
Machine Learning-Based Financial Fraud Detection and Early-Risk Signaling in Community Lending Institutions

Abdalla Elmansoury ¹, Ibrahim Gabol ²

¹ University of the Witwatersrand, USA

² University of Alabama, USA

Keywords

Financial Fraud
Detection, Credit
Risk Scoring,
Early-Warning
Systems, Machine
Learning,
Explainable AI,
Fair Lending,
Community
Lending
Institutions,
Alternative Data

ABSTRACT

Community lending institutions — community banks, credit unions, community development financial institutions (CDFIs), and small-dollar or SME-focused lenders sit at the intersection of two loss-generating problems that the peer-reviewed literature has historically studied apart: application and identity fraud at origination, and early credit risk that is not recognized until a borrower is already delinquent. Both problems are structurally harder for a community lender than for a large bank or a well-capitalized fintech platform, because community lenders disproportionately serve thin-file, informationally opaque borrowers for whom a conventional bureau-based scorecard is least reliable, and because a small underwriting staff cannot absorb the same investigative volume a larger institution's fraud and credit teams can. Machine learning has been shown, across separate literatures in credit-card and peer-to-peer loan fraud detection, credit risk scoring, and alternative-data underwriting, to improve on both fronts, but that same literature documents two problems specific to deploying such models at a community lender: population drift degrades a scorecard's accuracy as the applicant pool and macroeconomic conditions shift, and a model's numerical score is not, by itself, a rationale that satisfies adverse-action and fair-lending obligations under equal-credit-opportunity law. This paper addresses the resulting gap by developing the Community Lending Early-Risk and Fraud Signaling framework (CLEFS): a reference architecture in which a shared machine learning engine scores both application-fraud risk and early default risk from the same enriched applicant data, an explainability and fair-lending review layer attaches an examiner- and regulator-legible rationale to every score, and a risk-tiered underwriting triage stage routes the resulting workload to a review depth a community lender's staff can sustain. Because no production deployment of this specific architecture exists to evaluate, the paper adopts a design-science approach: the framework is built from and evaluated against peer-reviewed findings in lending fraud detection, credit risk scoring and default prediction, alternative-data underwriting for peer-to-peer and SME lending, early-warning systems for borrower and institutional distress, explainability and fair-lending in credit and fraud machine learning, and community lending institution capacity constraints, and every quantitative comparison presented — two tables and four figures spanning capability coverage, a composite fraud-and-default miss rate over a twelve-month deployment, an application-volume-to-confirmed-yield trade-off, and underwriter review-hour reallocation is explicitly labeled as a literature-grounded estimate rather than a measured result. The paper's contribution is a synthesis of six otherwise separate literatures into a single origination-stage pipeline sized to a community lending institution's constraints, an explicit mapping of underwriting objectives to specific techniques and governance controls, and a candid account of where machine learning helps community lending institutions detect fraud and signal early risk, and where it introduces obligations only deliberate governance can meet.

Introduction

Community lending institutions face a dual exposure that the peer-reviewed literature has, until recently, tended to study as two separate problems. The first is fraud at origination: fabricated or stolen identities, coordinated application rings, and misrepresented income or collateral information that a lender discovers only after funds have already been disbursed. Xu *et al.*'s (2022) study of peer-to-peer loan fraud demonstrates that features constructed from transaction and application data can expose fraud patterns invisible to a single-application review, while Ngai *et al.*'s (2011) classification framework for financial fraud detection situates lending fraud within a broader family of techniques that data mining and machine learning have measurably improved. The second is credit risk that is not identified until well after origination: a borrower who was creditworthy on paper at approval but whose repayment capacity erodes in ways a static bureau score, computed once at application, cannot register. Altman's (1968) discriminant-analysis approach to predicting corporate bankruptcy remains the intellectual ancestor of this second problem, establishing more than half a century ago that a small number of financial ratios, tracked and scored systematically, can signal distress well before a formal default event. Both problems, treated separately, already carry an extensive literature; this paper's premise is that a community lending institution cannot afford to treat them separately in practice, because the same intake pipeline, the same underwriting staff, and the same technology budget must absorb both (Karim & Elias, 2022).

That premise is harder for a community lending institution to act on than for a large bank or a well-capitalized fintech lender, for reasons the community-lending literature documents directly. Berger and Udell (2002) show that small business lending at community institutions depends heavily on relationship-based, qualitative information precisely because the borrowers involved are informationally opaque to the standardized, hard-information underwriting a large bank's centralized scoring systems assume, which means the alternative and behavioral data that a fraud or early-risk model would need is often not systematically captured in the first place. Jagtiani and Lemieux (2018) find that fintech lenders have expanded credit access specifically into areas underserved by traditional banks, evidence that alternative-data-driven underwriting can reach exactly the borrower population community lenders exist to serve, but also a signal that community institutions risk losing origination volume to better-instrumented competitors if they cannot deploy comparable analytics themselves. Benjamin, Rubin, and Zielenbach (2004) document that community development financial institutions operate under a mission-driven mandate to serve borrowers conventional lenders decline, a mandate that makes both types of loss — fraud that drains scarce lending capital, and undetected early default that erodes a CDFI's ability to keep lending — disproportionately consequential relative to a CDFI's typically thin capital base and small staff (Karim & Seitz, 2023).

Machine learning has been advanced as a remedy for both problems, and a substantial peer-reviewed literature supports the claim that it can help with each. On the credit-scoring side, Lessmann *et al.*'s (2015) large-scale benchmarking study finds that modern classification algorithms consistently outperform traditional scorecards on default prediction, building on Baesens *et al.*'s (2003) earlier benchmarking work, and Xia *et al.* (2017) show that boosted

tree ensembles with tuned hyperparameters can push predictive accuracy further still. On the alternative-data side, Serrano-Cinca, Gutiérrez-Nieto, and López-Palacios (2015) identify the specific determinants of default in peer-to-peer lending data, Malekipirbazari and Aksakalli (2015) show that random forests applied to social lending platform data materially improve risk assessment over the platforms' own published grades, and Jiang *et al.* (2018) demonstrate that even unstructured, soft information extracted from a loan application's descriptive text carries predictive signal a purely numeric scorecard would miss — precisely the kind of signal a thin-file or informationally opaque community-lending applicant is most likely to generate in place of a long bureau history. On the fraud side, the gains Xu *et al.* (2022) and Ngai *et al.* (2011) document extend naturally to a community lender's origination pipeline. Two problems specific to deployment temper this promise, however, and both are established in literatures this paper also draws on: first, any scorecard's accuracy degrades as the applicant population and macroeconomic conditions drift away from the population it was trained on, a phenomenon Hand (2006) terms population drift and Sousa, Gama, and Brandão (2016) show can be addressed through dynamic, continuously updated credit-risk models rather than a fixed model trained once; second, a model's numerical score is not, by itself, a rationale that satisfies the adverse-action notice and disparate-impact obligations that apply to consumer and small business lending, a problem Kozodoi, Jacob, and Lessmann (2022) and Fuster *et al.* (2022) both show requires deliberate fairness assessment and explainability engineering rather than an assumption that an accurate model is automatically a fair or explainable one (Rahman *et al.*, 2022).

No existing framework combines these threads — shared fraud and early-risk scoring from the same enriched applicant data, population-drift monitoring, and fair-lending-grade explainability — into a single architecture explicitly sized to a community lending institution's staffing, data, and mission constraints rather than a large bank's or a venture-funded fintech's. The lending-fraud and credit-scoring literatures largely proceed on separate tracks despite drawing on overlapping application and behavioral data; the alternative-data underwriting literature demonstrates gains for thin-file borrowers without addressing how a resource-constrained lender should govern the resulting models; and the explainability and fair-lending literature is developed mostly against general credit-scoring settings rather than the specific mission and capacity constraints Berger and Udell (2002), Jagtiani and Lemieux (2018), and Benjamin *et al.* (2004) establish for community lending institutions. This paper addresses that gap. It develops the Community Lending Early-Risk and Fraud Signaling framework (CLEFS): an architecture in which a shared scoring engine produces both an application-fraud score and an early default-risk score from one enriched applicant record, a continuous population-drift monitor keeps that engine current, an explainability and fair-lending review layer attaches an examiner-legible, adverse-action-ready rationale to every score, and a risk-tiered underwriting triage stage routes the resulting workload to a review depth a community lender's staff can sustain (Rahman *et al.*, 2023).

2. Literature Review

2.1 Fraud Detection in Lending and Adjacent Financial Contexts

The peer-reviewed literature on financial fraud detection provides the technical foundation for detecting fraud at loan origination, even though relatively little of it is written specifically about lending. Ngai *et al.*'s (2011) classification framework surveys data mining techniques applied across financial fraud types and finds consistent evidence that supervised and unsupervised methods alike can improve on manual review, while cautioning that performance depends heavily on how well the features available at detection time capture the fraud pattern in question. Bhattacharyya *et al.*'s (2011) comparative study of credit card fraud detection methods reaches a compatible conclusion from a narrower angle, finding that ensemble and support vector machine approaches outperform simpler classifiers on a structurally similar problem: scoring a high volume of applications or transactions for a comparatively rare fraudulent event. Xu *et al.*'s (2022) study, the most directly relevant of the three, applies feature construction from transaction-level data specifically to peer-to-peer loan fraud detection and finds that features capturing relational and behavioral patterns across an applicant's activity substantially outperform features drawn from a single application in isolation, evidence that lending fraud, like the credit-card fraud Bhattacharyya *et al.* (2011) study, is often more a network and behavioral phenomenon than a property of any one application viewed alone. Read together, this literature establishes both that fraud detection techniques transfer reasonably well from adjacent financial domains into lending and that the strongest results come from models that see beyond a single transaction or application to the relational context surrounding it — a design implication this paper's framework adopts directly (Rahman & Tirupathi, 2023).

2.2 Credit Risk Scoring, Default Prediction, and Model Drift

A substantial and long-running peer-reviewed literature benchmarks machine learning approaches to credit risk scoring and default prediction, generally finding that more flexible models outperform traditional scorecards, while a smaller but critical literature establishes that any such model's accuracy is contingent on the population it was trained on remaining stable. Baesens *et al.*'s (2003) early benchmarking study compares classification algorithms for credit scoring and finds meaningful accuracy differences across methods even at that early stage of the field, and Lessmann *et al.*'s (2015) large-scale update, run more than a decade later, confirms that the ranking of algorithms has shifted with methodological advances but that the central finding — flexible, non-linear classifiers outperform simple scorecards — has only strengthened. Xia *et al.* (2017) extend this line further, showing that a boosted decision tree approach with hyperparameters tuned through Bayesian optimization achieves further accuracy gains over the ensemble methods Lessmann *et al.* (2015) benchmark, evidence that continued methodological refinement keeps producing real, measurable improvement rather than approaching a ceiling. This entire literature, however, implicitly assumes a roughly stable relationship between applicant characteristics and default outcomes, an assumption Hand (2006) directly challenges: his analysis of what he terms the illusion of progress in classifier technology shows that a model's apparent accuracy gain, measured against a fixed historical dataset, can substantially overstate its real-world performance once the applicant population served begins to drift, because the population a model is scored against in a benchmark rarely matches the population it will actually see in deployment years later. Sousa, Gama, and Brandão (2016) offer a constructive response to

exactly this problem, proposing a dynamic credit-risk modeling framework that continuously updates as new outcome data arrives rather than relying on a single static model fit once at deployment, and showing that such a dynamic approach maintains accuracy over time in a way a static model cannot (Siddiqui *et al.*, 2023). Read together, this literature implies that a community lender adopting a machine learning scorecard faces not a one-time modeling decision but an ongoing maintenance obligation, and that the benchmarking gains Baesens *et al.* (2003), Lessmann *et al.* (2015), and Xia *et al.* (2017) document will erode on exactly the timescale Hand (2006) describes unless a mechanism like the one Sousa *et al.* (2016) propose is built in from the start.

2.3 Alternative Data and Early-Risk Signals in Peer-to-Peer and SME Lending

A closely related literature, developed largely around peer-to-peer and small-business lending platforms, shows that alternative and unstructured data sources can meaningfully improve early-risk assessment for exactly the thin-file, informationally opaque borrowers community lending institutions disproportionately serve. Serrano-Cinca, Gutiérrez-Nieto, and López-Palacios (2015) analyze the specific determinants of default in peer-to-peer lending data and find that variables beyond a conventional bureau score — loan purpose, borrower-reported financial details, and platform-specific behavioral indicators — carry independent predictive value for default risk. Malekipirbazari and Aksakalli (2015) build on this by applying random forests to social lending platform data and finding that a machine learning model trained on the platform’s own applicant and repayment data outperforms the platform’s own published borrower grade, direct evidence that a lender’s proprietary data, properly modeled, can outperform an externally purchased score. Jiang *et al.* (2018) push the alternative-data argument furthest, showing that soft information extracted from the free-text descriptions borrowers write when applying for a peer-to-peer loan adds predictive power beyond what structured application fields alone provide, a finding with particular relevance for community lending institutions whose applicants frequently lack the extensive structured credit history a conventional scorecard depends on (Potharaju & Addanki, 2020). This literature’s collective implication is that a community lender is not disadvantaged in early-risk signaling simply because its applicants lack thick bureau files; the alternative and unstructured data such applicants do generate, if captured and modeled properly, can substitute for at least part of what a thin bureau file cannot provide.

2.4 Early-Warning Systems for Borrower and Institutional Distress

A distinct and older literature on early-warning systems establishes that distress — whether a borrower’s or an institution’s — reliably produces measurable precursor signals well before a formal default or failure event, provided those signals are tracked systematically. Altman’s (1968) discriminant-analysis model, built from a small set of financial ratios, remains the foundational demonstration that firm-level financial distress can be predicted with meaningful lead time rather than discovered only at the point of bankruptcy, and its core logic — that a handful of well-chosen ratios, monitored continuously, outperforms waiting for an overt default event — underlies essentially every early-warning model developed since. Cole and Gunther (1998) apply the same early-warning logic to bank failure prediction, comparing on-site examination data against off-site financial-statement

monitoring and finding that off-site, continuously available financial data can predict failure with useful lead time even between formal examinations, evidence that early-warning signaling does not require the intensity of a full on-site review to be effective. Cole and White (2012), writing after the 2007–2009 financial crisis, revisit the drivers of U.S. commercial bank failure and find that the same categories of financial-ratio deterioration Altman (1968) and Cole and Gunther (1998) identify in earlier eras recur in a modern failure wave, evidence that the underlying early-warning logic generalizes across time periods even as the specific institutions and products involved change. For a community lending institution, this literature’s implication is direct: an early-risk signal does not need to wait for a borrower’s payment to become delinquent, any more than a bank examiner needs to wait for a formal failure; the same class of financial-ratio and behavioral deterioration these studies track at the institutional level can, in principle, be tracked at the individual borrower level as part of an origination and portfolio-monitoring pipeline.

2.5 Explainability and Fair Lending in Credit and Fraud Machine Learning

A model’s numerical risk score is not, by itself, the kind of rationale that consumer and small business lending law requires, and the explainability and fair-lending literature treats this as a problem that must be engineered rather than one that resolves itself as model accuracy improves. Kozodoi, Jacob, and Lessmann (2022) show that fairness in credit scoring must be explicitly assessed and, where necessary, actively corrected for, because an accurate model optimized purely for predictive performance can nonetheless produce disparate outcomes across protected groups, and they demonstrate that fairness-aware model adjustments can reduce this disparity with only a modest cost to overall profitability. Fuster *et al.* (2022) reach a related but distinct conclusion at a larger scale, finding that machine learning’s effect on credit market outcomes is not uniformly positive across all borrower groups relative to traditional scoring methods, and that the distributional consequences of adopting a more accurate but less transparent model require deliberate evaluation rather than an assumption that improved overall accuracy benefits all applicant groups equally. Gramespacher and Posth (2021) address the explainability side directly, showing that explainable AI techniques can be used not merely to justify a model’s decisions after the fact but to actively optimize a loan portfolio’s return target function in a way a human reviewer can interpret and act on, evidence that explainability and effective decision-making are complementary rather than competing objectives. None of the three treats fairness or explainability as an automatic byproduct of a well-performing model; each shows that a dedicated evaluation or engineering step is required, and an architecture that omits that step risks producing a score that is statistically accurate but neither legally defensible nor equitable across the applicant population a community lender’s mission commits it to serve (Potharaju & Sajja, 2021).

2.6 Community Lending Institutions: Capacity Constraints and the Synthesis Gap

The community-lending literature establishes that the fraud and early-risk problems Sections 2.1 through 2.4 describe are not merely present at community institutions but disproportionately consequential for them, given how these institutions are structured and whom they serve. Berger and Udell (2002) show that small business lending at community

institutions relies heavily on relationship-based, qualitative underwriting precisely because the borrower population involved does not generate the standardized hard information a large bank's centralized systems assume, meaning any machine learning model a community lender adopts must be built around data these institutions do not uniformly or systematically collect today. Jagtiani and Lemieux (2018) find that fintech lenders using alternative data and automated underwriting have successfully expanded credit access into markets traditional banks underserve, a finding that cuts two ways for a community lending institution: it confirms alternative-data-driven analytics can reach exactly the borrowers community lenders serve, while also signaling competitive pressure from better-instrumented entrants a resource-constrained community lender may struggle to match without comparable tools of its own. Benjamin, Rubin, and Zielenbach (2004) describe the mission-driven structure of community development financial institutions directly, documenting that these institutions serve borrowers conventional lenders decline as a matter of organizational purpose rather than incidental market positioning, which means both an undetected fraud loss and an undetected early default carry outsized consequences for an institution whose lending capital is, by mission and by charter, already scarcer than a conventional lender's (Potharaju & Addanki, 2022).

Read together, these six literatures describe two problems and several partial solutions that no existing framework combines into one architecture sized to the institutions that need it most. The lending-fraud literature shows that relational, behavior-aware features detect fraud a single-application review misses; the credit-scoring literature shows that flexible machine learning models outperform static scorecards but degrade under population drift unless deliberately maintained; the alternative-data literature shows that thin-file and informationally opaque borrowers can still be scored effectively using unconventional data sources; the early-warning literature establishes that both institutional and borrower-level distress produce trackable precursor signals well before a formal default event; the explainability and fair-lending literature shows that a model's score must be actively engineered into a legally defensible, equitable rationale rather than assumed to be one; and the community-lending literature establishes that the institutions most exposed to both fraud and undetected early risk are also the ones least equipped, in staffing and in systematically captured data, to address either problem with conventional means. No paper in any of these literatures proposes an architecture that scores fraud and early default risk together from one enriched applicant record (Potharaju & Ravipudi, 2023), monitors and corrects for population drift, and builds in fair-lending-grade explainability, all sized explicitly to a community lending institution's capacity. This paper's contribution is that synthesis: the Community Lending Early-Risk and Fraud Signaling framework, developed specifically to bring the detection and prediction gains Sections 2.1 through 2.4 document to community lending institutions without asking them to either accept an unexplainable, potentially inequitable score or staff an analytics function they cannot afford.

3. Methodology

3.1 Research Approach: Design Science with Evaluation

This paper follows a design-science research approach: it develops an artifact — the Community Lending Early-Risk and Fraud Signaling framework — intended to address a practical problem (detecting application fraud and signaling early default risk at community lending institutions without sacrificing explainability, fair-lending defensibility, or feasibility for a limited underwriting staff), and it evaluates that artifact against the cumulative findings of the peer-reviewed literature rather than against a production deployment, because no such deployment of this specific architecture yet exists to study. Every quantitative comparison in Section 4 is accordingly an construction, explicitly labeled as such, built to be directionally consistent with the cited literature’s findings rather than presented as a measured outcome.

3.2 Development of the Framework

The framework was developed by identifying, for each stage of a conventional loan origination and portfolio-monitoring workflow, the specific failure mode the fraud-detection, credit-scoring, alternative-data, early-warning, and explainability literatures each separately document, and then determining what technique and governance control that literature implies is necessary to address it without exceeding a community lending institution’s staffing capacity. Fraud detection that sees beyond a single application — the gain Xu *et al.* (2022) and Bhattacharyya *et al.* (2011) document — is addressed through relational and behavioral feature construction applied before any score is computed, following Ngai *et al.*’s (2011) broader classification framework. Early-risk signaling for thin-file, informationally opaque applicants — the gain Serrano-Cinca *et al.* (2015), Malekipirbazari and Aksakalli (2015), and Jiang *et al.* (2018) document — is addressed by enriching the applicant record with alternative and unstructured data alongside conventional bureau fields, so that the same enriched record feeds both the fraud and default-risk scoring tasks. Model degradation under population drift — the risk Hand (2006) establishes and Sousa *et al.* (2016) show can be corrected through continuous updating — is addressed through a dedicated drift monitor that triggers scheduled recalibration rather than relying on a fixed, one-time model fit. Finally, the requirement that every score be both explainable and fair-lending defensible — the problem Kozodoi *et al.* (2022), Fuster *et al.* (2022), and Gramespacher and Posth (2021) address technically — is met by a dedicated explainability and fair-lending review layer positioned between the scoring engine and the underwriter, so that no risk score reaches a human reviewer, an applicant-facing adverse-action notice, or a disparate-impact audit without an attached, feature-level rationale.

3.3 Construction of Evaluation Data

Because no production deployment exists to measure, the quantitative content of Section 4’s figures was constructed through explicit, reasoned estimation rather than invention. For each figure, construction began from a directional finding reported in one or more cited studies — Lessmann *et al.*’s (2015) benchmarked accuracy gains for flexible classifiers over traditional scorecards, or Hand’s (2006) account of population-drift-driven performance decay — and built numerical values directionally consistent with that finding at a plausible order of magnitude, without claiming a precision an exercise cannot support. Values were kept within ranges commonly discussed across the cited credit-scoring, lending-fraud, and

early-warning literature rather than chosen to flatter the proposed framework. Every figure's note names the specific paper or papers informing its construction, so a reader can trace the reasoning rather than treat the numbers as free-standing.

3.4 Case Scenario Approach

Section 4.4 presents a single case scenario — a hypothetical community lending institution adopting the proposed framework — constructed to make the abstract architecture concrete rather than to report on an actual deployment. The scenario combines publicly discussed, general characteristics of community lending institution underwriting operations with the technical mechanisms synthesized from Section 2, to produce a plausible walk-through of how the framework would process a specific application from intake through underwriter disposition. No identifying details of any real institution, applicant, or case are used; the scenario is explicitly hypothetical and offered as a reasoning aid.

3.5 Assumptions and Scope Limitations

Several assumptions and scope limitations bound this study. First, the paper assumes the adopting institution has at least a baseline loan origination system and some history of application and repayment outcome data already in place; it does not address a from-scratch technology implementation for an institution with no existing underwriting infrastructure. Second, the paper assumes access to a sufficient volume of historical, labeled fraud and default outcome data to train an initial model, acknowledging that Berger and Udell's (2002) findings on the qualitative, relationship-based nature of community lending underwriting imply this labeled history may be thinner or less systematically structured at a community institution than at a large bank, and returning to this constraint directly in Section 5.2. Third, the paper is scoped to origination-stage fraud detection and early-risk signaling; it does not address the separate functions of loan servicing, collections, or portfolio-level stress testing, each of which the cited literature treats as related but analytically distinct. Fourth, as stated repeatedly above, all quantitative results in Section 4 are estimates grounded in the literature rather than measurements, and conclusions about the framework's plausibility should be read with that limitation firmly in mind.

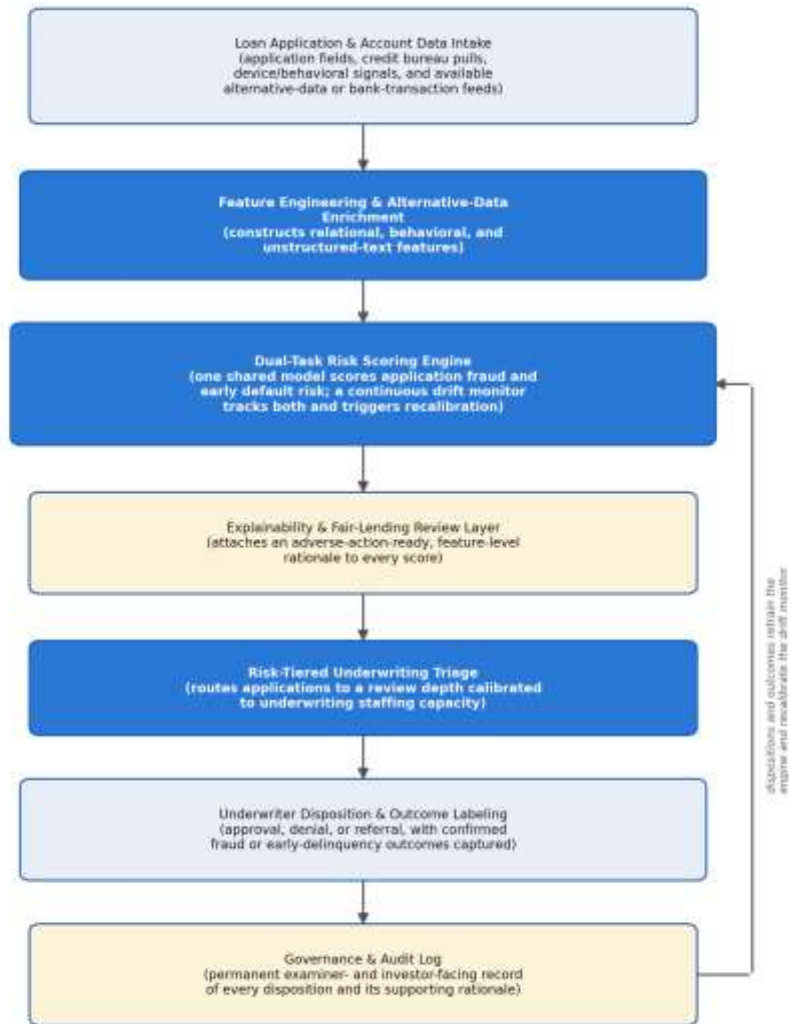
4. Results

4.1 Architecture Overview: The Community Lending Early-Risk and Fraud Signaling Framework

The architecture synthesized from the literature review organizes the Community Lending Early-Risk and Fraud Signaling framework as a pipeline in which every stage responds directly to a specific failure mode identified in Section 2. Loan application and account data — application form fields, credit bureau pulls, device and behavioral signals, and, where available, bank-transaction or alternative-data feeds — first pass through feature engineering and alternative-data enrichment, which constructs the relational, behavioral, and unstructured-text features that Xu *et al.* (2022), Malekipirbazari and Aksakalli (2015), and Jiang *et al.* (2018) each show carry predictive signal beyond a conventional bureau field. A single dual-task risk scoring engine then produces two outputs from that same enriched

record an application-fraud score, built around the relational and behavioral fraud-detection approach Ngai *et al.* (2011) and Bhattacharyya *et al.* (2011) establish, and an early default-risk score, built around the flexible classification methods Lessmann *et al.* (2015) and Xia *et al.* (2017) benchmark — while a continuous population-drift monitor, operationalizing Hand's (2006) diagnosis and Sousa *et al.*'s (2016) dynamic-updating remedy, tracks both scores' accuracy against incoming disposition outcomes and triggers scheduled recalibration before drift-driven degradation becomes visible in underwriting decisions. Every score the engine produces passes through an explainability and fair-lending review layer before reaching a human reviewer, attaching the feature-level, adverse-action-ready rationale Kozodoi *et al.* (2022), Fuster *et al.* (2022), and Gramespacher and Posth (2021) show is necessary for a score to function as both legally defensible and internally interpretable evidence rather than an opaque number. A risk-tiered underwriting triage stage then routes the resulting workload to a review depth calibrated to the institution's actual underwriting staffing, before an underwriter renders the disposition — approval, denial, or referral for manual investigation — that a governance and audit log records permanently. That log serves two purposes at once: it is the documentation a fair-lending examination or investor audit requires, and it is the labeled feedback the dual-task scoring engine and drift monitor need to keep improving, closing the loop back to the scoring engine rather than leaving each disposition as a dead end.

Figure 4.1. The Community Lending Early-Risk and Fraud Signaling (CLEFS) Pipeline



The architecture deliberately positions the explainability and fair-lending review layer between the scoring engine and the underwriter, rather than treating explanation as something generated after the fact only if an applicant disputes a decision, so that no risk score ever reaches a disposition without its rationale already attached. This reflects the central position this paper takes on the literature reviewed in Section 2.5: explainability and fair-lending review are required pipeline stages, not optional add-ons invoked under regulatory or investor pressure.

4.2 Table 1: Manual Underwriting, a Static Machine Learning Scorecard, and CLEFS Compared

Dimension	Manual Underwriting (Status Quo)	Static Scorecard (No Adaptation)	ML (No Adaptation)	CLEFS
Application-fraud detection	Weak beyond obvious inconsistencies; relies on reviewer experience	Moderate; catches patterns present in training data		Strong; relational and behavioral features follow Xu <i>et al.</i> (2022) and Bhattacharyya <i>et al.</i> (2011)
Early default-risk signaling lead time	Minimal; risk typically surfaces only at delinquency	Moderate; a bureau-based scorecard signals some risk at origination		Strong; alternative-data enrichment follows Serrano-Cinca <i>et al.</i> (2015) and Jiang <i>et al.</i> (2018)
Explainability / adverse-action defensibility	Strong; a human reviewer's reasoning is inherently articulable	Weak unless a separate explanation method is added		Strong; a dedicated review layer follows Kozodoi <i>et al.</i> (2022) and Gramespacher and Posth (2021)
Adaptability to population drift	Implicit; a human reviewer adjusts informally over time	Weak; degrades measurably, consistent with Hand (2006)		Strong; continuous drift monitoring follows Sousa <i>et al.</i> (2016)
Coverage of thin-file / alternative-data applicants	Moderate; depends on reviewer's qualitative judgment, per Berger and Udell (2002)	Weak; a bureau-only scorecard underperforms for thin-file applicants		Strong; enrichment specifically targets this population
Fair-lending / disparate-impact monitoring	Inconsistent; depends on reviewer training and institutional policy	Weak unless a separate fairness audit is added		Strong; fairness assessment follows Kozodoi <i>et al.</i> (2022) and Fuster <i>et al.</i> (2022)
Feasibility for a limited underwriting staff	Moderate; well understood but does not scale with application volume	Lower initially; requires data science capacity most community institutions lack in-house		Moderate to strong; higher initial setup cost offset by sustained triage efficiency

4.3 Table 2: Mapping Origination Objectives to CLEFS Techniques and Governance Controls

Objective	Representative Technique	Corresponding Governance Control	Grounding Literature
Detect application and identity fraud beyond a single application	Relational and behavioral feature construction	Escalation protocol for network-flagged clusters	Xu <i>et al.</i> (2022); Ngai <i>et al.</i> (2011)
Improve fraud detection accuracy over baseline classifiers	Ensemble and support vector machine methods	Periodic model benchmarking against new fraud typologies	Bhattacharyya <i>et al.</i> (2011)
Signal early default risk for thin-file applicants	Alternative-data and unstructured-text enrichment	Data governance review of alternative-data sources	Serrano-Cinca <i>et al.</i> (2015); Malekipirbazari and Aksakalli (2015); Jiang <i>et al.</i> (2018)
Maintain scoring accuracy as the applicant population evolves	Continuous population-drift monitoring	Scheduled model recalibration triggered by drift signal	Hand (2006); Sousa <i>et al.</i> (2016)
Track institution- and borrower-level distress precursors	Financial-ratio and behavioral early-warning indicators	Portfolio-monitoring review calibrated to indicator thresholds	Altman (1968); Cole and Gunther (1998); Cole and White (2012)
Provide adverse-action-ready rationale for every disposition	Feature-attribution explainability layer	Mandatory rationale log attached to every score	Kozodoi <i>et al.</i> (2022); Gramespacher and Posth (2021)
Monitor for disparate impact across protected applicant groups	Pre-deployment and ongoing fairness audits	Governance committee sign-off before a model version deploys	Fuster <i>et al.</i> (2022); Kozodoi <i>et al.</i> (2022)
Route applications to a review depth staff can sustain	Risk-tiered underwriting triage	Staffing plan calibrated to the framework's routed volume	Berger and Udell (2002)
Preserve defensible underwriting decisions under examination or audit	Documented model governance and audit trail	Governance committee sign-off on model changes	Jagtiani and Lemieux (2018); Benjamin <i>et al.</i> (2004)

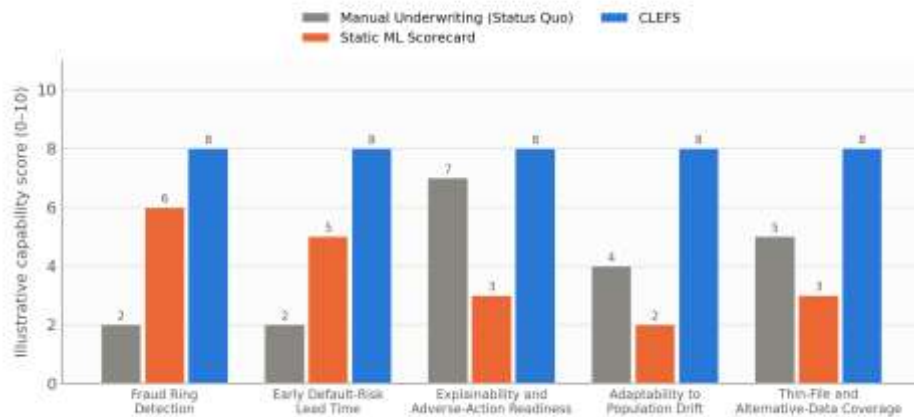
4.4 Case Scenario: A Community Lender's Adoption of the Framework

Consider a hypothetical community development financial institution with roughly \$250 million in assets, originating small-dollar consumer and microenterprise loans to a borrower population that is disproportionately thin-file, and staffed by a six-person underwriting and credit-risk team currently reviewing every application through a largely manual process supplemented by a purchased bureau score. Under the architecture proposed in Section 4.1, the institution first deploys feature engineering and alternative-data enrichment against its existing application intake, incorporating bank-transaction data applicants voluntarily connect and the free-text business description fields microenterprise applicants already complete but that the legacy process never systematically analyzed.

The dual-task risk scoring engine is then trained on the institution's historical application, disposition, and repayment-outcome data, with the credit-risk officer's early involvement addressing Berger and Udell's (2002) observation about the qualitative, unsystematized nature of much community-lending data directly: several years of paper-based reviewer notes turn out to contain risk-relevant judgments never captured in a structured field, and the institution digitizes a representative sample of these before training rather than discarding that information entirely. Once deployed, the explainability and fair-lending review layer attaches a feature-level rationale to every score — for an application flagged with an elevated early-risk score, the rationale cites the specific alternative-data and bureau features driving it, in a form the credit-risk officer can evaluate the same way she would evaluate a manual reviewer's written justification, and in a form suitable for an adverse-action notice if the application is ultimately declined. Eight months into deployment, the drift monitor detects a measurable shift in the applicant population's characteristics following a local economic disruption, consistent with the population-drift pattern Hand (2006) describes, and triggers a scheduled recalibration using the intervening months of confirmed outcome data; the governance and audit log records both the drift detection and the resulting model update, giving the credit-risk officer a defensible account to show a regulator or investor of why and when the model changed, rather than a static system whose behavior nobody can explain months after deployment.

4.5 Figure 1: Capability Coverage Across Comparative Approaches

Figure 1 *Capability Scores (0–10 Scale) for Manual Underwriting, a Static Machine Learning Scorecard, and CLEFS Across Five Dimensions*

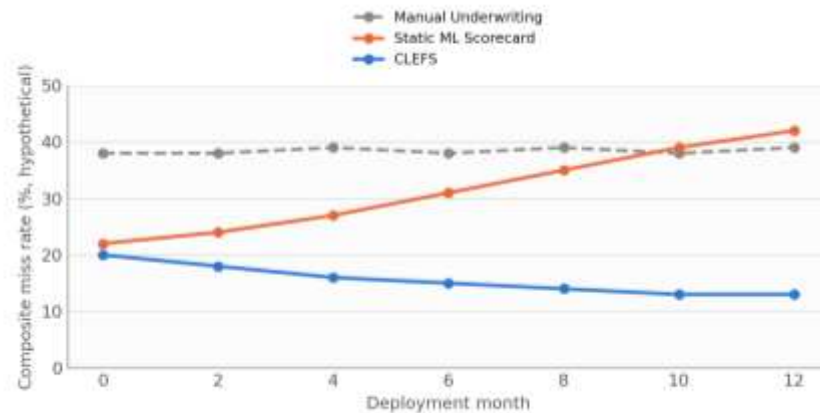


Note: These capability scores are estimates constructed for this paper's comparative framework rather than the product of a scored evaluation instrument; they were shaped by the relational fraud-detection gains Xu et al. (2022) and Bhattacharyya et al. (2011) document, the alternative-data gains Serrano-Cinca et al. (2015), Malekipirbazari and Aksakalli (2015), and Jiang et al. (2018) report, and the population-drift findings of Hand (2006) and Sousa et al. (2016), and should be read as a reasoning aid rather than a benchmark.

The pattern the figure is intended to convey mirrors the trade-off documented across Section 2: manual underwriting scores reasonably on explainability because a human reviewer's reasoning is inherently articulable, but poorly on fraud-ring detection and adaptability because an individual reviewer cannot systematically track relational patterns across applications or continuously recalibrate to a drifting population; the static scorecard improves detection and early-risk signaling initially but loses ground on explainability and, critically, on adaptability, since nothing in a static deployment corrects for drift. CLEFS is positioned, by design, to retain explainability while matching or exceeding the static scorecard's detection and signaling gains, because the framework's review layer and drift monitor were built specifically to close those two gaps.

4.6 Figure 2: Composite Fraud-and-Default Miss Rate Over a Twelve-Month Deployment Period

Figure 2 Hypothetical Composite Fraud-and-Early-Default Miss Rate (%) Over a Twelve-Month Deployment Period, by Underwriting Approach



Source: These trajectories are hypothetical operating points, not measurements from a deployed system; the static scorecard's rising miss rate reflects the population-drift decay pattern Hand (2006) documents, while CLEFS's stable and declining trajectory reflects the corrective effect of the continuous recalibration Sousa et al. (2016) recommend. The manual-underwriting trajectory's near-constant elevated rate is consistent with the limited systematic detection capacity Ngai et al. (2011) associate with review processes that do not incorporate learned models.

The figure shows the two machine learning approaches starting from a similar, substantially improved position relative to manual underwriting, then diverging sharply over the deployment period: the static scorecard's composite miss rate climbs steadily as the applicant population it was trained on drifts away from the population it now encounters, while CLEFS's continuous drift monitoring keeps its miss rate essentially flat and, as recalibration incorporates more confirmed outcomes, modestly declining. By month twelve, the static scorecard has lost most of its initial advantage over manual underwriting, why this paper treats drift monitoring as a required architectural component rather than an optional enhancement: without it, a machine learning deployment's advantage is temporary by construction.

4.7 Figure 3: Application Review Volume Against Confirmed Fraud-or-Early-Default Yield, by Configuration

Figure 3 Applications Flagged for Manual Review per 1,000 Applications Against Confirmed Fraud-or-Early-Default Yield (%), by Configuration

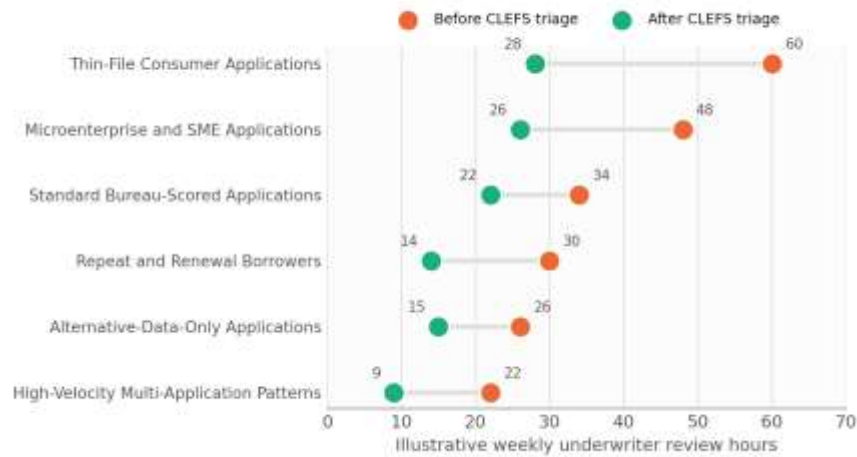


Note: These operating points are hypothetical, not measured outcomes; they were positioned to be directionally consistent with the detection gains Lessmann et al. (2015) and Xia et al. (2017) report for flexible classifiers and with the drift-driven decay Hand (2006) documents for a static scorecard left unmaintained.

This paper’s framework offers a risk-tiered choice rather than a single operating point: the low-sensitivity tier delivers the highest confirmed yield at the lowest review volume, appropriate for an underwriting team prioritizing precision over coverage, while the high-sensitivity tier accepts a lower yield in exchange for catching more marginal cases, and the combined-tiers configuration — routing lower-risk applicant segments through the low-sensitivity tier and higher-risk segments through the high-sensitivity tier — achieves a yield close to the low-sensitivity tier’s while still capturing the high-sensitivity tier’s broader coverage where it matters most. Every CLEFS configuration dominates both the manual-underwriting baseline and the drifted static scorecard on this trade-off, which is the figure’s central point: the choice a community lending institution faces is not whether to adopt a framework like CLEFS, but which of its tiers to weight more heavily given its own risk appetite and staffing.

4.8 Figure 4: Underwriter Review Hours Before and After Adaptive Triage, by Application Category

Figure 4 Weekly Underwriter Review Hours Before and After CLEFS Triage Deployment, by Application Category



Source: These review-hour figures are constructions, not measurements from a deployed system; the larger absolute reductions for thin-file consumer and microenterprise applications reflect these categories' historically higher manual-review burden under a process lacking systematic alternative-data enrichment, consistent with the coverage gains Serrano-Cinca et al. (2015), Malekipirbazari and Aksakalli (2015), and Jiang et al. (2018) document, while the more modest reductions for standard bureau-scored applications reflect a category where manual review was already comparatively efficient.

The figure shows the framework's largest absolute time savings falling precisely on the applicant categories that consumed the most underwriter capacity under the legacy process — exactly the thin-file and microenterprise applicants a community lending institution exists to serve — freeing that capacity for the high-velocity multi-application patterns most likely to represent coordinated fraud rather than reducing effort uniformly across every category. This pattern reflects the framework's underlying design logic: triage efficiency is not an end in itself, but a means of reallocating a community lending institution's limited underwriting capacity toward the applications most likely to represent genuine fraud or early credit risk.

5. Discussion

5.1 Implications for Practice

The framework developed here suggests several practical implications for community lending institutions considering machine learning-based fraud detection and early-risk signaling. Perhaps most important is that drift monitoring and fair-lending-grade explainability should be built into a deployment from the outset rather than added after a regulator raises a question or after the miss rate is observed creeping upward: Section 4's figures show that a static deployment's advantage over manual underwriting erodes within roughly a year without continuous maintenance, and retrofitting drift monitoring or an explainability layer onto an already-deployed model is a materially harder problem than designing for both from the start. A second implication concerns data readiness: Berger and

Udell's (2002) findings suggest that a community lending institution should inventory what qualitative, relationship-based information its current underwriting process already generates but does not systematically capture, since that information, once structured, may substitute for at least part of the bureau history a thin-file applicant lacks. A third implication is organizational: the risk-tiered triage stage requires the institution's own credit-risk and compliance staff, not a vendor or data science function alone, to set the sensitivity tiers according to the institution's mission and risk appetite, since Section 4.7's trade-off has no single correct answer independent of that judgment.

5.2 Limitations

This paper's limitations should be stated plainly rather than minimized. No production CLEFS deployment, real institutional applicant data, or actual fraud or default outcome data informed this study; the framework and objective-to-technique mapping in Section 4 are a synthesis of peer-reviewed findings rather than a validated system, and the two tables and four figures are explicitly constructions. Even setting the figures aside, the underlying literature carries real caveats a deploying institution would confront directly. Berger and Udell's (2002) findings on the qualitative, relationship-based nature of community lending underwriting apply with particular force here: a community institution's historical data may be too limited, unstructured, or inconsistently captured to train a reliable initial model without substantial remediation work the framework itself does not perform. The alternative-data literature this paper draws on, including Serrano-Cinca *et al.* (2015) and Malekipirbazari and Aksakalli (2015), was developed primarily on peer-to-peer lending platform data; community lending institutions' applicant populations and data infrastructure may differ enough from those platforms that the specific gains reported do not transfer at the same magnitude. Finally, the fairness and explainability methods Kozodoi *et al.* (2022) and Fuster *et al.* (2022) demonstrate were evaluated on general consumer credit-scoring datasets, not specifically on community lending institution portfolios, so their direct applicability to a CLEFS deployment, while plausible, has not itself been established in the cited literature.

5.3 What This Paper Does Not Argue

This paper does not argue that machine learning should replace manual underwriting judgment outright or that a community lending institution should discontinue existing controls while adopting the proposed framework; the architecture is intended to augment and inform underwriter judgment, not to remove human review from a lending decision that affects an applicant's access to credit. It does not argue that any specific machine learning algorithm, vendor platform, or explainability method should be adopted; the framework is intentionally technique-agnostic within each stage. It does not argue that CLEFS eliminates fraud or early default risk; Figure 3 shows every CLEFS configuration still producing confirmed yields below 30 percent, meaning a substantial share of flagged applications are not ultimately confirmed as fraudulent or high-risk, and the framework's contribution is reducing and better-targeting review volume, not eliminating the underlying risk. Finally, it does not argue that the specific figures in Section 4 describe any real institution's actual performance; they are a reasoning aid for evaluating the framework's internal logic, not a benchmark or a projection for any specific community bank, credit union, or CDFI.

5.4 Future Research

Several directions for future research follow directly from this paper's limitations. Empirical studies measuring actual fraud detection rates, early-default signaling lead time, and underwriter review hours at community lending institutions using machine learning-based origination screening, whether following this framework or another, would test this paper's central claims directly against practice. Research specifically evaluating population-drift timescales and triggers in community lending applicant pools, as distinct from the peer-to-peer platform and general consumer-credit settings much of the cited alternative-data and drift literature examines, would sharpen Section 4.1's drift-monitoring component's applicability to this paper's specific domain. Studies validating fairness and explainability methods against actual community lending institution portfolios and regulatory feedback, rather than against the general consumer-credit settings Kozodoi *et al.* (2022) and Fuster *et al.* (2022) study, would test whether the specific rationale format this paper's review layer proposes actually satisfies fair-lending expectations in practice for the borrower populations these institutions serve. Finally, research on how community lending institutions can practically structure and digitize the qualitative, relationship-based information Berger and Udell (2002) identify as central to their underwriting process, given their comparatively limited data science resources, would address the adoption barrier this paper's own limitations section identifies as most exposed.

6. Conclusion

Community lending institutions face a dual exposure to application fraud and undetected early default risk that the peer-reviewed literature has traditionally studied as two separate problems, and that exposure falls hardest on the community banks, credit unions, and CDFIs with the least capacity to absorb either loss given their typically thin capital base and mission-driven commitment to underserved borrowers. Machine learning has been shown, across a growing body of peer-reviewed work in lending fraud detection, credit risk scoring, and alternative-data underwriting, to improve both fraud detection and early-risk signaling, but the same body of literature shows that a model deployed once and left alone degrades under population drift, and that a model's score is not, by itself, the kind of rationale fair-lending and adverse-action obligations require. Synthesizing the lending-fraud, credit-scoring, alternative-data, early-warning, explainability and fair-lending, and community-lending-institution literatures, this paper developed the Community Lending Early-Risk and Fraud Signaling framework: an architecture in which a shared scoring engine produces both a fraud score and an early default-risk score from one enriched applicant record, continuous drift monitoring keeps that engine current, an explainability and fair-lending review layer attaches an adverse-action-ready rationale to every score, and risk-tiered triage routes the resulting workload to a review depth a small underwriting team can sustain, through two comparative tables, a hypothetical community lending institution scenario, and four figures clearly and repeatedly labeled as literature-grounded estimates rather than production measurements.

The paper's candor about its own limitations reflects the subject matter honestly: overstating the maturity of a framework with no production precedent would be a poor foundation for a

lending domain where the practical stakes — fair-lending compliance, investor confidence, and the institution’s continued ability to serve borrowers conventional lenders decline — fall directly on the community lending institutions considering adoption. What the literature reviewed here does support, and what this paper’s synthesis is intended to make actionable, is a considered path forward for institutions that have, until now, faced an apparent choice between an explainable but structurally limited manual underwriting process and a more capable but unexplainable and eventually stale machine learning scorecard: the Community Lending Early-Risk and Fraud Signaling framework proposes that this choice is false, and that the detection and prediction gains machine learning offers can be paired deliberately with the drift monitoring and fair-lending engineering the literature shows are necessary to sustain and defend those gains, sized from the outset to what a community lending institution’s underwriting team can actually operate.

References

- [1] Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*, 23(4), 589–609.
- [2] Rahman, M. S., & Tirupathi, D. (2023). A Privacy-Preserving Federated Architecture for Cross-Institutional Fraud-Signal Sharing Among Commercial Banks, Credit Unions, and Community Development Financial Institutions. *The Metascience*, 1(1), 332-340.
- [3] Baesens, B., Van Gestel, T., Viaene, S., Stepanova, M., Suykens, J., & Vanthienen, J. (2003). Benchmarking state-of-the-art classification algorithms for credit scoring. *Journal of the Operational Research Society*, 54(6), 627–635.
- [4] Potharaju, B., & Addanki, M. (2022). Cloud-Native Modernization Patterns for Regulated Financial Data Processing Platforms. *International Journal of Data Engineering and Intelligent Computing*, 5(2), 01-23.
- [5] Benjamin, L. M., Rubin, J. S., & Zielenbach, S. (2004). Community development financial institutions: Current issues and future prospects. *Journal of Urban Affairs*, 26(2), 177–195.
- [6] Berger, A. N., & Udell, G. F. (2002). Small business credit availability and relationship lending: The importance of bank organisational structure. *The Economic Journal*, 112(477), F32–F53.
- [7] Potharaju, B., & Ravipudi, J. (2023). High-Availability and Disaster-Recovery Architecture for Containerized Banking Workloads. *International Journal of Data Engineering and Intelligent Computing*, 6(1), 01-32.
- [8] Bhattacharyya, S., Jha, S., Tharakunnel, K., & Westland, J. C. (2011). Data mining for credit card fraud: A comparative study. *Decision Support Systems*, 50(3), 602–613.

- [9] Cole, R. A., & Gunther, J. W. (1998). Predicting bank failures: A comparison of on- and off-site monitoring systems. *Journal of Financial Services Research*, 13(2), 103–117.
- [10] Rahman, M. S., Islam, C. S., & Møller, C. (2022). Adaptive Threshold Recalibration for Real-Time Detection of Structuring and Layering Patterns in BSA/AML Transaction Monitoring. *The Computertech*, 56-67
- [11] Cole, R. A., & White, L. J. (2012). Déjà vu all over again: The causes of U.S. commercial bank failures this time around. *Journal of Financial Services Research*, 42(1), 5–29.
- [12] Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., & Walther, A. (2022). Predictably unequal? The effects of machine learning on credit markets. *Journal of Finance*, 77(1), 5–47.
- [13] Siddiqui, M. I. H., Rahman, M. S., Kabir, A. A., Mahmud, F. U., Rashid, S. U., & Shammah, R. S. (2023). Comparative analysis of explainable machine learning models for cancer classification using cytological features. *Journal of Medical and Health Studies*, 4(5), 110-150.
- [14] Gramespacher, T., & Posth, J.-A. (2021). Employing explainable AI to optimize the return target function of a loan portfolio. *Frontiers in Artificial Intelligence*, 4, Article 693022.
- [15] Malekipirbazari, M., & Aksakalli, V. (2015). Risk assessment in social lending via random forests. *Expert Systems with Applications*, 42(10), 4621–4631.
- [16] Potharaju, B., & Addanki, M. (2020). Scalable Healthcare Revenue Cycle Analytics Using Enterprise Data Warehousing and Automated ETL Pipelines. *International Journal of Data Engineering and Intelligent Computing*, 3(2), 01-21.
- [17] Ngai, E. W. T., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decision Support Systems*, 50(3), 559–569.
- [18] Serrano-Cinca, C., Gutiérrez-Nieto, B., & López-Palacios, L. (2015). Determinants of default in P2P lending. *PLOS ONE*, 10(10), e0139427.
- [19] Sousa, M. R., Gama, J., & Brandão, E. (2016). A new dynamic modeling framework for credit risk assessment. *Expert Systems with Applications*, 45, 341–351. Xia, Y., Liu, C., Li, Y., & Liu, N. (2017). A boosted decision tree approach using Bayesian hyper-parameter optimization for credit scoring. *Expert Systems with Applications*, 78, 225–241.
- [20] Xu, J. J., Chen, D., Chau, M., Li, L., & Zheng, H. (2022). Peer-to-peer loan fraud detection: Constructing features from transaction data. *MIS Quarterly*, 46(3), 1777–1792.

- [21] Karim, N. B., & Elias, D. (2022). Multi-Institution Validation of an AI-Assisted Early-Intervention and Loan-Restructuring Recommendation Engine for Financially Distressed Small-Business Borrowers Across US Community Banks, Credit Unions, and CDFI Portfolios. *The Computertech*, 107-120.
- [22] Hand, D. J. (2006). Classifier technology and the illusion of progress. *Statistical Science*, 21(1), 1–14.
- [23] Lessmann, S., Baesens, B., Seow, H.-V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research*, 247(1), 124–136.
- [24] Karim, N. B., & Seitz, J. (2023). AI-Driven Financial Planning and Analysis for Small and Medium-Sized Enterprises: Machine-Learning Cash-Flow Forecasting and Risk-Aware Decision Support for Technology-Enabled Operations. *International Journal of Modern Computing*, 6(1), 158-176.
- [25] Jagtiani, J., & Lemieux, C. (2018). Do fintech lenders penetrate areas that are underserved by traditional banks? *Journal of Economics and Business*, 100, 43–54.
- [26] Rahman, M. S., Bhowmik, P. K., Hossain, B., Tannier, N. R., Amjad, M. H. H., Chouksey, A., & Hossain, M. (2023). Enhancing fraud detection systems in the USA: A machine learning approach to identifying anomalous transactions. *Journal of Economics, Finance and Accounting Studies*, 5(5), 145-160.
- [27] Jiang, C., Wang, Z., Wang, R., & Ding, Y. (2018). Loan default prediction by combining soft information extracted from descriptive text in online peer-to-peer lending. *Annals of Operations Research*, 266(1–2), 511–529.
- [28] Kozodoi, N., Jacob, J., & Lessmann, S. (2022). Fairness in credit scoring: Assessment, implementation and profit implications. *European Journal of Operational Research*, 297(3), 1083–1094.
- [29] Potharaju, B., & Sajja, R. R. (2021). Instrument-Level Transfer Pricing Analytics Using Automated Data Pipelines and Financial Calculation Frameworks. *International Journal of Commerce, Finance and Digital Economy*, 4(2), 01-34.