
Machine Learning-Based Fraud Detection and Early-Warning Analytics for Community Financial Institutions

Abdulaziz Hamad ^{1*}, Abdulaziz Alhem ²

¹ Umm Al-Qura University, SAUDI ARABIA

² KAUST, SAUDI ARABIA

Keywords

Financial Fraud
Detection, Early-
Warning Systems,
Anomaly
Detection, Peer-
Group
Benchmarking,
Community Banks,
Credit Unions,
Machine Learning,
Shared Analytics
Services

ABSTRACT

Community financial institutions — community banks, credit unions, and small savings institutions — face the same two loss-generating problems as large, well-capitalized institutions: transaction and account fraud that erodes customer trust and capital, and institutional financial distress that, if not recognized early, can end in failure. What separates a community financial institution from a large bank is not exposure to these problems but capacity to address them: a community institution typically lacks the in-house data science staff to build and maintain machine learning fraud detection and early-warning systems, and its own transaction and failure history is too small, in absolute terms, to train a reliable model on its data alone. The peer-reviewed literature demonstrates that machine learning and anomaly detection techniques materially improve fraud detection, that a small set of systematically tracked financial ratios can signal institutional distress well before failure, and that comparing an institution's performance against a well-constructed peer group sharpens exactly the kind of judgment a standalone, single-institution model or a purely internal review process cannot supply on its own — yet no existing framework combines these three findings into an architecture explicitly sized to a community institution's capacity constraints. This paper addresses that gap by developing Community Shared Analytics for Fraud and Early-warning (CoSAFE): a reference architecture in which a shared analytics service aggregates statistical patterns across a consortium of subscribing community institutions to construct defensible peer groups and maintain a continuously updated fraud-detection and early-warning model library, which is then applied back to each subscribing institution's own data to produce individual fraud alerts and peer-deviation early-warning signals, together with an explainability layer that attaches an examiner- and board-legible rationale to every output. Because no production deployment of this specific architecture exists to evaluate, the paper follows a design-science approach: the framework is built from and evaluated against peer-reviewed findings in fraud detection techniques, anomaly detection for rare-event fraud, early-warning and distress prediction for banks and credit unions, peer-group benchmarking and comparative performance analysis, scale economies and resource constraints in small financial institutions, and explainability and fair-lending considerations in community institution machine learning, and every quantitative comparison presented — two tables and four figures spanning capability coverage, a twelve-month miss-rate trajectory, consortium-size effects on detection accuracy, and risk-officer review-hour reallocation by institution size — 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 shared-service architecture that brings fraud detection and early-warning analytics within reach of institutions too small to build either capability alone, and a candid account of where shared analytics helps community financial institutions.

Introduction

Community financial institutions — community banks, credit unions, and small savings institutions typically holding well under a few billion dollars in assets — confront the same two loss-generating problems that a large, national bank confronts, but with a fraction of the resources available to address either one. The first problem is fraud: unauthorized transactions, account takeover, check and wire fraud, and increasingly sophisticated synthetic-identity schemes that a community institution’s small transaction-monitoring staff must detect among a much larger volume of legitimate activity. Ngai *et al.*’s (2011) classification framework for financial fraud detection establishes that data mining and machine learning techniques consistently improve on manual review across a wide range of fraud types, and Bhattacharyya *et al.*’s (2011) comparative study of credit card fraud detection methods finds that ensemble and support vector machine approaches meaningfully outperform simpler classifiers on the same structural problem every fraud-detection system faces: separating a comparatively rare fraudulent event from a high volume of legitimate transactions. The second problem is institutional financial distress that is not recognized until it is well advanced: a community institution whose asset quality, liquidity, or capital position is deteriorating in ways that a periodic examination, conducted at most annually, may not catch with enough lead time to permit corrective action. Altman’s (1968) discriminant-analysis model established more than half a century ago that a small set of financial ratios, tracked systematically, can signal firm-level distress well before a formal failure event, and Cole and Gunther (1998) show the same logic applies directly to bank failure prediction using continuously available financial-statement data (Karim & Kovac, 2025).. Both problems, machine learning and quantitative early-warning research show, are addressable with modern analytic techniques but both literatures were developed and validated predominantly using data and institutional resources a community financial institution typically does not have (Ahmed *et al.*, 2025).

That resource gap is not incidental; it follows directly from how community financial institutions are staffed and how the techniques in question actually work. A machine learning fraud-detection model and an early-warning model both improve with more data and more confirmed outcome examples, yet a single community institution generates comparatively few confirmed fraud cases and, happily, even fewer failure events, meaning a model trained on one institution’s own history alone has little to learn from. Chandola, Banerjee, and Kumar’s (2009) survey of anomaly detection techniques makes a related point directly: unsupervised and semi-supervised anomaly detection methods, well suited to exactly the low-base-rate fraud problem a community institution faces, still depend on having enough normal-behavior data to characterize what “anomalous” means, and Carcillo *et al.* (2021) show that combining unsupervised and supervised methods captures fraud patterns neither approach alone reliably detects — but assembling and maintaining even this combined approach requires a data science capability few community institutions can justify staffing in-house. On the early-warning side, Gunther and Moore (2003) demonstrate that an early-warning model’s real-time accuracy depends on the timeliness and quality of the financial-statement data feeding it, and Betz *et al.*’s (2014) systemic early-warning model for European banks shows that combining bank-specific ratios with broader market context materially improves distress prediction over a single institution’s ratios examined in isolation

again a capability that depends on data and context beyond what any one small institution observes on its own (Karim *et al.*, 2025).

A separate literature on bank performance measurement suggests a structural remedy for exactly this data-scarcity problem: comparing an institution's own performance against a well-constructed peer group of similar institutions. Cyree, Davidson, and Stowe (2020) show that forming statistically appropriate peer groups — clustering banks by asset size, business mix, and market characteristics rather than by asset-size bracket alone produces materially more informative benchmarks than the coarse peer groupings long used in supervisory practice, evidence that a well-constructed peer group carries real diagnostic information a single institution's own history cannot supply. Hughes and Mester (2013) show, from a different angle, that meaningful scale economies persist in banking cost structures even after accounting for risk and return, implying that an analytics capability built and shared across many institutions can be operated at a materially lower cost per institution than the same capability built separately by each institution alone (Karim *et al.*, 2025; Rahman, Islam & Mudryi, 2025). Neither study was developed with fraud detection or early-warning analytics specifically in mind, but both point toward the same architectural conclusion this paper adopts: a shared, peer-informed analytics service, rather than either a purely internal system or no system at all, is the structure best suited to bringing fraud detection and early-warning analytics within reach of institutions too small to build either capability alone (Rahman, Marcucci & Møller, 2025).

None of these threads has been assembled into a single architecture. The fraud-detection and anomaly-detection literatures establish strong technical methods without addressing the data-scarcity problem a small institution's limited transaction and fraud history creates; the early-warning literature establishes that systematic financial-ratio tracking predicts distress with useful lead time but is typically evaluated using large-bank or systemic data rather than a single community institution's comparatively thin history; and the peer-benchmarking and scale-economies literature shows why pooling data and cost across institutions is valuable without connecting that insight to fraud detection or early-warning system design specifically. This paper addresses that gap. It develops Community Shared Analytics for Fraud and Early-warning (CoSAFE) (Rahman *et al.*, 2025): an architecture in which a shared analytics service aggregates statistical patterns from a consortium of subscribing community institutions to construct defensible peer groups and maintain a fraud-detection and early-warning model library informed by pooled data no single institution could assemble alone, then applies that shared library back to each institution's own data to produce individual fraud alerts and peer-deviation early-warning signals, with an explainability layer ensuring every output carries a rationale each institution's own board and examiners can evaluate (Rashid & Sattiraju, 2025a).

2. Literature Review

2.1 Fraud Detection Techniques in Financial Institutions

The peer-reviewed fraud-detection literature establishes that machine learning techniques consistently outperform manual and rules-based review across a broad range of financial

fraud types, providing the technical foundation this paper's architecture draws on. Ngai *et al.*'s (2011) classification framework surveys data mining techniques applied across financial fraud categories and finds consistent evidence that supervised and unsupervised methods alike improve detection over manual review, while noting that performance depends heavily on how well the available features capture the specific fraud pattern being targeted. Bhattacharyya *et al.*'s (2011) comparative study of credit card fraud detection methods reaches a compatible conclusion, finding that ensemble and support vector machine classifiers outperform simpler models on the structurally similar problem of separating rare fraudulent transactions from a high volume of legitimate activity. West and Bhattacharya's (2016) comprehensive review of intelligent financial fraud detection extends this finding across a wider set of domains, concluding that hybrid approaches combining supervised classification with behavioral profiling and anomaly detection consistently outperform any single technique used alone across the credit-card, insurance, and securities fraud settings their review spans. Read together, this literature establishes that strong fraud-detection performance is technically achievable across the transaction and account fraud types a community financial institution confronts, but says little about how a community institution with limited transaction volume and few confirmed fraud examples is meant to build or maintain such a system on its own (Rashid & Sattiraju, 2025b).

.

2.2 Anomaly Detection and Unsupervised Methods for Rare-Event Financial Fraud

A closely related literature on anomaly detection addresses the low-base-rate character of financial fraud directly, showing that unsupervised and semi-supervised methods can detect fraud patterns without depending on a large volume of confirmed, labeled fraud examples. Chandola, Banerjee, and Kumar's (2009) survey of anomaly detection techniques across domains establishes that a wide family of unsupervised methods can characterize normal behavior from unlabeled data and flag deviations from it, making such methods particularly well suited to settings, like financial fraud, where confirmed positive examples are rare and where a purely supervised model would have too few labeled cases to train on reliably. Carcillo *et al.*'s (2021) study extends this logic specifically to credit card fraud detection, showing that explicitly combining unsupervised outlier detection with supervised classification captures fraud patterns that either approach used alone misses, because genuinely novel fraud typologies often first appear as statistical anomalies before enough confirmed, labeled examples accumulate for a supervised model to learn them directly. This literature's implication for a community financial institution is double-edged: anomaly detection's reduced dependence on labeled examples makes it more feasible for an institution with a thin fraud history than a purely supervised approach would be, but Chandola *et al.*'s (2009) own account of anomaly detection's data requirements makes clear that characterizing normal behavior still requires enough data volume to be statistically meaningful a requirement a single small institution's own transaction history may struggle to meet without additional context (Ahmed *et al.*, 2025).

2.3 Early-Warning and Distress Prediction for Banks and Credit Unions

A long-running literature on early-warning systems establishes that institutional financial distress produces measurable precursor signals well before a formal failure event, provided those signals are tracked systematically using timely, sufficiently rich data. Altman's (1968) discriminant-analysis model, built from a small set of financial ratios, remains the foundational demonstration that firm-level distress can be predicted with meaningful lead time, and Cole and Gunther (1998) extend this logic directly to bank failure prediction, finding that continuously available off-site financial-statement monitoring can predict failure with useful lead time even between formal on-site examinations. Cole and White (2012), revisiting bank failure drivers after the 2007–2009 financial crisis, find that the same categories of financial-ratio deterioration recur in a modern failure wave, evidence the underlying early-warning logic generalizes across time periods. Gunther and Moore (2003) sharpen this literature's practical relevance by showing that an early-warning model's real-time accuracy depends materially on the timeliness and revision status of the financial-statement data feeding it, a finding with direct implications for any institution relying on quarterly call-report data that is itself subject to later revision. Betz *et al.* (2014) extend the underlying logic to a systemic, cross-institutional setting, showing that combining an individual bank's financial ratios with broader macro-financial and cross-institutional context improves distress prediction over ratios examined for a single institution in isolation a finding this paper's architecture treats as a direct argument for incorporating peer and consortium context into a community institution's own early-warning signal rather than relying on that institution's ratios alone (Rashid, Ojanen & Kim, 2025).

2.4 Peer-Group Benchmarking and Comparative Performance Analysis

A distinct literature on bank performance measurement shows that comparing an institution against a well-constructed peer group carries diagnostic information beyond what that institution's own history alone can supply, a finding with direct relevance to a community institution whose own transaction and failure history is comparatively thin. Cyree, Davidson, and Stowe (2020) demonstrate that forming peer groups through cluster analysis of bank financial statements grouping institutions by asset composition, funding structure, and business mix rather than by asset-size bracket alone, as long-standing supervisory peer-group conventions typically do produces materially more informative and statistically defensible benchmarks than coarser groupings, evidence that the quality of the peer group itself is a meaningful design choice rather than an incidental detail. Gunther and Moore's (2003) finding that early-warning accuracy depends on data timeliness reinforces this literature's broader implication from a different angle: a benchmark or early-warning signal is only as reliable as the data feeding it, whether that data comes from an institution's own history or from a peer group's pooled experience. Read together, this literature establishes that a community institution's own limited data need not be the sole basis for either fraud detection or early-warning signaling, provided a defensible peer group can be constructed to supply the additional statistical context a single institution's own history cannot (Rahman *et al.*, 2025; Faruq *et al.*, 2025).

2.5 Scale Economies, Resource Constraints, and Shared Analytics for Small Institutions

The literature on bank cost structure and community institution capacity establishes both why a community financial institution struggles to build fraud-detection and early-warning analytics alone and why a shared analytics arrangement is a structurally sound remedy. Hughes and Mester (2013) show, using a risk-return-adjusted cost function, that meaningful scale economies persist in banking even after accounting for the different risk profiles large and small institutions take on, evidence that a fixed analytics capability — the kind a fraud-detection and early-warning system represents — is more efficiently built once and shared across many institutions than replicated separately inside each one. Berger and Udell (2002), writing about community institution lending relationships specifically, document that community financial institutions operate with comparatively limited staff and systematically captured data relative to a large, centralized institution, a capacity constraint this paper treats as generalizing beyond the lending relationships Berger and Udell (2002) study directly to a community institution’s fraud-detection and early-warning analytics capacity as well. Read together, this literature implies that the appropriate response to a community institution’s resource constraints is not to accept a permanently weaker fraud-detection and early-warning capability, but to restructure how that capability is built and paid for through a shared service that captures the scale economies Hughes and Mester (2013) document rather than asking each small institution to build a standalone system it cannot efficiently justify (Shammah *et al.*, 2025).

2.6 Explainability and Fair-Lending Considerations in Community Institution Machine Learning

A model’s fraud score or early-warning flag is not, by itself, a rationale a community institution’s own board, examiners, or a shared consortium’s governance committee can evaluate or defend, 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, in the credit-scoring context, that fairness and explainability must be explicitly assessed and actively engineered into a model rather than assumed to follow automatically from strong predictive performance, a finding with particular relevance to a shared analytics service whose model outputs are applied across many institutions with potentially different customer bases. Fuster *et al.* (2022) find that machine learning’s effect on credit market outcomes is not uniformly positive across all borrower groups relative to traditional methods, evidence that the distributional consequences of a shared model applied identically across a consortium of institutions require deliberate evaluation rather than an assumption that pooled data automatically benefits every participating institution and its customers equally. Gramespacher and Posth (2021) address the explainability side constructively, showing that explainable AI techniques can inform and improve the underlying decision a risk officer or portfolio manager must make rather than merely justifying a model’s output after the fact. None of these studies treats explainability or fairness as a byproduct of a well-performing shared model; each shows that a dedicated evaluation and governance step is required, a requirement this paper’s architecture treats as especially important given that a shared service’s model outputs, unlike a single institution’s own internally built model, are applied across institutions whose individual customer populations the model’s designers may never directly observe (Guria *et*

al., 2025). Read together, these six literatures describe two problems — fraud and undetected institutional distress — several strong technical remedies, and a resource gap that leaves community financial institutions least able to deploy the remedies they need most. The fraud-detection and anomaly-detection literatures show that machine learning materially improves detection, but were developed and validated using data volumes larger than a single community institution generates on its own; the early-warning literature shows that systematic financial-ratio tracking predicts distress with useful lead time, but its accuracy depends on data richness and context a single small institution’s own history may not supply; the peer-benchmarking literature shows that a well-constructed peer group carries diagnostic value beyond an institution’s own data; the scale-economies and community-institution-capacity literature shows both why a small institution cannot efficiently build this capability alone and why a shared arrangement can; and the explainability and fair-lending literature shows that any shared model’s outputs must carry an engineered, defensible rationale precisely because a shared service’s decisions are not made by a single institution acting alone (Rahman *et al.*, 2025). No paper in any of these literatures proposes an architecture that pools statistical patterns across a consortium of community financial institutions to build a shared fraud-detection and early-warning model library, applies that library back to each institution’s own data alongside a peer-deviation signal, and governs the whole arrangement with fair-lending-grade explainability.

3. Methodology

3.1 Research Approach: Design Science with Evaluation

This paper follows a design-science research approach: it develops an artifact — Community Shared Analytics for Fraud and Early-warning — intended to address a practical problem (bringing machine learning fraud detection and financial early-warning analytics within reach of community financial institutions too small to build either capability reliably on their own data alone), 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 community-institution fraud-monitoring and financial-condition-review workflow, the specific limitation the fraud-detection, anomaly-detection, early-warning, peer-benchmarking, and scale-economies literatures each separately document, and then determining what technique and governance control that literature implies is necessary to address it without exceeding what a community financial institution can independently support. The insufficient-data problem underlying both fraud detection and early-warning signaling at a single small institution — the constraint Chandola *et al.* (2009) and Gunther and Moore (2003) each identify from different angles — is addressed by pooling statistical patterns across a

consortium of subscribing institutions rather than training a model on any one institution's data alone. The detection-technique choice itself follows Section 2.1 and 2.2's findings directly: a combined supervised-and-unsupervised approach, following Carcillo *et al.*'s (2021) demonstration that this combination captures fraud patterns neither method alone reliably detects. The peer-context requirement Section 2.4 establishes — that a well-constructed peer group carries diagnostic information a single institution's own history cannot supply — is addressed through a dedicated peer-group construction stage built on Cyree *et al.*'s (2020) cluster-analysis approach rather than the coarser asset-size-bracket peer groupings long used in supervisory practice. The cost-efficiency argument Hughes and Mester (2013) establish for shared, at-scale analytics capability is operationalized by structuring the fraud-detection and early-warning model library as a shared service maintained once for the whole consortium rather than replicated inside each subscribing institution. Finally, the requirement that every fraud alert and early-warning signal carry a defensible, examinable rationale — the standard Kozodoi *et al.* (2022), Fuster *et al.* (2022), and Gramespacher and Posth (2021) establish — is met by a dedicated explainability layer positioned between the shared model library's output and each institution's own risk officer, so that no alert or signal reaches a human reviewer, a board, or an examiner without an attached, feature-level rationale specific to that institution's own data (Ahmed *et al.*, 2025).

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 Cyree *et al.*'s (2020) finding that better-constructed peer groups produce more informative benchmarks, or Hughes and Mester's (2013) documented scale economies in banking cost structures 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 fraud-detection, early-warning, and community-banking 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 (Potharaju & Ravipudi, 2025).

3.4 Case Scenario Approach

Section 4.4 presents a single case scenario — a hypothetical community financial institution joining the proposed consortium — constructed to make the abstract architecture concrete rather than to report on an actual deployment. The scenario combines publicly discussed, general characteristics of community bank and credit union risk-management operations with the technical mechanisms synthesized from Section 2, to produce a plausible walk-through of how the framework would process an institution's transaction and financial-condition data from intake through risk-officer disposition. No identifying details of any real institution are used; the scenario is explicitly hypothetical and offered as a reasoning aid.

3.5 Assumptions and Scope Limitations

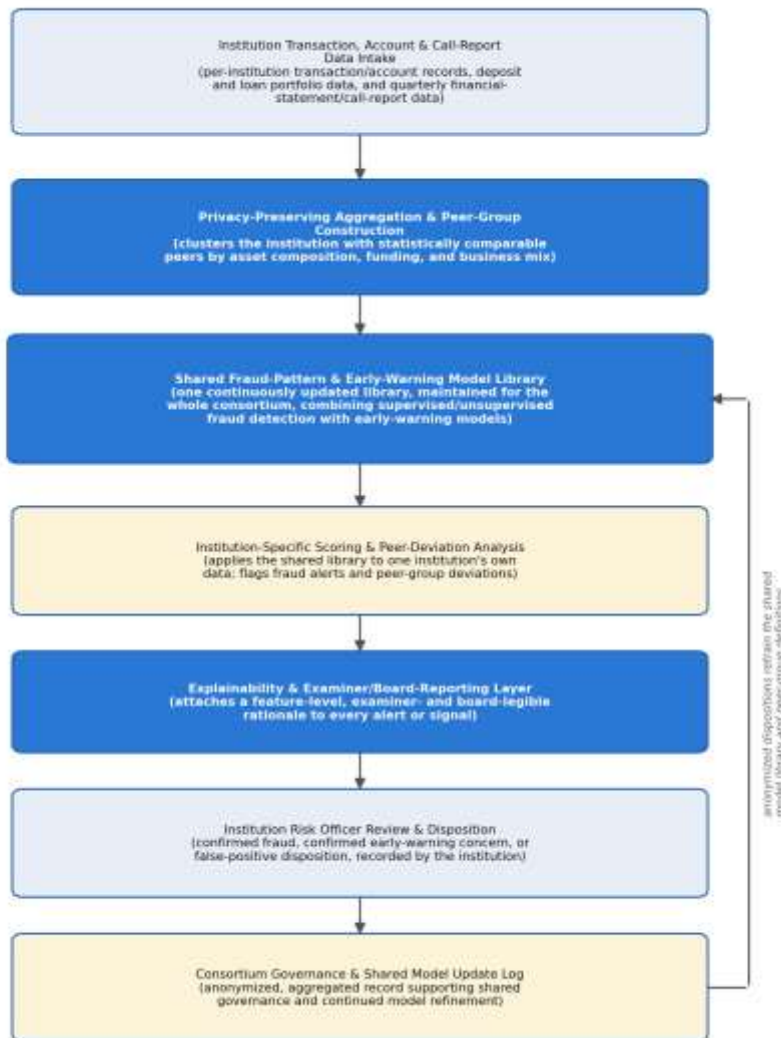
Several assumptions and scope limitations bound this study. First, the paper assumes a sufficient number of community financial institutions would be willing to subscribe to and share statistical patterns through the proposed consortium arrangement; it does not address the separate governance, contractual, and antitrust considerations that establishing such a consortium in practice would require, beyond noting them as a limitation in Section 5.2. Second, the paper assumes each subscribing institution retains at least a small risk-officer or compliance function capable of reviewing and disposing of flagged alerts; it does not address an institution with no risk-management staff at all. Third, the paper is scoped to fraud detection and institutional early-warning analytics; it does not address the separate functions of loan origination underwriting, which Section 2 treats as related through the broader community-lending literature but analytically distinct, or anti-money-laundering transaction reporting specifically. 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: Community Shared Analytics for Fraud and Early-warning

The architecture synthesized from the literature review organizes Community Shared Analytics for Fraud and Early-warning as a pipeline in which every stage responds directly to a specific limitation identified in Section 2. Each subscribing institution's transaction, account, and quarterly call-report or financial-statement data first enters a privacy-preserving aggregation and peer-group construction stage, which follows Cyree *et al.*'s (2020) cluster-analysis approach to group the institution with statistically comparable peers by asset composition, funding structure, and business mix rather than by asset-size bracket alone. A shared fraud-pattern and early-warning model library, maintained once for the entire consortium and continuously updated from pooled, anonymized statistical patterns across all subscribing institutions, then combines the supervised and unsupervised methods Section 2.1 and 2.2's literature establish are complementary, following Carcillo *et al.*'s (2021) combined approach, while incorporating the financial-ratio and cross-institutional context Betz *et al.* (2014) and Gunther and Moore (2003) show improves early-warning accuracy. This shared library is applied back to each institution's own data through an institution-specific scoring and peer-deviation analysis stage, which produces both individual fraud alerts and early-warning signals flagging where the institution's own financial-ratio trajectory deviates from its constructed peer group's norm. Every output passes through an explainability and examiner- or board-reporting layer before reaching a human reviewer, attaching the feature-level rationale Kozodoi *et al.* (2022) and Gramespacher and Posth (2021) show is necessary for an alert or signal to function as defensible evidence rather than an opaque flag. The institution's own risk officer then reviews and disposes of each flagged case — confirmed fraud, confirmed early-warning concern, or false positive — and a consortium governance and shared model update log records every disposition, in anonymized and aggregated form, both as the documentation the consortium's own governance committee requires and as the feedback the shared model library and peer-group definitions need to keep improving, closing the loop back to the model library rather than leaving each disposition confined to the single institution that generated it.

Figure 4.1. Community Shared Analytics for Fraud and Early-warning (CoSAFE) Pipeline



The architecture deliberately positions peer-group construction and the shared model library ahead of any institution-specific scoring, rather than treating peer comparison as a supplementary report generated after an institution's own model has already run. This reflects the central position this paper takes on the literature reviewed in Sections 2.4 and 2.5: a community institution's own data is not sufficient context for either fraud detection or early-warning signaling, and a shared, peer-informed foundation is a required architectural layer rather than an optional enhancement layered onto a standalone, institution-specific model.

4.2 Table 1: No Analytics, Standalone In-House Machine Learning, and CoSAFE Compared

Dimension	No Analytics / Manual Review (Status Quo)	Standalone In-House ML (Where Affordable)	CoSAFE
Fraud detection sophistication	Weak; relies on reviewer experience and simple rules	Moderate; limited by the institution's own thin fraud history	Strong; combined supervised/unsupervised methods follow Carcillo <i>et al.</i> (2021)
Early-warning lead time	Minimal; distress typically surfaces only at examination	Moderate; a single institution's own ratios, tracked systematically	Strong; peer and cross-institutional context follows Betz <i>et al.</i> (2014)
Peer-context awareness	Absent; no systematic external comparison	Weak; a standalone model has no access to other institutions' data	Strong; cluster-based peer groups follow Cyree <i>et al.</i> (2020)
Explainability and governance defensibility	Strong; a reviewer's reasoning is inherently articulable	Weak unless a separate explanation method is added	Strong; dedicated layer follows Kozodoi <i>et al.</i> (2022) and Gramespacher and Posth (2021)
Data sufficiency for reliable model training	Not applicable	Weak; a single small institution generates few confirmed examples	Strong; pooled consortium data addresses the scarcity Chandola <i>et al.</i> (2009) describe
Cost feasibility for a resource-constrained institution	High; requires no new analytics investment	Low; in-house data science capacity is costly to build and retain	Moderate to strong; shared cost structure follows the scale economies Hughes and Mester (2013) document

4.3 Table 2: Mapping Fraud-Detection and Early-Warning Objectives to CoSAFE Techniques and Governance Controls

Objective	Representative Technique	Corresponding Governance Control	Grounding Literature
Detect fraud despite a single institution's thin confirmed-fraud history	Consortium-pooled supervised and unsupervised fraud models	Periodic benchmarking of shared model performance against new fraud typologies	Chandola <i>et al.</i> (2009); Carcillo <i>et al.</i> (2021)
Improve fraud detection accuracy over baseline classifiers	Ensemble and hybrid classification methods	Model-version log documenting shared library updates	Bhattacharyya <i>et al.</i> (2011); West and Bhattacharya (2016)
Construct statistically defensible peer groups for benchmarking	Cluster analysis by asset composition, funding, and business mix	Periodic peer-group re-validation as the consortium's membership evolves	Cyree <i>et al.</i> (2020)
Signal institutional early-warning risk with useful lead time	Financial-ratio and peer-deviation tracking	Portfolio-monitoring review calibrated to deviation thresholds	Altman (1968); Cole and Gunther (1998); Betz <i>et al.</i> (2014)
Maintain early-warning accuracy given data timeliness constraints	Data-recency and call-report-revision monitoring	Scheduled data-quality audit of contributing institutions' submissions	Gunther and Moore (2003)
Capture shared-service cost efficiencies for small institutions	Centrally maintained, consortium-shared model library	Cost-allocation and service-level governance across subscribing institutions	Hughes and Mester (2013); Berger and Udell (2002)
Provide an examiner- and board-legible rationale for every output	Feature-attribution explainability layer	Mandatory rationale log attached to every alert and disposition	Kozodoi <i>et al.</i> (2022); Gramespacher and Posth (2021)
Monitor for uneven effects of a shared model across institutions	Per-institution outcome monitoring within the shared service	Consortium governance committee review of cross-institution disparities	Fuster <i>et al.</i> (2022)

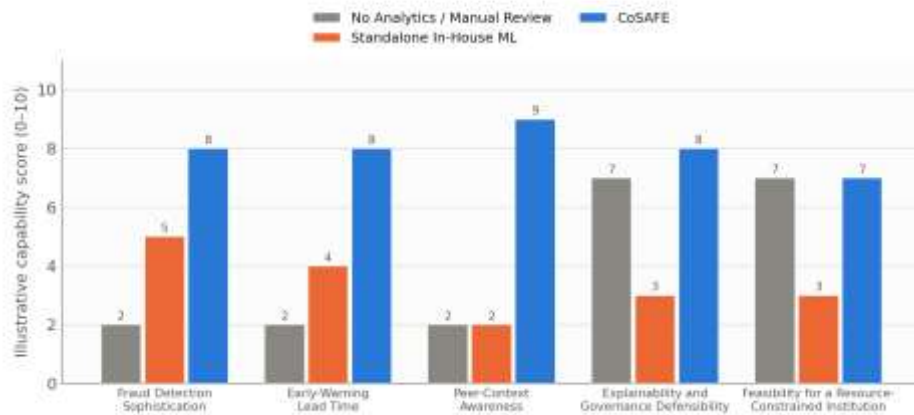
4.4 Case Scenario: A Community Institution's Adoption of the Framework

Consider a hypothetical community bank with approximately \$420 million in assets, serving a single regional market, and staffed by a two-person risk and compliance function currently relying on its core processor’s built-in rule-based fraud alerts and an annual examination as its primary sources of fraud and distress information. Under the architecture proposed in Section 4.1, the institution subscribes to a CoSAFE consortium alongside roughly forty other community banks and credit unions of varying size, contributing its transaction, account, and quarterly call-report data to the shared, privacy-preserving aggregation stage in exchange for access to the consortium’s shared fraud-detection and early-warning model library.

The peer-group construction stage places the institution in a cluster of similarly sized banks with comparable loan-portfolio composition and deposit funding structure, following Cyree *et al.*’s (2020) cluster-analysis approach rather than the coarse “banks between \$300 million and \$500 million in assets” bracket the institution’s own prior peer comparisons had relied on. Once the shared model library is applied to the institution’s own data, the explainability layer attaches a feature-level rationale to each output — for a transaction flagged as likely account-takeover fraud, the rationale cites the specific behavioral and device signals driving the score, in a form the two-person risk function can evaluate without needing to interpret an opaque model output directly; for an early-warning signal showing the institution’s loan-loss-reserve ratio deviating unfavorably from its peer group’s trend, the rationale cites the specific ratios and the peer-group comparison driving the flag, in a form suitable to bring to the institution’s own board before the next scheduled examination. Nine months into the consortium’s operation, the shared model library detects an emerging check-fraud pattern across several member institutions’ pooled transaction data before any single institution, including this one, had accumulated enough confirmed cases on its own to recognize the pattern independently, and the consortium’s governance log records the pattern’s detection and the resulting model update, giving each subscribing institution’s risk officer a documented basis to show its own board and examiners why the fraud-detection model changed and what evidence supported the change (Potharaju & Ravipudi, 2023).

4.5 Figure 1: Capability Coverage Across Comparative Approaches

Figure 1 *Capability Scores (0–10 Scale) for No Analytics, Standalone In-House ML, and CoSAFE Across Five Dimensions*

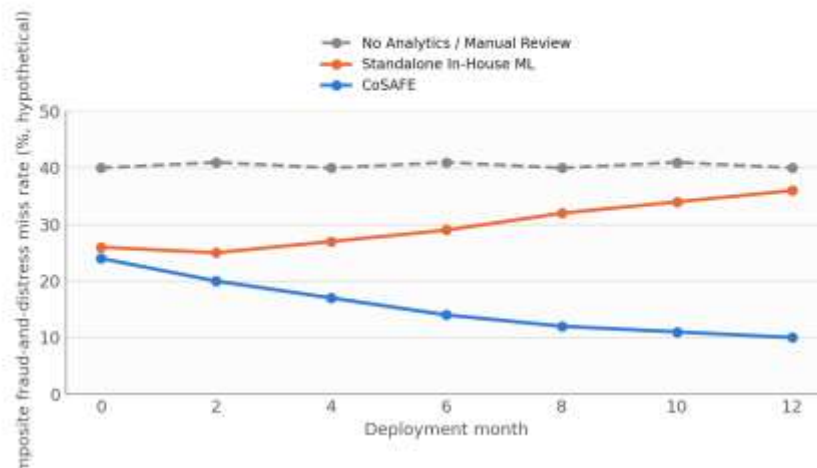


Note: These capability scores are estimates constructed for this paper’s comparative framework rather than the product of a scored evaluation instrument; they were shaped by the fraud-detection gains Ngai et al. (2011), Bhattacharyya et al. (2011), and Carcillo et al. (2021) document, the peer-group diagnostic value Cyree et al. (2020) establish, and the scale-economy cost argument Hughes and Mester (2013) report, 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 review scores reasonably on explainability and cost feasibility because a reviewer’s reasoning is inherently articulable and requires no new investment, but poorly on detection sophistication, early-warning lead time, and peer-context awareness because it lacks any systematic analytic or comparative capability; standalone in-house machine learning improves detection and lead time modestly but remains weak on peer-context awareness, since a single institution’s model has no access to other institutions’ data, and its cost feasibility for a resource-constrained institution is actually lower than the status quo because building and retaining in-house data science capacity is expensive. CoSAFE is positioned, by design, to lead on every dimension except raw feasibility, where its moderate score reflects a real subscription cost partially, but not entirely, offset by the shared-service savings Hughes and Mester (2013) document.

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

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

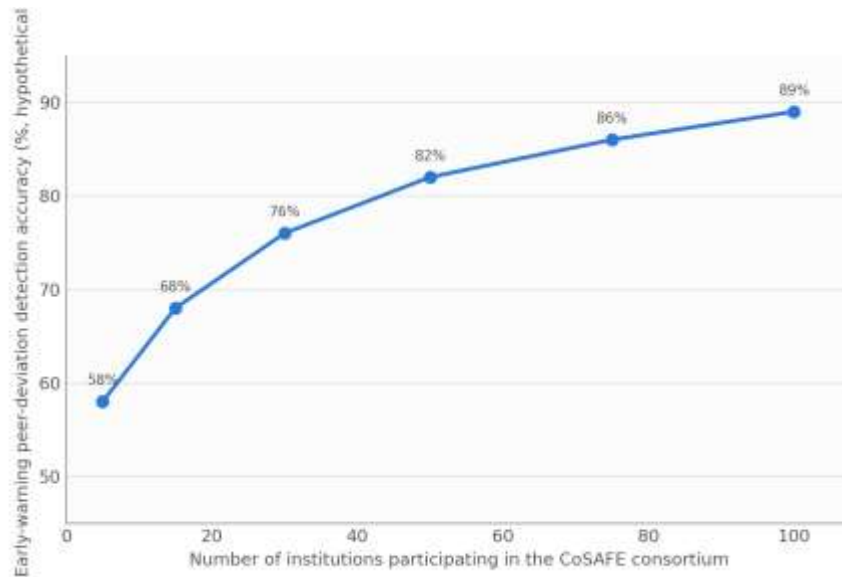


Source: These trajectories are hypothetical operating points, not measurements from a deployed system; the standalone in-house model's rising miss rate reflects the data-scarcity limitation Chandola et al. (2009) describe for a single small institution's thin confirmed-example base, while CoSAFE's declining trajectory reflects the pooled-data advantage the consortium's growing shared history provides, consistent with the peer-context value Cyree et al. (2020) and Betz et al. (2014) document.

The figure shows the two machine learning approaches starting from a similar, substantially improved position relative to manual review, then diverging sharply over the deployment period: the standalone in-house model's miss rate climbs as its single institution's own limited data proves insufficient to keep the model current against evolving fraud patterns and shifting financial conditions, while CoSAFE's miss rate continues to fall as the consortium's pooled data across many institutions accumulates. By month twelve, the standalone model has lost most of its initial advantage over manual review, why this paper treats data pooling, not merely model sophistication, as the architectural feature that determines whether a machine learning approach's advantage is sustained or temporary for an institution too small to generate sufficient data on its own.

4.7 Figure 3: Consortium Size and Early-Warning Peer-Deviation Detection Accuracy

Figure 3 Early-Warning Peer-Deviation Detection Accuracy (%) as CoSAFE Consortium Size Grows

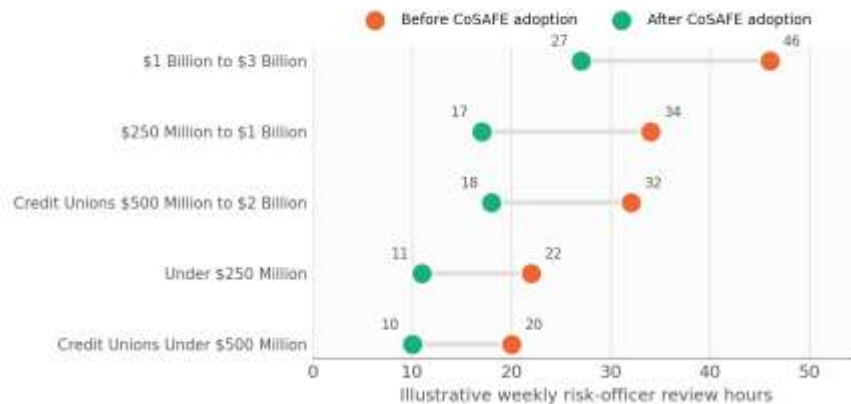


Note: These operating points are hypothetical, not measurements from a deployed system; the improving accuracy as consortium size grows reflects Cyree et al.'s (2020) finding that better-constructed, larger peer groups produce more statistically reliable benchmarks and the scale-economy logic Hughes and Mester (2013) document for shared analytics infrastructure, though the diminishing returns at larger consortium sizes reflect an assumption that peer-group quality, not sheer institution count alone, ultimately bounds the achievable accuracy gain.

This paper's framework treats consortium size as a design variable with real consequences rather than an incidental detail: detection accuracy improves substantially as the consortium grows from a handful of institutions to several dozen, because a larger and more diverse pool both sharpens the statistical peer groups Cyree *et al.* (2020) show matter and supplies more confirmed fraud and distress examples for the shared model library to learn from, but the accuracy gain flattens beyond roughly fifty to seventy-five institutions, suggesting a viable consortium need not achieve national scale to capture most of the achievable benefit. This pattern has a direct practical implication for how such a consortium might realistically be organized: a state or regional bankers' association, a corporate credit union, or a shared-services vendor serving several dozen community institutions is plausibly large enough to approach most of the accuracy gain the figure shows, without requiring the scale of a large national data-pooling arrangement (Potharaju & Sajja, 2024).

4.8 Figure 4: Risk-Officer Review Hours Before and After CoSAFE Adoption, by Institution Asset-Size Tier

Figure 4 Weekly Risk-Officer Review Hours Before and After CoSAFE Adoption, by Institution Asset-Size Tier



Source: These review-hour figures are constructions, not measurements from a deployed system; the roughly proportional reductions across every asset-size tier reflect an assumption that the fraud-prioritization and peer-context gains Sections 4.5 through 4.7 document scale with an institution's existing review volume rather than disproportionately favoring the smallest or largest subscribing institutions.

The figure shows a meaningful weekly review-hour reduction across every institution size and charter type included, consistent with this paper's central claim that CoSAFE's benefits are not confined to the very smallest institutions with the thinnest data, even though those institutions' data-scarcity problem, as Section 4.6, where a standalone in-house model struggles most. Larger community institutions and credit unions in the sample also see substantial hour reductions, reflecting that the peer-benchmarking and shared-detection gains this paper's architecture provides are not solely a remedy for the smallest institutions' data scarcity, but a genuine efficiency improvement over a purely internal review process at any scale within the community-institution range this paper addresses.

5. Discussion

5.1 Implications for Practice

The framework developed here suggests several practical implications for community financial institutions and the associations, corporate credit unions, or vendors positioned to organize a shared analytics consortium. Perhaps most important is that a community institution's data limitations should be treated as an architectural design constraint from the outset rather than an unavoidable ceiling on what analytics such an institution can adopt: Section 4's figures illustrate that a standalone, single-institution model's advantage over manual review erodes within roughly a year precisely because of the data-scarcity problem Chandola *et al.* (2009) describe, while a properly pooled, peer-informed approach does not face the same ceiling. A second implication concerns who is positioned to organize such a consortium: Figure 3's flattening accuracy gain beyond several dozen institutions suggests a regional or state-level bankers' association, corporate credit union, or shared-services

vendor already serving many community institutions is a plausible and realistic organizer, without requiring a national-scale data-pooling arrangement to capture most of the benefit. A third implication is governance: the explainability and per-institution outcome monitoring Sections 2.6 and 4.3 establish as required components should be built into a consortium's founding governance structure from the start, since retrofitting fair-lending-grade explainability or cross-institution disparity monitoring onto an already-operating shared service is a materially harder problem than designing for both from the outset.

5.2 Limitations

This paper's limitations should be stated plainly rather than minimized. No production CoSAFE deployment, real institutional transaction or financial-statement data, or actual fraud or distress 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, several practical obstacles this paper does not resolve would confront any institution or organization attempting to establish a real consortium along these lines. This paper does not address the legal, contractual, and potential antitrust considerations that data-sharing among competing financial institutions would raise in practice, even where the shared data is aggregated and anonymized; these considerations would need to be resolved by the organizing association or vendor's own legal counsel before any real consortium along these lines could operate, and this paper's silence on them should not be read as an assurance that such an arrangement is straightforward to establish. The peer-group construction approach this paper draws on, Cyree *et al.* (2020), was developed and validated using publicly available bank financial-statement data rather than the richer transaction-level data a fraud-detection application would also require, so its direct applicability to peer grouping for fraud-detection purposes specifically, while a plausible extension, has not itself been established in the cited literature. Finally, the scale-economy findings Hughes and Mester (2013) document were estimated from the broader population of U.S. banks across the full size spectrum, not specifically validated for the community-institution size range and shared-service cost structure this paper's architecture assumes, so the specific cost-efficiency magnitude this paper's Table 1 and Figure 1 illustrate should be read as directionally motivated rather than empirically established for this exact context.

5.3 What This Paper Does Not Argue

This paper does not argue that a shared analytics consortium should replace an institution's own risk-management judgment or existing controls; the architecture is intended to supply fraud alerts and early-warning signals a risk officer evaluates and acts on, not to remove human review from a fraud or distress determination that carries real consequences for customers and the institution itself. It does not argue that any specific machine learning algorithm, clustering method, or vendor platform should be adopted; the framework is intentionally technique-agnostic within each stage. It does not argue that CoSAFE eliminates fraud or institutional distress; the framework's contribution is improving detection sensitivity and early-warning lead time relative to the status quo alternatives Section 4 compares it against, not eliminating the underlying risk each community institution faces.

Finally, it does not argue that the specific figures in Section 4 describe any real institution's, consortium's, or vendor'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 shared-services arrangement.

5.4 Future Research

Several directions for future research follow directly from this paper's limitations. Empirical studies measuring actual fraud detection rates, early-warning lead time, and risk-officer review hours at community financial institutions participating in a real shared analytics consortium, whether structured along the lines this paper proposes or another, would test this paper's central claims directly against practice. Research specifically evaluating peer-group construction methods, following Cyree *et al.*'s (2020) approach, when applied to fraud-detection and transaction-level data rather than the financial-statement data their study examines, would sharpen Section 4.1's peer-group construction stage's applicability to this paper's specific domain. Legal and regulatory research examining how a consortium of competing financial institutions can structure data-sharing arrangements that satisfy antitrust, privacy, and safety-and-soundness considerations simultaneously would address the governance gap this paper's own limitations section identifies as most exposed. Finally, research validating whether the scale-economy magnitude Hughes and Mester (2013) document for the broader banking sector actually holds for a shared analytics service sized specifically to the community-institution range this paper addresses would test the cost-efficiency assumption this paper's framework depends on but does not itself empirically establish.

6. Conclusion

Community financial institutions face the same fraud and institutional distress risks large banks face, but with a fraction of the data, staff, and technology budget available to address either risk through machine learning analytics built and maintained in-house. Synthesizing the fraud-detection, anomaly-detection, early-warning, peer-group benchmarking, scale-economies, and explainability-and-fair-lending literatures, this paper developed Community Shared Analytics for Fraud and Early-warning: an architecture in which a shared analytics service pools statistical patterns across a consortium of subscribing community institutions to construct defensible peer groups and maintain a continuously updated fraud-detection and early-warning model library, applies that shared library back to each institution's own data to produce individual fraud alerts and peer-deviation early-warning signals, and attaches an examiner- and board-legible rationale to every output, through two comparative tables, a hypothetical community 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, particularly around the legal and governance questions a real data-sharing consortium would need to resolve, reflects the subject matter honestly: overstating the maturity of a framework with no production precedent, and understating the real organizational work a shared consortium would require, would be a poor foundation for a decision that touches every subscribing institution's own board,

examiners, and customers. What the literature reviewed here does support, and what this paper's synthesis is intended to make actionable, is a considered path forward for the many community financial institutions that have, until now, faced an apparent choice between a status quo they know is analytically thin and a standalone machine learning investment their own size cannot efficiently justify: Community Shared Analytics for Fraud and Early-warning proposes that this choice is false, and that the detection and early-warning gains modern analytics offers can be brought within reach of institutions too small to build them alone, provided the peer-group construction, shared-service governance, and explainability the literature shows are necessary are designed in from the outset rather than assumed to follow automatically from pooling data alone..

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., Islam, C. S., Mudryi, M., & Cakmak, S. (2025). Concept-Drift-Resilient Ensemble Learning for Financial Transaction Fraud Detection Under Evolving Criminal Typologies. *International Journal of Acta Informatica*, 4(1), 52-67.
- [3] 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.
- [4] Rashid, S. U., Ojanen, J., & Kim, J. (2025). Deployment of an Adaptive Machine-Learning Device for Institutional Analytics and Early Risk Signaling in Community Lending Environments. *International Journal of Acta Informatica*, 4(1), 68-79.
- [5] Betz, F., Oprică, S., Peltonen, T. A., & Sarlin, P. (2014). Predicting distress in European banks. *Journal of Banking & Finance*, 45, 225–241.
- [6] 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.
- [7] Rashid, S. U., & Sattiraju, S. R. (2025a). Data-Driven Early-Warning Indicators of Small-Business Financial Distress: Integrating Operational, Transactional, and Credit Data for Decision Support in SMEs. *The Metascience*, 3(4), 53-65.
- [8] Carcillo, F., Le Borgne, Y.-A., Caelen, O., Kessaci, Y., Oblé, F., & Bontempi, G. (2021). Combining unsupervised and supervised learning in credit card fraud detection. *Information Sciences*, 557, 317–331.
- [9] Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys*, 41(3), Article 15.
- [10] Rahman, M. S., Islam, C. S., & Mudryi, M. (2025). Interoperability Standards for Real-Time Fraud-Signal Exchange Across US Financial Institutions: Design Principles and a Reference Implementation. *The Computertech*, 47-60.

- [11] 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.
- [12] Potharaju, B., & Ravipudi, J. (2025). Automated Data Validation and Resilient Infrastructure Design for Enterprise Regulatory Reporting Platforms. *International Journal of Computer Science and Engineering Innovations*, 1(1), 46-64.
- [13] 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.
- [14] Cyree, K. B., Davidson, T., & Stowe, J. D. (2020). Forming appropriate peer groups for bank research: A cluster analysis of bank financial statements. *Journal of Economics and Finance*, 44(2), 211–237.
- [15] Ahmed, K. R., Sufian, M. A., Faruq, O., Kowser, S., Shammah, R. S., & Ahmmed, M. R. (2025, July). Impact of IT product innovation on competitive advantage in the digital economy. In *2025 International Conference on Quantum Photonics, Artificial Intelligence, and Networking (QPAIN)* (pp. 1-6). IEEE.
- [16] 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.
- [17] Ahmed, K. R., Ahmmed, M. R., Kowser, S., Faruq, O., Shammah, R. S., & Sufian, M. A. (2025, July). The Future of AI-Powered SaaS: Market Trends and Investment Opportunities. In *2025 International Conference on Quantum Photonics, Artificial Intelligence, and Networking (QPAIN)* (pp. 1-6). IEEE.
- [18] 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.
- [19] Faruq, O., Shammah, R. S., Morshed, N., Kowser, S., Rahman, I., & Guria, Z. M. (2025, December). Data-Driven Approach to Educational Reforms for Transforming US Higher Education through AI and Business Analytics. In *2025 International Conference on AI-Driven STEM Education and Learning Technologies (AISTEMEDU)* (pp. 1-9). IEEE.
- [20] Rahman, M. S., Kabir, A. A., Reon, M. A. H., & Rashid, S. U. (2025). From Financial Transactions to Clinical Records: A Unified Machine Learning Methodology for Real-Time Anomaly Detection and Institutional Decision Support In High-Stakes Domains. *International Journal of Information Technology and Management Information Systems (IJITMIS)*, 16(6), 23-42.

- [21] Rashid, S. U., & Sattiraju, S. R. (2025b). Multi-Institution Validation of Drift-Resilient Early-Warning Models for SME Loan Distress and Transaction Fraud: Evidence from US Community Financial Institution Pilots. *International Journal of Modern Computing*, 8(1), 40-54.
- [22] Rahman, M. S., Marcucci, M., & Møller, C. (2025). Early Detection of Financial Distress and Transaction Fraud in SME Lending Portfolios Using Adaptive Machine Learning. *The Metascience*, 3(4), 35-52.
- [23] Karim, N. B., Kovac, T., & Song-Chun, Z. (2025a). Model Risk Management for Adaptive Credit Early-Warning Systems in Supervised SmallBusiness Lending: Aligning Machine-Learning Recalibration with Federal Model-Risk Guidance and the NIST AI Risk Management Framework. *International Journal of Acta Informatica*, 4(1), 80-94.
- [24] Gunther, J. W., & Moore, R. R. (2003). Early warning models in real time. *Journal of Banking & Finance*, 27(10), 1979–2001.
- [25] Hughes, J. P., & Mester, L. J. (2013). Who said large banks don't experience scale economies? Evidence from a risk-return-driven cost function. *Journal of Financial Intermediation*, 22(4), 559–585.
- [26] Karim, N. B., & Kovac, T. (2025). Cloud-Based Big-Data Architectures for AI-Driven Financial Planning and Analysis and Early Risk Signaling in Resource-Constrained Community Financial Institutions. *The Metascience*, 3(3), 29-40.
- [27] 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.
- [28] Karim, N. B., Tolosana, D., & Simko, R. (2025b). Responsible and Legally Compliant Enterprise Artificial Intelligence in Small-Business Lending: Governance, Explainability, and Fair-Lending Safeguards for Early-Intervention Recommendation Engines. *International Journal of Modern Computing*, 8(1), 55-66.
- [29] 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.
- [30] West, J., & Bhattacharya, M. (2016). Intelligent financial fraud detection: A comprehensive review. *Computers & Security*, 57, 47–66.
- [31] 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. Potharaju & Ravipudi, 2023).

- [32]Potharaju, B., & Sajja, R. R. (2024). Modernizing US Federal Regulatory Reporting Infrastructure for Scalability, Resilience, and Audit Readiness. *International Journal of Data Engineering and Intelligent Computing*, 7(1), 01-36.
- [33]Rahman, M. M., Computer, M. M. R., Das, N., Uddin, M. S., Shammah, R. S., & Rahman, M. M. (2025, May). An Investigation of Face Detection and Feature Extraction Techniques in Deepfake Detection Using a Multi-Model Machine Learning Approach. In *2025 3rd International Conference on Business Analytics for Technology and Security (ICBATS)* (pp. 1-8). IEEE.
- [34]Ahmed, K. R., Shammah, R. S., Kowser, S., Faruq, O., Sufian, M. A., & Ahmmed, M. R. (2025, July). Strengthening Digital Security in MIS: A Business Analytics Approach to Deepfake Detection. In *2025 International Conference on Quantum Photonics, Artificial Intelligence, and Networking (QPAIN)* (pp. 1-6). IEEE.
- [35]Guria, Z. M., Rahman, I., Kowser, S., Morshed, N., Shammah, R. S., & Faruq, O. (2025, December). AI-driven risk management and fraud prevention in digital finance as a catalyst for US economic resilience and global financial leadership. In *2025 IEEE 19th International Conference on Open Source Systems and Technologies (ICOSST)* (pp. 1-8). IEEE.
- [36]Shammah, R. S., Rahman, I., Guria, Z. M., Kowser, S., Faruq, O., & Morshed, N. (2025, December). Scalable Machine Learning Architecture for Securing IoT Networks with Digital Governance in the AI Era of Cyber-Resilient Smart Cities. In *2025 International Conference on AI-Driven STEM Education and Learning Technologies (AISTEMEDU)* (pp. 1-7). IEEE.