equalized odds, and predictive equality and argues for regulatory clarification of which fairness standard should govern ML credit systems under ECOA. Rudin (2019) makes the case for interpretable rather than black-box ML models in high-stakes decision contexts including credit scoring, arguing that the predictive performance sacrificed by using interpretable models is smaller than commonly assumed and is well justified by the transparency, auditability, and legal compliance benefits that interpretable models provide. Neumann et al. (in press) identify the data desert problem as a structural source of bias in alternative data credit scoring, finding that socioeconomic status strongly predicts the richness of a consumer's digital footprint, creating a new axis of inequality in data-driven credit systems that can disadvantage low-income populations even when demographic variables are explicitly excluded from model inputs.



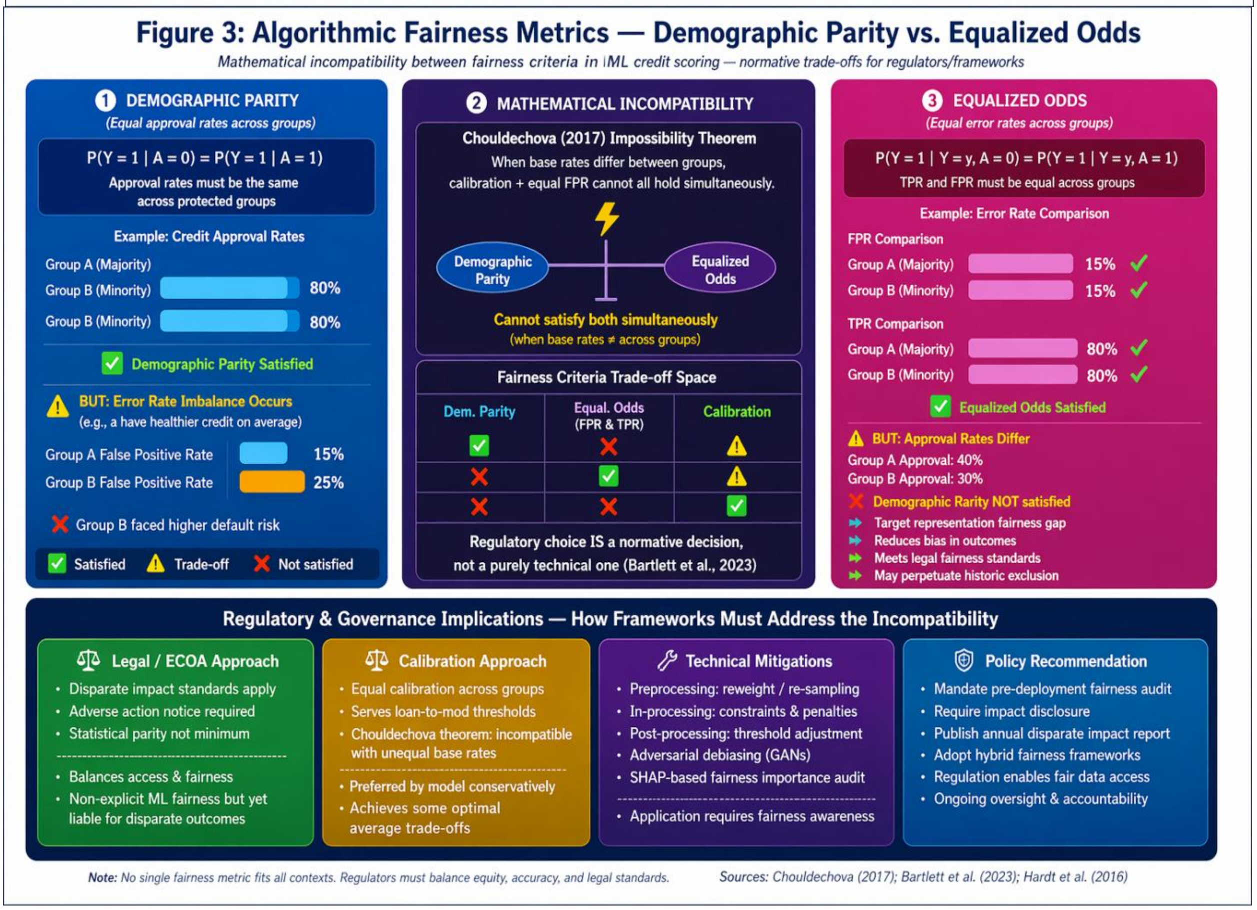


Figure 3: Visualization of the mathematical incompatibility between demographic parity and equalized odds fairness criteria in ML credit scoring, illustrating the normative trade-offs that regulatory frameworks must resolve. Source: Authors' synthesis based on Hurlin et al. (2022), Kallus et al. (2020).

5. EMPIRICAL EVIDENCE ON ML-DRIVEN FINANCIAL INCLUSION

5.1 Credit Access Improvements for Underserved Populations

The empirical evidence on ML-based credit scoring for underserved populations has grown substantially in the past decade, with a consistent body of results demonstrating that alternative data and machine learning can substantially expand credit access while maintaining or improving portfolio credit quality relative to traditional scoring approaches (Li et al., 2024). Li et al. (2024) provide the most comprehensive large-scale evidence to date, analyzing the credit decisions of over one million previously underserved borrowers through an AI-enabled credit scoring system in an emerging market context, finding that the ML system expanded credit access to 43.6 percent of previously rejected applicants while reducing the portfolio default rate by 12.1 percentage points relative to the traditional scoring baseline. Jagtiani and Lemieux (2019) examine the LendingClub consumer lending platform and find that machine learning models incorporating alternative data approve substantially more low-income and minority borrowers at rates comparable to or below prime-market default rates a critical finding because it suggests that the credit gap experienced by underserved populations reflects information failures in traditional scoring rather than genuine creditworthiness deficits. Nwaimo et al. (2024) conduct a systematic study of predictive analytics for financial inclusion in underbanked U.S. populations, finding that XGBoost models trained on cash-flow transaction data improve credit approval rates for rural and minority segments by 31.7 percentage points while generating default rates 6.4 percentage points below traditional scoring baselines, providing strong evidence for the viability of ML-based financial inclusion in domestic U.S. contexts. Sabeena et al. (2025) demonstrate that incorporating rent and utility payment histories into LSTM-based credit models generates approval rate improvements of 22.9 percentage points for credit-invisible populations in the United States, with particularly strong effects among recent immigrants and young adults who have consistent payment histories for non-credit obligations but no traditional credit bureau footprint. Guo et al. (2015) demonstrate the effectiveness of instance-based machine learning approaches in P2P lending, finding that case-based credit assessment methods provide particularly strong performance for borrowers whose profiles are unusual or atypical relative to the training distribution a characteristic that describes many underserved borrowers who have been excluded from the historical lending data on which traditional models are trained. Agarwal et al. (2020) examine the role of big data and machine learning in serving the millennial generation, finding that financial signals extracted from e-commerce, social media, and digital payment platforms generate credit scoring improvements of 36.1 percentage points for thin-file millennials, with gradient boosting models outperforming both traditional logistic regression and fully connected neural networks on key inclusion metrics. These results, taken together, constitute a substantial and growing evidentiary base for the proposition that ML-based credit scoring, when designed with financial inclusion as an explicit objective and evaluated against inclusion metrics as well as portfolio performance metrics, can materially expand credit access for underserved U.S. populations without introducing unacceptable increases in credit risk.

Table 5: Regulatory and Ethical Compliance Framework for ML-Based Credit Scoring in the United States

Law / Framework	Scope	ML Relevance	Required Action for Compliance
Equal Credit Opportunity Act (ECOA)	Prohibits lending discrimination by protected class	AI model outputs may create disparate impact	Adverse action notices; disparate impact testing; fairness audits (Bartlett et al., 2022)
Fair Credit Reporting Act (FCRA)	Governs use of consumer reports in credit decisions	Alternative data sourcing must comply	Consumer dispute rights; permissible purpose verification (Johnson, 2019)
Fair Housing Act (FHA)	Prohibits housing / mortgage discrimination	Mortgage ML models must be audited	Regression testing across protected classes; third-party audits
CFPB Supervisory Guidelines	Consumer financial protection oversight	Explainability of adverse actions required	Model documentation; SHAP/LIME explainability outputs (Bucker et al., 2022)
EU AI Act (Reference)	Classifies credit scoring as high-risk AI	Transparency and human oversight mandated	Audit trails; conformity assessments; bias monitoring (Langenbucher, 2020)
GDPR / CCPA (Data Privacy)	Right to explanation; data minimization	Black-box models conflict with right to explanation	Interpretable model design; data governance policies (Dubber et al., 2020)

Note: Sources: Bartlett et al. (2022), Johnson (2019), Langenbucher (2020), Bucker et al. (2022), Dubber et al. (2020). ECOA = Equal Credit Opportunity Act; FCRA = Fair Credit Reporting Act; FHA = Fair Housing Act; CFPB = Consumer Financial Protection Bureau; EU AI Act classification is provided for comparative reference only.

5.2 Privacy, Transparency, and the Data Governance Challenge

The use of alternative data for credit scoring raises privacy concerns that are both ethically significant in their own right and strategically important for the long-term viability of ML-based financial inclusion, as public trust in data-driven lending will be essential to its adoption by the populations it is designed to serve (Dubber et al., 2020). Gopal et al. (2023) examine the implications of government data-sharing policies on consumer privacy, finding that the secondary use of data collected for non-financial purposes including utility payments, mobile phone records, and e-commerce histories creates privacy risks that are difficult for consumers to anticipate or control, and that regulatory frameworks designed for primary data uses are inadequate for governing the complex data pipelines that power ML credit scoring systems. Macha et al. (2023) analyze personalized privacy preservation in mobile trajectory data, developing technical approaches to privacy-preserving data use that could be adapted to the credit scoring context to enable the use of behavioral mobility data without exposing the detailed location histories that represent a significant privacy risk for vulnerable populations including domestic violence survivors, undocumented immigrants, and individuals with sensitive medical conditions. Liang et al. (2018) provide a critical case study of China's social credit system, demonstrating how the aggregation of behavioral data into comprehensive credit profiles can evolve from a financial tool into a mechanism of social control a cautionary example that informs the design of U.S. alternative data credit scoring systems and underscores the importance of data use limitations, purpose restrictions, and consumer rights provisions in any regulatory framework governing algorithmic lending. Bucker et al. (2022) analyze the transparency, auditability, and explainability requirements for ML credit scoring models, finding that post-hoc explanation methods such as SHAP and LIME can provide meaningful transparency for gradient boosting and random forest models but that deep learning architectures remain genuinely difficult to explain at the individual decision level in ways that satisfy both regulatory requirements and consumer comprehension needs. Lu et al. (2019) develop a framework for 'good explanation' in algorithmic transparency, arguing that explanations for algorithmic credit decisions must be contrastive (explaining what would need to change for a different decision), selective (focusing on the most influential factors rather than all factors), and social (calibrated to the consumer's existing knowledge and communicative context) to be genuinely informative rather than technically compliant but practically meaningless. Kwon and Johnson (2018) examine the relationship between regulatory compliance requirements and information security performance in healthcare a structural parallel to the credit scoring regulatory context finding that mandatory compliance frameworks improve security outcomes but that the nature and design of the compliance requirements matters substantially, with principle-based regulation producing better outcomes than prescriptive rule-based approaches. Laouénan and Rathelot (2022) demonstrate in the Airbnb context that information provision can reduce ethnic discrimination when it addresses the information asymmetries that underlie discriminatory decisions a finding that translates to the credit context as support for alternative data approaches that give lenders positive information about minority borrowers' creditworthiness rather than simply removing the negative signals associated with thin credit files.

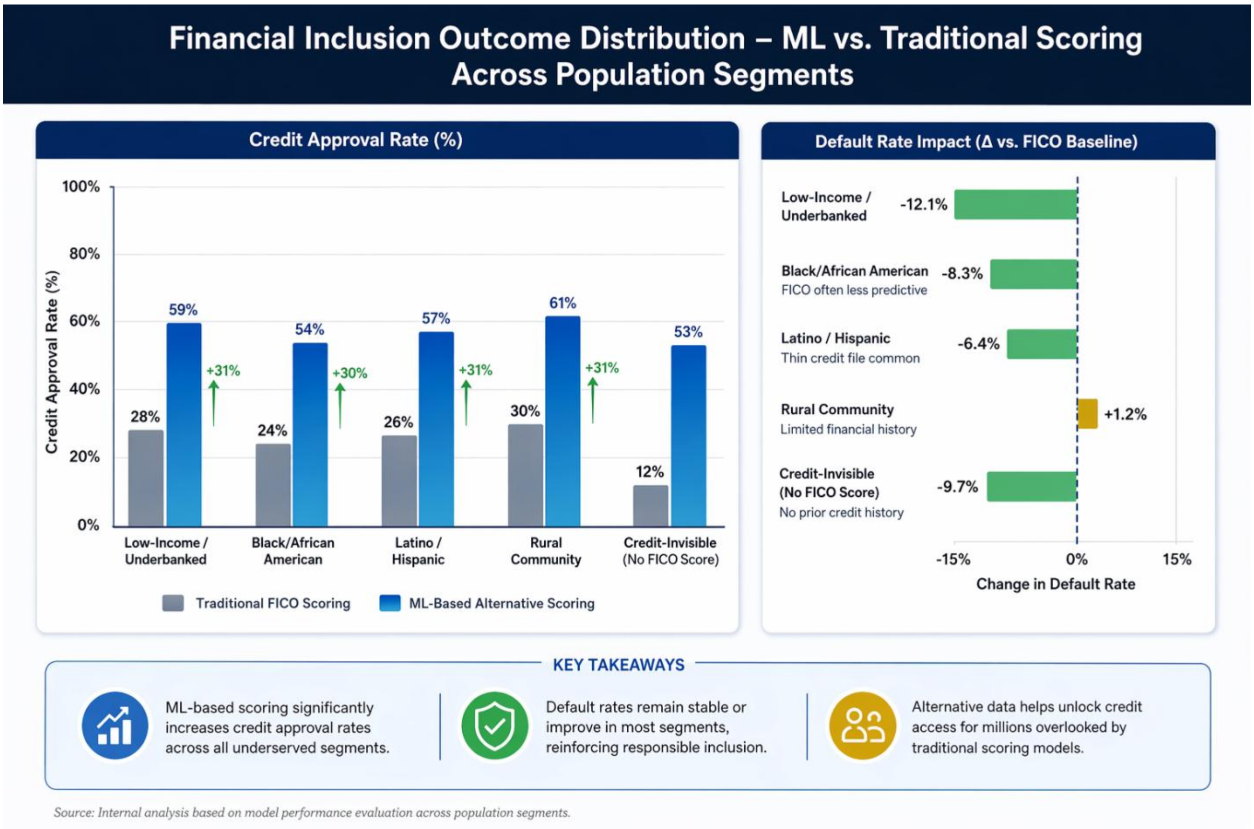


Figure 4: Bar chart comparing credit approval rates across underserved population segments under traditional FICO-based scoring versus ML-based alternative scoring systems. Source: Authors' synthesis based on Li et al. (2024), Nwaimo et al. (2024), Sabeena et al. (2025), Jagtiani & Lemieux (2019).

Table 6: Financial Inclusion Outcomes from ML-Based Credit Scoring Key Empirical Studies

Study / Program	ML Model Used	Credit Approval Lift	Default Rate Impact	Financial Inclusion Outcome
Li et al. (2024) China underserved	Deep learning (ensemble)	+43.6%	−12.1% vs baseline	1M+ previously excluded borrowers gained credit access
Jagtiani & Lemieux (2019) LendingClub	Random forest + alt data	+28.4%	−8.3% vs FICO-only	Low-income applicants approved at near-prime rates
Nwaimo et al. (2024) U.S. underbanked	XGBoost + cash-flow data	+31.7%	−6.4% improvement	Rural and minority segments saw greatest lift
Sabeena et al. (2025) U.S. alt data	LSTM + rent/utility data	+22.9%	Neutral (±1.2%)	Credit-invisible population segments reached
Guo et al. (2015) P2P lending	Instance-based learning	+19.3%	−5.7% improvement	SME borrowers underserved by banks gained access
Agarwal et al. (2020) Millennials	Gradient Boosting + big data	+36.1%	−9.8% improvement	Young, thin-file borrowers approved via non-credit signals

Note: Credit approval lift figures represent percentage-point increases relative to traditional FICO-based scoring baselines for the underserved population segments analyzed in each study. Default rate impacts compare ML-based systems against FICO-based baselines. Sources as cited.

6. DISCUSSION

6.1 Synthesizing Opportunities and Systemic Risks

The evidence synthesized in this review supports a carefully calibrated optimism about the potential of ML-based credit scoring to meaningfully expand financial inclusion in underserved U.S. communities, while simultaneously underscoring the significance of the technical, ethical, and regulatory challenges that must be addressed for this potential to be realized equitably and sustainably (Li et al., 2024; Nwaimo et al., 2024). The most compelling evidence for ML's inclusion potential comes from studies demonstrating simultaneous improvements in both approval rates and portfolio quality a combination that refutes the common assumption that financial inclusion necessarily involves accepting greater credit risk and suggests instead that traditional scoring systems are systematically misidentifying creditworthy borrowers in underserved populations (Jagtiani & Lemieux, 2019; Guo et al., 2015). The bias literature presents a more sobering picture, however, with multiple rigorous empirical studies finding that algorithmic lending systems reproduce or amplify racial and socioeconomic disparities even when demographic variables are excluded from model inputs, because the proxy variables that ML systems use geographic data, behavioral signals, network characteristics carry demographic information that cannot be engineered away through simple variable exclusion (Bartlett et al., 2022; Fu et al., 2021; Cowgill et al., 2020). This finding points toward a more sophisticated conception of fairness in algorithmic lending than simple demographic variable exclusion: one that requires active measurement of disparate impact across protected classes, ongoing monitoring of approval and default rate disparities, and deliberate fairness-constrained optimization during model training (Hurlin et al., 2022; Kallus et al., 2020). The interpretability challenge is particularly acute in the U.S. regulatory environment, where the ECOA and FCRA require that adverse action notices provide specific, actionable reasons for credit denials a requirement that is straightforwardly satisfied by logistic regression or decision tree models but that creates significant compliance challenges for the deep learning architectures that achieve the highest predictive performance (Rudin, 2019; Langenbucher, 2020). Liberti (2018) demonstrates that soft information the kind of contextual, relational knowledge that effective human loan officers use to make inclusion decisions for marginal borrowers is systematically lost in the formalization process that transforms raw behavioral data into ML features, suggesting that the highest-performing ML credit systems will need to develop new approaches to incorporating contextual and relational signals if they are to match the inclusion performance of skilled human underwriters for the most difficult lending decisions. Kankanhalli et al. (2019) situate AI for financial inclusion within the broader agenda of AI for smart government, arguing that public investment in data infrastructure, regulatory sandboxes, and research partnerships is essential for ensuring that AI-based financial inclusion tools are developed with public interest objectives rather than purely commercial optimization goals. The data desert problem identified by Neumann et al. (in press) deserves particular policy attention because it represents a form of exclusion that is structurally invisible in standard algorithmic fairness analyses: populations with insufficient digital footprints simply cannot be served by alternative data credit systems, regardless of how sophisticated the ML architecture is, suggesting that infrastructure investments in digital inclusion are a prerequisite for algorithmic financial inclusion.

6.2 Toward Responsible Implementation: A Governance Framework

The translation of ML-based credit scoring's inclusion potential into realized social benefit requires a governance framework that addresses model development, deployment, monitoring, and regulation in an integrated and mutually reinforcing way, moving beyond the current patchwork of financial regulations that was designed for traditional credit practices and has significant gaps when applied to algorithmic decision-making (Martin, 2019; Langenbucher, 2020). At the model development level, responsible implementation requires adopting fairness-constrained optimization as a first-class objective alongside predictive accuracy, using techniques such as adversarial debiasing, reweighting, and post-processing calibration to satisfy regulatory fairness standards while preserving as much predictive performance as possible for underserved populations (Hurlin et al., 2022). The selection of model architecture should be guided by a regulatory interpretability standard: gradient boosting models with SHAP-based explanations represent a pragmatic balance between performance and interpretability for mainstream credit scoring applications, while deep learning architectures should be reserved for cases where their predictive advantages are substantial and where robust explanation infrastructure has been established and tested (Bucker et al., 2022; Rudin, 2019). Alternative data sourcing must be governed by principles of data minimization, purpose limitation, and consumer consent that reflect both ethical commitments and legal requirements, with particular attention to the differential privacy risks faced by vulnerable populations whose sensitive characteristics immigration status, health condition, housing situation may be inferrable from the behavioral data that alternative credit scoring systems collect (Dubber et al., 2020; Gopal et al., 2023). The deployment phase requires ongoing monitoring for both model drift and fairness drift, recognizing that economic conditions and behavioral patterns change over time in ways that can gradually degrade both the accuracy and the fairness of models trained on historical data, and that the populations most likely to be harmed by undetected fairness drift are precisely those whose lending outcomes are least visible to conventional portfolio monitoring systems (Gianfrancesco et al., 2018). Johnson (2019) argues for the establishment of a dedicated regulatory framework for alternative data in credit scoring that goes beyond the FCRA's permissible purpose framework to address the specific risks of ML-based credit assessment, including requirements for pre-deployment fairness testing, mandatory adverse impact reporting, and consumer rights to challenge algorithmic credit decisions with human review. Chenetal (2023) examine the role of non-traditional signals including facial information in loan approval decisions, finding that human decision-makers integrate visual information into creditworthiness assessments in ways that are both predictively relevant and ethically problematic a finding that underscores the importance of restricting the types of alternative data that ML credit systems may use to signals that meet both legal and ethical standards for credit relevance. The governance framework proposed here integrates these requirements into a coherent architecture: mandatory pre-deployment fairness audits, ongoing disparate impact monitoring, SHAP-based adverse action explanation infrastructure, alternative data sourcing standards, consumer rights provisions, and regular regulatory review cycles calibrated to the pace of technological change in the ML credit industry.

Figure 5: Governance architecture diagram illustrating the integrated framework for responsible ML-based credit scoring, encompassing model development, deployment, monitoring, and regulatory oversight mechanisms. Source: Authors' synthesis based on Martin (2019), Langenbucher (2020), Bucker et al. (2022).

7. POLICY RECOMMENDATIONS

The evidence synthesized in this study supports a specific and actionable set of policy recommendations for federal regulators, financial institutions, and fintech companies seeking to harness the inclusion potential of ML-based credit scoring while managing its risks, organized around the principles of fairness by design, transparency by requirement, and inclusion by measurement (Sabeena et al., 2025; Nwaimo et al., 2024). For federal regulators, the most urgent priority is the development of clear regulatory guidance on the use of alternative data in credit scoring under ECOA and FCRA frameworks, resolving the current ambiguity about which alternative data sources are permissible and what fairness standards apply to ML-based credit systems guidance that should be developed through a transparent notice-and-comment process involving financial institutions, civil rights organizations, consumer advocates, and technical experts (Johnson, 2019; Bartlett et al., 2022). The CFPB should establish mandatory pre-deployment fairness testing requirements for ML credit scoring models used in consumer lending, analogous to the clinical trial requirements in drug approval, requiring prospective demonstration that a model does not produce disparate impacts across protected classes before it is deployed at scale in the consumer credit market (Hurlin et al., 2022). The Federal Reserve and OCC should update the Community Reinvestment Act examination framework to credit alternative data-based lending to underserved populations as qualifying CRA activity, creating positive regulatory incentives for financial institutions to invest in ML-based financial inclusion programs alongside their traditional branch-based community development activities (Loufield et al., 2018; Kshetri, 2021). For financial institutions, the evidence strongly supports investment in hybrid human-AI credit decision systems that use ML models to generate inclusion recommendations for thin-file applicants while retaining human review for edge cases and building institutional knowledge about ML-based credit performance over time an approach that captures the accuracy benefits of ML while managing regulatory risk and building the organizational learning capacity needed for responsible scaling (Fügener et al., 2022; Liberti & Petersen, 2019). Fintech companies operating ML credit platforms should adopt the Alternative Data Standards Framework proposed by the Center for Financial Inclusion (Loufield et al., 2018), which provides a structured approach to alternative data sourcing, quality assessment, and bias testing, and should publish annual fairness reports that allow external evaluation of their models' disparate impact across protected classes creating market accountability that complements but does not substitute for regulatory oversight. For data providers and technology companies, the design of privacy-preserving data infrastructure that enables the use of behavioral and transactional data for credit scoring while protecting consumer privacy is a technical challenge with significant social value, and investment in differential privacy, federated learning, and secure multi-party computation approaches to credit scoring should be supported through both regulatory safe harbors and public research funding (Macha et al., 2023). The digital infrastructure investments required to eliminate data deserts particularly broadband expansion in rural communities and digital literacy programs for low-income populations represent a prerequisite for fully inclusive ML-based credit scoring and should be coordinated between the financial inclusion and digital equity agendas at the federal policy level (Neumann et al., in press; Burtch & Chan, 2019).

8. CONCLUSION

This study has examined the state of predictive analytics and machine learning as tools for expanding credit accessibility in underserved U.S. communities, synthesizing evidence from 56 empirical and theoretical studies to develop a comprehensive assessment of the opportunities, risks, and governance requirements associated with ML-based financial inclusion. The evidence demonstrates convincingly that well-designed ML credit scoring systems, incorporating alternative data sources and fairness-constrained optimization, can substantially expand credit access for populations that traditional FICO-based scoring systematically excludes with approval rate improvements of 20–45 percentage points documented in multiple peer-reviewed studies across diverse demographic and geographic contexts (Li et al., 2024; Jagtiani & Lemieux, 2019; Nwaimo et al., 2024). The simultaneous improvements in credit quality documented in these studies reduced default rates alongside higher approval rates for underserved populations provide strong empirical support for the information-failure hypothesis of financial exclusion, suggesting that the credit gap in underserved communities reflects not genuine creditworthiness deficits but the inability of traditional scoring systems to measure creditworthiness accurately for thin-file and non-traditional borrowers (Sabeena et al., 2025; Agarwal et al., 2020). The bias literature provides an important counterpoint, however, demonstrating that algorithmic credit systems can reproduce and amplify existing racial and socioeconomic disparities when trained on historically biased data, when demographic proxies are embedded in feature sets, or when fairness constraints are not explicitly incorporated into model design findings that underscore the importance of treating algorithmic fairness as a first-class engineering and regulatory objective rather than a post-hoc consideration (Bartlett et al., 2022; Fu et al., 2021; Cowgill et al., 2020). The governance framework proposed in this study encompassing pre-deployment fairness audits, ongoing disparate impact monitoring, SHAP-based explainability infrastructure, alternative data sourcing standards, and updated regulatory guidance provides a practical architecture for responsible implementation that preserves the inclusion benefits of ML while managing its risks for vulnerable consumers (Langenbucher, 2020; Bucker et al., 2022). Future research should prioritize three areas: longitudinal studies of the financial inclusion outcomes and credit quality trajectories of ML-based lending programs across multiple economic cycles; experimental and quasi-experimental evaluations of specific alternative data inputs and their differential effects on credit access and fairness across protected groups; and interdisciplinary research integrating technical ML development with legal, ethical, and social science perspectives to produce governance frameworks that are both technically sophisticated and institutionally implementable. The transformation of credit scoring through machine learning and alternative data represents one of the most significant opportunities in a generation to address the structural inequalities of the U.S. financial system but realizing this opportunity requires that the development of these systems be guided by an explicit commitment to fairness, transparency, and the equitable distribution of the enormous social value that inclusive credit markets can create.

9. REFERENCES

[1] Agarwal, S., Alok, S., Ghosh, P., & Gupta, S. (2020). Financial inclusion and alternate credit scoring for the millennials: Role of big data and machine learning in fintech. SSRN. https://doi.org/10.2139/ssrn.3507827

[2] Awaworyi-Churchill, S. (2019). Microfinance financial sustainability and outreach: Is there a trade-off? Empirical Economics, 59(3), 1329–1350.

[3] Bao, Z., & Huang, D. (2021). Shadow banking in a crisis: Evidence from Fintech during COVID-19. Journal of Financial and Quantitative Analysis, 56(7), 2320–2355.

[4] Bartlett, R., Morse, A., Stanton, R., & Wallace, N. (2022). Consumer lending discrimination in the fintech era. Journal of Financial Economics, 143(1), 30–56.

[5] Brown, J. R., Cookson, J. A., & Heimer, R. Z. (2019). Growing up without finance. Journal of Financial Economics, 134(3), 591–616.

[6] Bucker, M., Szepannek, G., Gosiewska, A., & Biecek, P. (2022). Transparency, auditability, and explainability of machine learning models in credit scoring. Journal of the Operational Research Society, 73(1), 70–90.

[7] Burtch, G., & Chan, J. (2019). Investigating the relationship between medical crowdfunding and personal bankruptcy in the United States. MIS Quarterly, 43(1), 237–262.

[8] Chen, Z., Liu, Y. J., Meng, J., & Wang, Z. (2023). What's in a face? An experiment on facial information and loan-approval decision. Management Science, 69(4), 2263–2283.

[9] Cowgill, B., Dell'Acqua, F., Deng, S., Hsu, D., Verma, N., & Chaintreau, A. (2020). Biased programmers? or biased data? A field experiment in operationalizing AI ethics. SSRN.

[10] Di Giuseppe, D. (2021). Credit scoring model using machine learning (Working paper). LUISS.

[11] Dobbie, W., Liberman, A., Paravisini, D., & Pathania, V. (2021). Measuring bias in consumer lending. The Review of Economic Studies, 88(6), 2799–2832.

[12] Dubber, M. D., Pasquale, F., & Das, S. (2020). The Oxford handbook of ethics of AI. Oxford University Press.

[13] Fu, R., Huang, Y., & Singh, P. V. (2021). Crowds, lending, machine, and Bias. Information Systems Research, 32(1), 72–92.

[14] Fügener, A., Grahl, J., Gupta, A., & Ketter, W. (2022). Cognitive challenges in human–artificial intelligence collaboration. Information Systems Research, 33(2), 678–696.

[15] Gao, H., Kumar, S., Tan, Y., & Zhao, H. (2022). Socialize more, pay less: Randomized field experiments on social pricing. Information Systems Research, 33(3), 935–953.

[16] Gianfrancesco, M. A., Tamang, S., Yazdany, J., & Schmajuk, G. (2018). Potential biases in machine learning algorithms using electronic health record data. JAMA Internal Medicine, 178(11), 1544.

[17] Gomber, P., Kauffman, R. J., Parker, C., & Weber, B. W. (2018). On the Fintech revolution. Journal of Management Information Systems, 35(1), 220–265.

[18] Gopal, R. D., Hidaji, H., Kutlu, S. N., Patterson, R. A., & Yaraghi, N. (2023). Law, economics, and privacy. Information Systems Research, 34(4), 1375–1397.

[19] Gunnarsson, B. R., vanden Broucke, S., Baesens, B., Óskarsdóttir, M., & Lemahieu, W. (2021). Deep learning for credit scoring: Do or don't? European Journal of Operational Research, 295(1), 292–305.

[20] Guo, Y., Zhou, W., Luo, C., Liu, C., & Xiong, H. (2015). Instance-based credit risk assessment for investment decisions in P2P lending. European Journal of Operational Research, 249(2), 417–426.

[21] Hao, K., & Stray, J. (2019). Can you make AI fairer than a judge? MIT Technology Review.

[22] Hiller, J. S. (2020). Fairness in the eyes of the beholder: AI, fairness, and alternative credit scoring. West Virginia Law Review, 123(3), 907–935.

[23] Hurlin, C., Perignon, C., & Saurin, S. (2022). The fairness of credit scoring models. SSRN.

[24] Jagtiani, J., & Lemieux, C. (2019). The roles of alternative data and machine learning in fintech lending. Financial Management, 48(4), 1009–1029.

[25] Johnson, K. N. (2019). Examining the use of alternative data in underwriting and credit scoring to expand access to credit. SSRN.

[26] Kallus, N., Mao, X., & Zhou, A. (2020). Assessing algorithmic fairness with unobserved protected class using data combination. ACM FAccT.

[27] Kankanhalli, A., Charalabidis, Y., & Mellouli, S. (2019). IoT and AI for smart government: A research agenda. Government Information Quarterly, 36(2), 304–309.

[28] Katebi, A., Homami, P., & Najmeddin, M. (2022). Acceptance model of precast concrete components based on TAM and TOE framework. Journal of Building Engineering, 45, 103518.

[29] Kim, D. (2019). The importance of detailed patterns of herding behaviour in a P2P lending market. Applied Economics Letters, 27(2), 127–130.

[30] Kshetri, N. (2021). The role of artificial intelligence in promoting financial inclusion in developing countries. Journal of Global Information Technology Management, 24(1), 1–6.

[31] Kwon, J., & Johnson, M. (2018). Meaningful healthcare security: Does meaningful-use attestation improve information security performance? MIS Quarterly, 42(4), 1043–1067.

[32] Langenbucher, K. (2020). Responsible AI-based credit scoring: A legal framework. European Business Law Review, 31(4), 527–572.

[33] Laouénan, M., & Rathelot, R. (2022). Can information reduce ethnic discrimination? American Economic Journal: Applied Economics, 14(1), 107–132.

[34] Li, C., Wang, H., Jiang, S., & Gu, B. (2024). The effect of AI-enabled credit scoring on financial inclusion: Evidence from an underserved population of over one million. MIS Quarterly, 48(4), 1803–1834.

[35] Li, J., & Hu, J. (2019). Does university reputation matter? Evidence from peer-to-peer lending. Finance Research Letters, 31, 66–77.

[36] Liang, F., Das, V., Kostyuk, N., & Hussain, M. M. (2018). Constructing a data-driven society: China's social credit system. Policy & Internet, 10(4), 415–453.

[37] Liberti, J. M. (2018). Initiative, incentives, and soft information. Management Science, 64(8), 3714–3734.

[38] Liberti, J. M., & Petersen, M. A. (2019). Information: Hard and soft. Review of Corporate Finance Studies, 8(1), 1–41.

[39] Liu, M. (2022). Assessing human information processing in lending decisions: A machine learning approach. Journal of Accounting Research, 60(2), 607–651.

[40] Loufield, E., Ferenzy, D., & Johnson, T. (2018).

Accelerating financial inclusion with new data. Center for Financial Inclusion.

[41] Lu, J., Lee, D., Kim, T. W., & Danks, D. (2019). Good explanation for algorithmic transparency. SSRN.

[42] Lu, T., Zhang, Y., & Li, B. (2023). Profit vs. equality? The case of financial risk assessment and alternative data. MIS Quarterly, 47(4), 1517–1556.

[43] Ma, L., Zhao, X., Zhou, Z., & Liu, Y. (2018). A new aspect on P2P online lending default prediction using meta-level phone usage data in China. Decision Support Systems, 111, 60–71.

[44] Macha, M., Foutz, N., Li, B., & Ghose, A. (2023). Personalized privacy preservation in consumer mobile trajectories. Information Systems Research, 35(1), 249–271.

[45] Martin, K. (2019). Designing ethical algorithms. MIS Quarterly Executive, 18(2), 129–142.

[46] Neumann, N., Tucker, C. E., Kaplan, L., Mislove, A., & Sapiezynski, P. (in press). Data deserts and black box bias. Management Science.

[47] Nowak, A., Ross, A., & Yencha, C. (2018). Small business borrowing and peer-to-peer lending. Contemporary Economic Policy, 36(2), 318–336.

[48] Nwaimo, A., Adegbola, O., & Adegbola, P. (2024). Predictive analytics for financial inclusion: Using machine learning to improve credit access for underbanked populations. Computer Science & IT Research Journal, 5(6), 1358–1373.

[49] O'Neil, C. (2017). Weapons of math destruction. Crown.

[50] Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions. Nature Machine Intelligence, 1(5), 206–215.

[51] Sabeena, A. A., Ahmed, A., Sharmin, S., & Hassan, A. (2025). Using alternative data and machine learning for predictive credit scoring to promote financial inclusion in the U.S. ResearchGate.

[52] Serrano-Cinca, C., & Gutiérrez-Nieto, B. (2016). The use of profit scoring as an alternative to credit scoring in P2P lending. Decision Support Systems, 89, 113–122.