INTERNATIONAL JOURNAL OF ACTA INFORMATICA
VOLUME (2024)

ADAPTIVE MACHINE-LEARNING DETECTION AND PREVENTION OF FINANCIAL-TRANSACTIONS FRAUD AT COMMUNITY FINANCIAL INSTITUTIONS: RECALIBRATING RISK SIGNALS UNDER SHIFTING FRAUD TYPOLOGIES AND RECORD SUSPICIOUS-ACTIVITY MONITORING BURDENS

Naima Bintay Karim[1*], Raj Kumar Pentyala[2], Rafael Simko[3]

[1]Arkansas State University, USA

[2] Financial Analytics, JP Morgan Chase, UNITED STATES

[3] University of Zaragoza, SPAIN

*Corresponding Author Email: naimabintay980@gmail.com

ABSTRACT

Community financial institutions are facing a simultaneous rise in fraud incidence and in the regulatory monitoring burden that accompanies it. U.S. Suspicious Activity Report (SAR) filings reached a record 3.6 million across all filer groups in 2022, up 18% year-over-year, with banks, savings associations, and credit unions accounting for roughly half of that volume [1]. Against this backdrop, industry benchmark surveys found that 79% of credit union and community bank leaders reported direct fraud losses exceeding $500,000 in 2023, a larger share than at midsize or large financial institutions [7]. This article examines what adaptive machine-learning approaches have demonstrated for fraud detection accuracy, why static, rule-based detection struggles against shifting fraud typologies such as synthetic identity fraud and authorized push payment (APP) fraud, and what recalibration and governance practices community financial institutions need to sustain detection performance as fraud tactics evolve faster than manual rule updates can track. Synthetic identity fraud in which a fabricated identity built from a mix of real and invented data passes standard verification and is cultivated over months or years before exploitation has grown U.S. unsecured-credit losses from $1.80 billion in 2020 to a projected $2.42 billion in 2023, a roughly 10% compound annual growth rate [3-5]. Published fraud-detection studies report ensemble machine-learning models achieving accuracy in the 0.91–0.94 range on realistic, class-imbalanced transaction data [12], while models trained and evaluated without adequately accounting for severe class imbalance can report accuracy figures above 99% that are structurally misleading given how rare fraud is relative to legitimate transactions [13][14]. This article synthesises this evidence into a practical framework for recalibrating fraud-detection models under shifting typologies while managing a suspicious-activity monitoring workload that is growing faster than most community institutions' compliance staffing.

KEYWORDS: Fraud Detection; ML; Community Banks; Credit Unions; Concept Drift; Model Recalibration; AML; Adaptive Learning

INTRODUCTION

This article draws on Bank Secrecy Act (BSA) filing statistics, applied machine-learning fraud-detection research, and community-financial-institution fraud-survey data to examine a problem that is simultaneously technical and operational: fraud typologies are shifting faster than static, rule-based detection systems can be manually updated, while the volume of suspicious-activity monitoring and reporting these institutions must sustain has reached a record high. Each claim in this article is tied to a specific source rather than presented as general industry consensus.

The article is organised to move from problem scale to solution architecture. Sections 2 and 3 establish the size of the compliance burden and the specific fraud typologies driving it. Section 4 reviews what machine-learning research has demonstrated for detection accuracy, and Section 5 addresses the harder, less commonly discussed problem of sustaining that accuracy as fraud tactics shift. Sections 6 through 9 address the operational, governance, and resourcing realities specific to community-scale institutions, and Sections 10 through 12 consolidate limitations, recommendations, and conclusions.

The scale of the reporting burden alone is substantial. U.S. Suspicious Activity Report filings reached a record 3.6 million across all seven filer categories in 2022, an increase of 18% over 2021, with banks, savings associations, and credit unions accounting for roughly half of these reports [1]. Figure 1 shows this year-over-year growth. Since mandatory SAR filing began in 1996, well over 30 million SARs have been filed in total, with a substantial majority filed by depository institutions specifically [1].

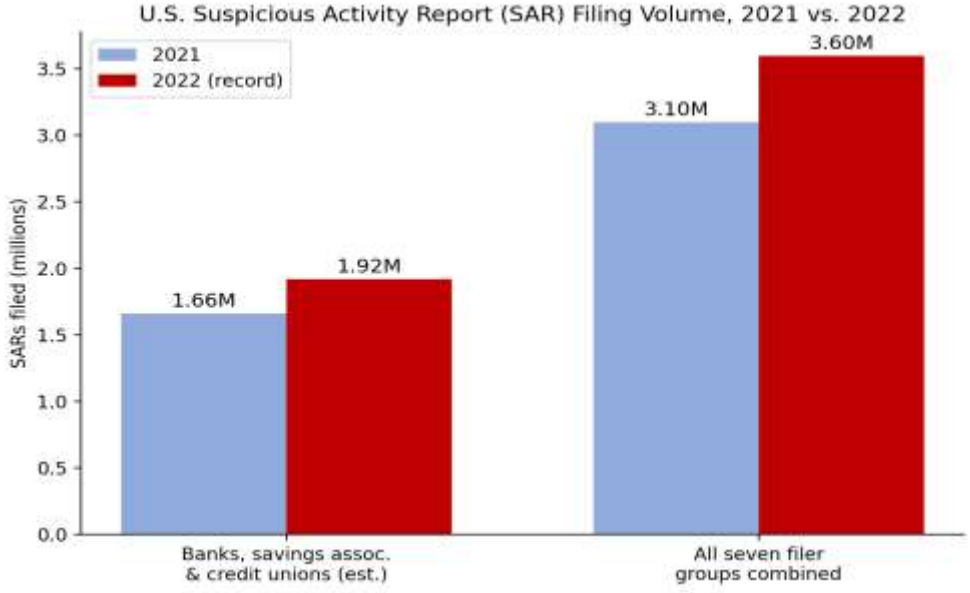

Figure 1. U.S. Suspicious Activity Report (SAR) filing volume, 2021 vs. 2022 [1].

This growth in filing volume is not simply an administrative artifact; a separate analysis of FinCEN's staffing found that in the record 2022 year of 3.6 million SARs, the agency's roughly 300 staff — even generously assuming half were fully dedicated to SAR processing — would need to review approximately 66 reports per analyst per day to keep pace, a workload the analysis characterises as unrealistic to sustain through manual review alone [8]. This structural mismatch between filing volume and review capacity is precisely the condition under which better-targeted, adaptive detection at the point of filing — reducing false-positive-driven over-filing while improving true-positive capture — has the greatest potential value.

THE FRAUD AND COMPLIANCE BURDEN AT COMMUNITY FINANCIAL INSTITUTIONS

Table 1 consolidates the key statistics establishing the scale of the current fraud and compliance burden specifically at community-scale institutions. Industry benchmark reporting found that 79% of credit union and community bank leaders reported direct fraud losses exceeding $500,000 in 2023 — a larger share than at midsize or large financial institutions, underlining the disproportionate vulnerability of smaller institutions with thinner fraud-technology budgets [7]. Figure 2 visualises these findings.

Table 1. Key statistics on fraud and compliance burden at community financial institutions.

Statistic	Value	Source
Total SARs filed, 2022 (all seven filer groups)	3.6 million (+18% vs. 2021)	FinCEN, via Thomson Reuters [1]
SARs filed by banks, savings associations, credit unions, 2022	Roughly half of total volume	FinCEN, via Thomson Reuters [1]
Cumulative SARs filed since 1996 (26-year total)	Well over 30 million (majority by depository institutions)	FinCEN, via Thomson Reuters [1]
Credit unions and community banks with fraud losses exceeding $500,000, 2023	79% of surveyed institutions	Alloy State of Fraud Benchmark Report [7]
Credit unions/community banks planning to invest in an identity-risk solution, 2023	88% of surveyed institutions	Alloy State of Fraud Benchmark Report [7]

U.S. consumer fraud losses, 2023	More than $10 billion (+14% vs. 2022)	Alloy State of Fraud Benchmark Report [7]

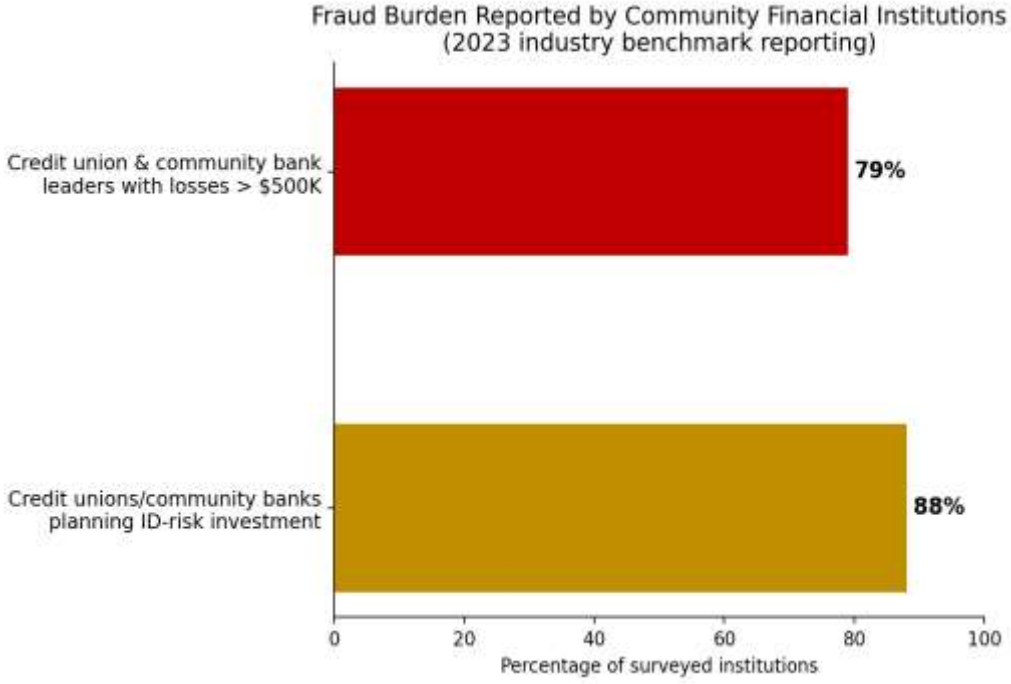

Figure 2. Fraud burden reported by community financial institutions, 2023 industry benchmark reporting [7].

The most frequently reported SAR category among banks, savings associations, and credit unions in 2022 was "Check Fraud," which overtook other categories as check-fraud SAR volume roughly doubled from approximately 350,000 in 2021 to approximately 683,000 in 2022; other frequently reported categories included suspicion concerning the source of funds, transactions with no apparent lawful purpose, transactions below the CTR threshold, and suspicious EFTs/wire transfers [1]. Figure 3 shows this category breakdown, which is a useful diagnostic: the dominance of check fraud is a useful reminder that established, non-AI-driven fraud typologies remain a major share of filings even as newer typologies draw more industry attention, while the remaining categories reflect transaction patterns that are, by design, exactly the kind of behavioural anomaly machine-learning models are well suited to flag, as distinct from simple threshold-based rules.

This diagnostic value cuts both ways, however. A SAR category built around "no apparent lawful purpose" is, by construction, a judgment call rather than a hard threshold, meaning its high filing volume partly reflects genuine behavioural-anomaly detection and partly reflects defensive filing — institutions filing a SAR when uncertain, precisely because the cost of under-filing (regulatory criticism, potential

enforcement action) is asymmetrically higher than the cost of over-filing a report that turns out not to indicate genuine criminal activity. This asymmetry is itself part of why SAR volume has grown alongside, and arguably faster than, actual fraud incidence, and is a structural factor any institution evaluating its own filing patterns against the industry figures in Table 1 should keep in view.

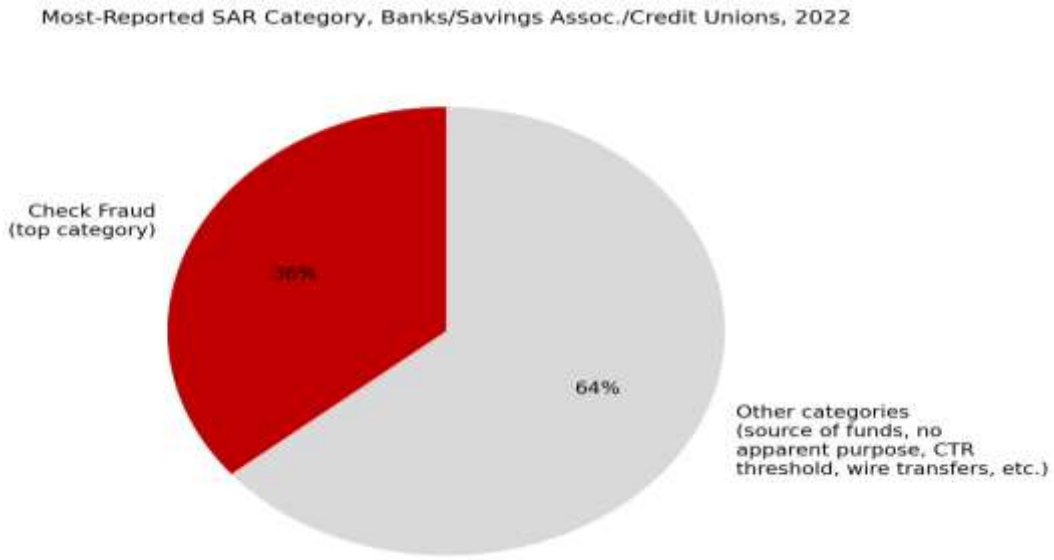

Figure 3. Most-reported SAR categories, banks/savings associations/credit unions, 2022 [1].

SHIFTING FRAUD TYPOLOGIES

Static, rule-based fraud detection is built around known patterns: a transaction above a threshold, a mismatch between billing and shipping address, a rapid sequence of small transactions. These rules work well against fraud typologies that were common when the rules were written, but degrade as fraud tactics shift — a phenomenon closely related to what the machine-learning literature calls concept drift, where the statistical relationship between input features and the fraud/non-fraud label changes over time. Table 2 summarises four fraud typologies that have grown fastest in recent reporting and explains why each specifically strains legacy rule-based detection.

Table 2. Shifting fraud typologies and their strain on legacy rule-based detection.

Fraud typology	Reported scale	Why it strains legacy rule-based detection
Synthetic identity fraud	$2.42B projected for 2023 U.S. unsecured-credit losses, ~10% annual growth [3][4][5]	Fabricated identities pass standard verification and build credit history over months or years before exploitation [10]

Bust-out fraud	21% of fraud cases, 16% of total losses [2]	Often built on synthetic or counterfeit identities that appear creditworthy until the account is deliberately maxed out [11]
Authorized Push Payment (APP) fraud	17% of incidents, 16% of losses [2]; consumer losses growing ~20% YoY [6]	Victim authorizes the transfer themselves, so transaction-level rules see a legitimate, authenticated payment
AI-enabled/deepfake-assisted fraud	Early-stage but growing concern following high-profile AI-voice-cloning fraud cases [2][3]	Generative tools defeat static document- and voice-verification rules faster than those rules can be updated manually

Synthetic identity fraud is the clearest illustration of this dynamic. Unlike traditional identity theft, which uses a real, stolen identity in ways that can eventually be flagged by the genuine identity-holder noticing suspicious activity, a synthetic identity combines real information (such as a stolen Social Security number) with fabricated details (a fictitious name, invented employment history) to create an identity that does not correspond to any actual person and can pass standard verification checks [10]. Because this fabricated identity is often cultivated over months or years — used responsibly at first to build a credit history before being exploited in a bust-out scheme — a rule-based system evaluating each transaction or application in isolation, without a longitudinal view of the identity's behaviour over time, has no natural trigger point at which to flag it. Figure 4 shows the growth trajectory of reported synthetic identity fraud losses.

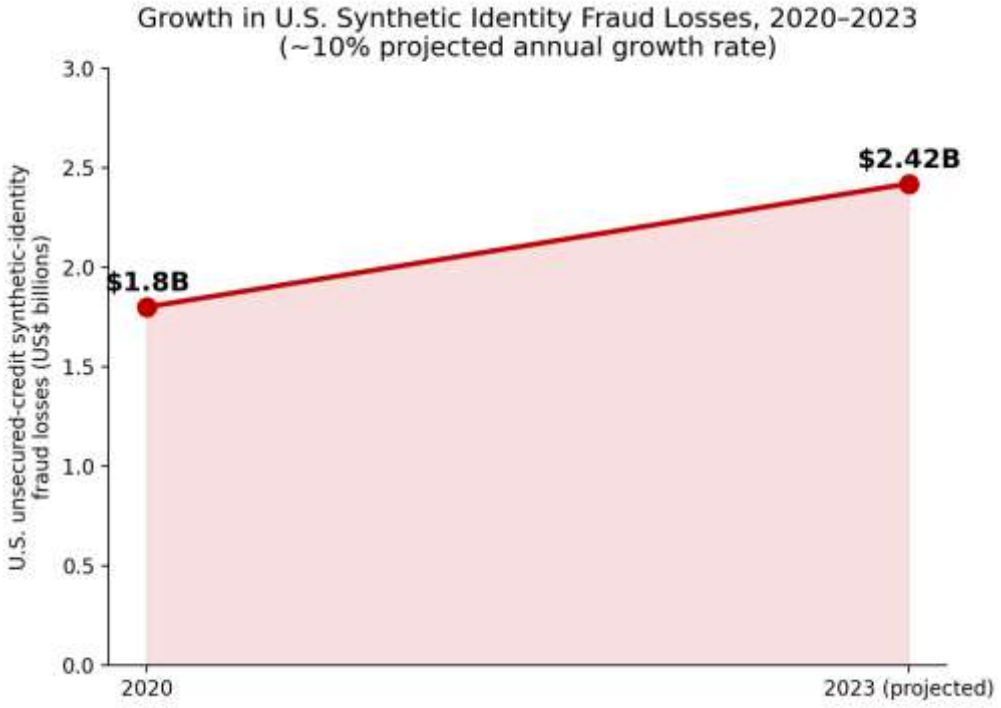

Figure 4. Growth in U.S. synthetic identity fraud losses (unsecured credit), 2020–2023 (2023 projected) [4][5].

Authorized Push Payment fraud presents a structurally different challenge: because the account holder themselves authorizes the transfer — typically after being deceived by a scammer into believing the payment is legitimate — the transaction itself looks exactly like a normal, authenticated payment from the perspective of transaction-level rules, which is precisely why it now accounts for 17% of fraud incidents and 16% of total losses in recent industry data [2]. Figure 5 compares bust-out fraud and APP fraud on incidence and loss share, showing that both typologies contribute a share of total losses roughly proportional to, or in APP fraud's case exceeding proportionally, their share of incidents — indicating neither is a low-value nuisance category that can be deprioritised.

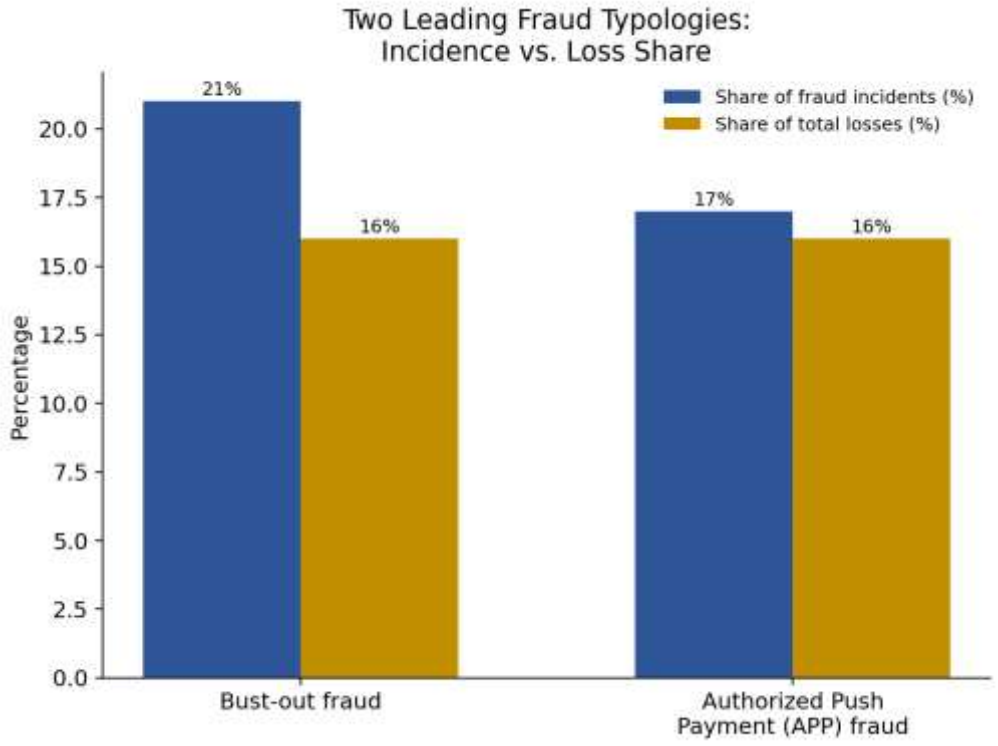

Figure 5. Bust-out fraud and Authorized Push Payment fraud: share of incidents vs. share of total losses [2].

Generative AI is an early but growing accelerant for both typologies. Industry commentary following high-profile AI-voice-cloning and social-engineering fraud cases has flagged generative tools as an emerging identity-fraud concern, even though comprehensive industry-wide measurement of AI-enabled fraud's share of losses is still immature relative to established typologies [2][3]. AI-generated synthetic documents, cloned voices for social-engineering-driven APP scams, and automated, scaled account-creation attempts all share a common implication for detection strategy: static rules calibrated against last year's fraud patterns can be defeated faster than periodic manual rule updates can track, which argues for more frequent, ideally automated, recalibration.

It is also worth noting that older, non-AI-driven fraud tactics have not disappeared even as AI-enabled tactics grow; check-fraud SAR volume itself roughly doubled between 2021 and 2022, per Section 3's category breakdown [4]. This matters for detection strategy because it means community institutions cannot simply redirect all detection investment toward AI-era typologies at the expense of maintaining coverage for established fraud categories like check fraud, which — per the SAR category breakdown in Figure 3 — remains one of the most frequently reported categories in its own right. The practical implication is that a recalibration programme (Section 5) needs

to expand detection coverage for emerging typologies while maintaining, not deprioritising, coverage for established ones.

MACHINE-LEARNING APPROACHES TO FRAUD DETECTION

Applied research on machine-learning fraud detection spans traditional classifiers, ensemble methods, and deep-learning architectures, with reported performance summarised in Table 3. A comparative study of online fraud detection found the Voting Classifier, SVM, and Random Forest achieving the highest accuracy at 0.94, with Stacking and Logistic Regression close behind at 0.93, and XGBoost at a still-strong 0.91 [12]. On precision specifically — the metric that determines how often a flagged transaction is actually fraudulent, which directly drives the false-positive burden on compliance staff — the Voting Classifier and Stacking models achieved 0.99 and 0.98 respectively [12].

Table 3. Reported machine-learning fraud-detection model performance across published studies.

Model / study	Reported performance	Dataset / context
Voting Classifier / SVM / Random Forest	0.94 accuracy; 0.99 precision (Voting Classifier)	Online fraud-detection benchmark comparison [12]
Stacking ensemble / Logistic Regression	0.93 accuracy	Same benchmark comparison [12]
XGBoost	0.91 accuracy	Same benchmark comparison [12]
TPOT + Random Forest + Tomek links resampling	99.99% accuracy (imbalanced-data caveat applies)	CreditCardFraud public dataset, 492 fraud / 284,315 legitimate [14]

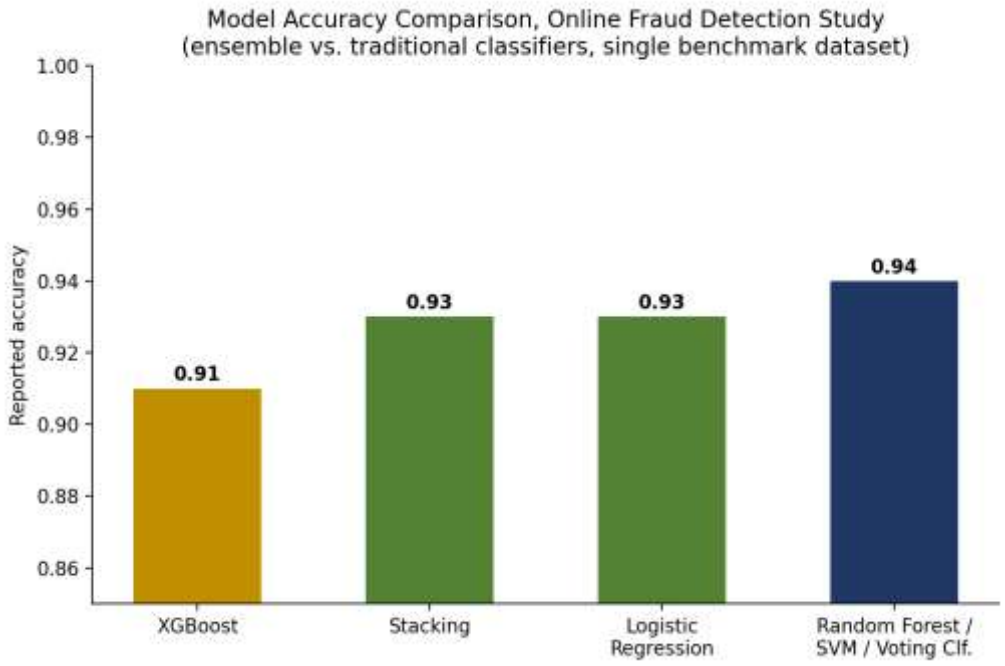

Figure 6. Model accuracy comparison, online fraud-detection benchmark study [12].

A separate study using the well-known CreditCardFraud public dataset — a highly imbalanced set of 284,315 legitimate and only 492 fraudulent transactions reported 99.99% accuracy using a TPOT-optimised Random Forest model combined with Tomek links resampling to address class imbalance [14]. This figure requires careful interpretation: on a dataset where fraud represents roughly 0.17% of all transactions, a trivial model that classified every transaction as legitimate would already achieve 99.83% accuracy while catching zero fraud, which is precisely why accuracy alone is described in the broader methodological literature as a misleading metric for imbalanced fraud classification, and why precision, recall, and F1-score not accuracy in isolation are the metrics that actually determine whether a fraud-detection model is operationally useful [13]. The reported 99.99% figure should be read alongside the study's precision and recall results specifically, not treated as a standalone headline number.

Beyond traditional supervised classifiers, more recent research applies deep-learning architectures convolutional and recurrent neural networks, autoencoders, and graph neural networks that can capture non-linear and relational patterns rule-based and even traditional ML systems miss. Graph-based approaches in particular represent account holders, devices, and counterparties as connected nodes, which is directly relevant to bust-out and synthetic-identity fraud, since these schemes typically involve networks of related fabricated identities and mule accounts rather than a single isolated bad actor, and a graph-aware model can surface these network-level connections that a

transaction-by-transaction rule engine structurally cannot see [9][11].

A separate methodological point emerging from this literature deserves emphasis for community-institution deployments specifically: several studies note that ensemble methods combining multiple model types voting or stacking classifiers that pool predictions from several base models consistently rank among the top performers on both accuracy and precision, per Table 3, without requiring the specialised infrastructure that deep-learning or graph-based architectures demand [12][15]. For an institution deciding where to start, this suggests ensembling existing, well-understood classifiers is frequently a higher-return investment than jumping directly to a more exotic architecture, a point developed further in the architecture-selection guidance in Section 8.

A comparative banking-transaction study working with a sample of 141,174 bank transfers of which 141,079 were legitimate and only 95 fraudulent, an even more extreme imbalance than the CreditCardFraud dataset discussed above — similarly found boosting-based ensemble methods such as AdaBoost and Gradient Boosting particularly effective in this imbalanced setting, aggregating multiple weak classifiers to build a stronger overall model that proved more resilient to class imbalance than single-classifier approaches, while other tested methods (KNN, SVM, and neural networks) also achieved strong, though somewhat less favourable, results [15]. This additional data point reinforces the pattern from Table 3: across multiple independent studies using different datasets, ensemble and boosting methods consistently outperform single-classifier baselines specifically on the imbalanced-data conditions that characterise real-world fraud detection, which is the single most consistent finding across the fraud-detection literature reviewed for this article.

ADAPTIVE RECALIBRATION UNDER CONCEPT DRIFT

The central technical challenge this article addresses is not simply choosing an accurate model at deployment, but sustaining that accuracy as fraud typologies shift — a requirement that goes beyond standard model validation practice. A model trained on a snapshot of historical transaction and fraud-label data implicitly assumes the relationships in that data remain stable going forward; the growth trajectory in Figure 4 and the rapid AI-driven typology shifts described in Section 3 both indicate this stability assumption breaks down within a timeframe materially shorter than the annual or multi-year retraining cycles common in less mature fraud-detection deployments.

Three practical recalibration disciplines follow from this. First, a defined, ideally automated, model-monitoring cadence should track precision and recall separately and continuously, rather than only accuracy, since a drift in false-positive or false-negative rate can occur well before overall accuracy visibly degrades, particularly on the heavily imbalanced data typical of fraud detection (Section 4). Second, retraining triggers

should be tied to detected performance degradation or to known typology shifts (such as a documented rise in a specific fraud category) rather than to a fixed calendar schedule alone, since the AI-driven typology shifts described in Section 3 are not occurring on a convenient annual cycle. Third, feature sets need periodic review as new signal types become available or as previously predictive features lose their discriminative power — for example, as verification checks that once reliably distinguished synthetic from genuine identities are increasingly defeated by AI-generated supporting documentation [2][3], features built around those checks need to be supplemented or replaced rather than assumed permanently reliable.

Table 4 sets out a practical monitoring framework operationalising these three disciplines, tying specific tracked metrics to the recalibration action each should trigger.

Table 4. Recalibration monitoring framework for adaptive fraud-detection models.

Monitoring metric	What it tracks	Recalibration trigger
Precision (rolling window)	Share of flagged transactions that are genuinely fraudulent	Sustained decline signals rising false-positive burden on compliance staff
Recall (rolling window, where ground truth available)	Share of actual fraud the model successfully flags	Sustained decline signals a fraud typology the model no longer detects
Feature drift indicators	Statistical distance between current and training-period feature distributions	Significant drift signals the model's training data no longer represents current behaviour
Emerging-typology alerts	Qualitative signals from industry advisories, FinCEN alerts, and peer-institution reporting	A documented new typology (e.g., a FinCEN alert) triggers proactive feature/rule review even absent measured drift

A useful discipline for smaller institutions without dedicated data-science staff is to treat the fourth row of Table 4 emerging-typology alerts as the lowest-cost, highest-leverage monitoring input available, since it requires no in-house statistical monitoring infrastructure, only a documented process for reviewing FinCEN advisories and peer-institution reporting and translating a relevant new alert into a concrete review of the institution's own feature set and rules, following the community-bank-specific guidance

already published on typologies such as synthetic identity fraud [10].

BALANCING DETECTION SENSITIVITY AGAINST THE SUSPICIOUS-ACTIVITY MONITORING BURDEN

Improving fraud-detection recall catching more true fraud — is only half of the practical challenge community institutions face; the other half is managing the false-positive volume that any sufficiently sensitive detection system generates, given the direct connection between flagged transactions and the SAR-filing burden documented in Sections 1 and 2. A detection system tuned purely to maximise recall, without corresponding attention to precision, risks materially increasing an already-record SAR volume with reports that add compliance-staff workload without adding investigative value, which following the staffing-versus-volume mismatch described in Section 1 a resource-constrained community institution is poorly positioned to absorb.

This creates a specific optimisation problem for community-institution deployments that differs from the pure detection-accuracy problem addressed in most published fraud-detection benchmarks: the objective is not simply the highest achievable recall or the highest achievable precision in isolation, but the operating point on the precision-recall trade-off curve that a given institution's investigative staffing can actually sustain. A large money-center bank with a large compliance team can reasonably operate at a more recall-heavy, higher-false-positive-tolerant setting than a community bank or credit union with a handful of BSA/AML analysts, even if both institutions are using the same underlying model architecture.

A practical way to operationalise this institution-specific optimum is to work backward from actual investigative capacity rather than forward from a generic target precision or recall figure. If an institution's BSA/AML team can realistically investigate a fixed number of alerts per analyst per day without the kind of unsustainable workload FinCEN's own staff face at national scale (Section 1), that capacity figure — not an abstract accuracy target — should set the model's alert-volume threshold, with the model's precision then tuned to maximise recall subject to that volume constraint rather than the reverse. This reframes model tuning as a capacity-planning exercise as much as a statistical one, which is a departure from how most published fraud-detection benchmarks report results, since those benchmarks typically optimise for the accuracy metrics in Table 3 without reference to any specific institution's investigative throughput.

GOVERNANCE AND REGULATORY CONSIDERATIONS

Fraud-detection models at a regulated financial institution sit within an existing Bank Secrecy Act and anti-money-laundering compliance framework, and adaptive, continuously recalibrated models introduce governance questions that static,

infrequently-updated rule sets do not. A model that is retrained or recalibrated in response to observed performance drift, per Section 5, needs a documented process for validating that each recalibration genuinely improves detection rather than simply overfitting to recent data, together with an audit trail a regulator can review to understand why the model's behaviour changed at a given point in time. This is a materially different governance burden than validating a static rule set once at deployment and revisiting it only periodically.

This audit-trail requirement is not merely a best practice but a practical necessity given how BSA examinations work: an examiner reviewing an institution's fraud and AML program will typically ask not only whether current detection performance is adequate, but why the model behaves as it does and how the institution knows a given change improved rather than degraded detection. An institution that cannot answer these questions for an adaptively recalibrated model — because recalibration happened without a documented before/after comparison — faces materially greater examination risk than one using a simpler, static rule set whose behaviour, however less accurate, is at least fully explainable at every point in time. This is a genuine trade-off, not a reason to avoid adaptive models, but institutions should budget for the documentation and validation overhead alongside the modelling work itself.

A second governance consideration follows directly from the fraud-and-AML convergence trend noted in current industry reporting: fraud risk, credit risk, and AML monitoring are often owned by separate teams with different tools and success metrics within the same institution, which allows a synthetic identity — spanning application fraud, ongoing account behaviour, and eventual bust-out — to persist undetected as it moves across these organisational boundaries [6]. Addressing this specifically requires shared ownership of identity-risk signals across the customer lifecycle rather than point-in-time checks owned by a single team, a structural recommendation that applies with particular force to smaller institutions where the same limited staff often already wear multiple compliance hats and could, in principle, share this signal more easily than a large institution's more siloed functions — provided the underlying detection tooling itself is integrated rather than fragmented across separate fraud, credit, and AML systems.

A third governance consideration concerns model explainability specifically in the fair-lending and consumer-protection context, distinct from the BSA/AML examination context discussed above. Where a fraud-detection flag leads to an account restriction, a declined transaction, or a closed account, the affected customer may be entitled to an explanation under applicable consumer-protection frameworks, which means the same explainability standard developed for AI-assisted credit decisions in community lending applies with comparable force here: a fraud model's flag needs to be traceable

to specific, articulable reasons, not merely a high-confidence score from an opaque model, regardless of how strong that model's aggregate precision and recall figures are.

IMPLEMENTATION CONSIDERATIONS FOR RESOURCE-CONSTRAINED INSTITUTIONS

The precision/recall balancing problem in Section 6 and the governance requirements in Section 7 both point toward the same practical conclusion for community institutions with limited data-science and compliance capacity: model sophistication should not be pursued for its own sake, and a well-governed, well-monitored ensemble model of the kind reported in Table 3 — rather than a more complex, harder-to-explain deep-learning or graph-based architecture — is often the more appropriate starting point, with more advanced architectures reserved for specific typologies (such as network-aware detection for bust-out rings) where the added complexity is justified by a documented gap in the simpler model's coverage. This mirrors the same maintainability-versus-marginal-performance trade-off that applies to AI-assisted early-warning systems in the community-lending context more broadly.

Table 5 sets out this architecture-selection logic more explicitly, mapping the three broad architecture families discussed in this article to the fraud typology each is best suited for and the resource commitment each realistically requires, to help a resource-constrained institution prioritise where added model complexity is actually justified rather than adopting complexity uniformly across its entire detection stack.

Table 5. Fraud-detection architecture selection by typology and resource requirement.

Architecture	Best-suited typology	Resource requirement
Tree-based ensembles (Random Forest, XGBoost, Voting/Stacking)	General transaction-level fraud; strong default choice	Low-to-moderate; widely supported tooling
Graph neural networks	Bust-out rings, mule-account networks, synthetic-identity clusters	High; requires relationship data and specialised infrastructure
Autoencoders / anomaly detection	Novel typologies with no labelled fraud examples yet	Moderate; useful complement where labelled data is scarce

FUTURE DIRECTIONS

Three developments are likely to shape adaptive fraud detection at community financial

institutions over the coming several years. First, as the fraud-and-AML convergence trend noted in Section 7 matures, expect growing vendor and platform support for genuinely unified fraud/credit/AML signal-sharing, reducing the organisational-silo problem that currently allows synthetic identities to persist undetected across functions. Second, as generative-AI-enabled fraud tactics continue to accelerate already flagged as an emerging concern in industry commentary [2] detection research is likely to shift further toward behavioural and relationship-based signals (of the kind graph-based architectures capture) that are inherently harder for a generated, synthetic identity to fake than a single static document or verification check. Third, growing SAR volume (Section 1) combined with continued FinCEN staffing constraints is likely to increase regulatory and industry pressure toward AI-assisted SAR narrative drafting and triage tools that help compliance teams prioritise the highest-value reports for human review, rather than treating every generated alert as requiring equal manual attention.

A fourth development, specific to the community-institution segment this article focuses on, is the likely emergence of shared-infrastructure or managed-service fraud-detection offerings purpose-built for smaller institutions' resource constraints. Just as core-banking-platform vendors already serve many community banks and credit unions with shared infrastructure rather than requiring each institution to build and maintain its own systems, a comparable managed-service model for adaptive fraud detection — where the recalibration discipline of Section 5 and the monitoring framework of Table 4 are maintained by a shared vendor across many participating institutions, each contributing anonymised training signal would directly address the data-scale limitation discussed in Section 11 while avoiding the in-house data-science staffing burden discussed in Section 8. Whether such offerings mature quickly enough to keep pace with the typology-shift rate documented in Section 3 remains an open question this article's evidence base cannot resolve, but the direction of demand — smaller institutions facing enterprise-grade fraud sophistication without enterprise-grade compliance staffing — makes some form of shared-infrastructure response likely.

LIMITATIONS OF THE EVIDENCE BASE

Several limitations of the evidence assembled here should be stated plainly. The model-performance figures in Table 3 are drawn from studies using different underlying datasets, fraud definitions, and class-imbalance-handling techniques, which are not directly comparable to one another or necessarily representative of any specific community institution's actual transaction mix; the 99.99% accuracy figure in particular illustrates why a single headline metric, taken out of context, can materially mislead rather than inform a deployment decision. The community-financial-institution fraud-burden statistics in Section 2 come from industry benchmark survey reporting rather than a single audited source; while these surveys draw on hundreds of fraud decision-

makers, they have not been independently replicated by a study of comparable scope using a different methodology, and self-reported survey data on fraud losses may be subject to reporting or recall bias in either direction. Finally, the synthetic-identity-fraud loss figures in Section 3 are specific to U.S. unsecured-credit losses and should not be extrapolated to represent total synthetic-identity fraud losses across all product types and channels without qualification.

A further limitation concerns the pace of change in the underlying phenomenon this article describes. Fraud typology and AI-capability trends, by the evidence's own account, are shifting within single-year windows the 20-point year-over-year jump in institutions reporting rising fraud (Section 2) is itself an illustration of how quickly baseline figures can move. Readers applying this article's specific percentages and dollar figures at a materially later date should treat the qualitative patterns described (typology shift outpacing static rules, the precision-recall capacity trade-off, the value of shared fraud/credit/AML ownership) as more durable than the specific point-in-time statistics used to illustrate them, and should verify whether more recent data has superseded the figures cited in Tables 1 and 2 specifically.

REGIONAL AND CONSORTIUM-BASED APPROACHES

An option available to community financial institutions that is under-discussed in the vendor-facing fraud-detection literature is consortium-based or regional data-sharing for fraud-model training and signal exchange. Because individual community institutions each see only their own transaction volume, a single small institution's own historical fraud examples may be too sparse to train or validate a model with the confidence larger institutions can bring to bear on the same architecture. Information-sharing frameworks that allow participating institutions to pool anonymised fraud-pattern signals while respecting applicable privacy and competitive-sensitivity constraints can partially offset this data-scale disadvantage, giving a community bank or credit union access to a broader, more representative training signal than its own transaction history alone could provide.

This is particularly relevant for the emerging-typology-alert channel described in Table 4: a regional consortium or information-sharing association is well positioned to aggregate and distribute exactly the kind of qualitative, fast-moving typology intelligence (a new bust-out pattern observed at one member institution, an emerging synthetic-identity ring operating across several institutions in a region) that no single community institution could observe on its own, and that moves faster than formal regulatory advisories are typically published. Institutions evaluating their fraud-detection strategy should treat participation in such information-sharing arrangements as a complement to, not a substitute for, the in-house recalibration discipline described in Section 5, since shared signals still need to be translated into an individual

institution's own model and feature set to be actionable.

The legal basis for this kind of sharing already exists for U.S. institutions in the form of Section 314(b) information-sharing safe harbours, which permit participating financial institutions to share information about suspected money laundering or terrorist financing without certain civil-liability exposure that would otherwise attach to sharing customer information across institutions. Many community institutions that are legally eligible to participate in 314(b) sharing arrangements do not actively use them for fraud-typology intelligence specifically, treating the mechanism as an AML-only tool rather than as a channel that could also carry the kind of emerging bust-out or synthetic-identity pattern intelligence described above. Institutions that already have 314(b) participation in place for AML purposes should evaluate whether that same channel, or a parallel regional fraud-specific consortium, can be extended to serve the typology-intelligence function this section describes, since the marginal cost of extending an existing information-sharing relationship is typically lower than establishing a new one from scratch.

RECOMMENDATIONS

Five recommendations follow for community financial institutions and their technology partners. First, treat precision and recall as the primary monitored metrics for any fraud-detection model, not accuracy, given the severe class imbalance inherent to fraud data documented in Section 4. Second, build a defined, ideally automated recalibration cadence triggered by detected performance drift or documented typology shifts rather than a fixed annual review, given the pace of AI-accelerated typology change documented in Section 3. Third, size detection sensitivity to the institution's actual investigative capacity, following the precision-recall trade-off logic in Section 6, rather than defaulting to a maximum-recall setting that a small compliance team cannot realistically action. Fourth, pursue shared ownership of identity-risk signals across fraud, credit, and AML functions specifically to address synthetic identity fraud's tendency to persist across organisational silos, per Section 7. Fifth, favour explainable, well-governed ensemble architectures as the default deployment choice, reserving more complex network-aware or deep-learning architectures for specific, documented typology gaps rather than as a default starting point.

CONCLUSION

Community financial institutions face a fraud environment that is deteriorating faster than static, rule-based detection systems and, in many cases, faster than compliance staffing can absorb through manual review alone a record 3.6 million SARs were filed in 2022, and 79% of credit union and community bank leaders reported direct fraud losses exceeding $500,000 in 2023. Adaptive machine-learning approaches have demonstrated real, measurable detection-accuracy gains, but the evidence reviewed

here indicates that accuracy alone is an incomplete and sometimes misleading measure of a fraud-detection system's real-world value; recalibration discipline under concept drift, precision-recall balance sized to actual investigative capacity, and governance that spans fraud, credit, and AML functions together determine whether a model's benchmark performance translates into fewer sustained losses and a manageable, rather than overwhelming, suspicious-activity monitoring workload. The throughline connecting this article's sections is that detection accuracy and operational sustainability are separate problems requiring separate solutions. A model can achieve excellent published benchmark performance, of the kind summarised in Table 3, and still fail a community institution in practice if it generates more alerts than the institution's investigative staff can review, if its behaviour cannot be explained to an examiner or an affected customer, or if it is never recalibrated as the specific fraud typologies described in Section 3 continue to evolve. Institutions that treat model accuracy as the whole problem, rather than one input into a broader system encompassing capacity planning, governance, and recalibration discipline, are the ones most likely to find a well-validated model underperforming its own benchmark once deployed against their specific transaction volume, staffing, and typology mix.

REFERENCES

[1] Abdallah, A., Maarof, M. A., & Zainal, A. (2016). Fraud detection system: A survey. *Journal of Network and Computer Applications, 68*, 90–113.

[2] Aljawarneh, S., Vangipuram, R., & Alawneh, A. (2023). The imbalanced classification of fraudulent bank transactions using machine learning. *Mathematics, 11*(13), Article 2862.

[3] Alwadain, A., Ali, R. F., & Muneer, A. (2023). Estimating financial fraud through transaction-level features and machine learning. *Mathematics, 11*(5), Article 1184.

[4] Carcillo, F., Dal Pozzolo, A., Le Borgne, Y.-A., Caelen, O., Mazzer, Y., & Bontempi, G. (2018). SCARFF: A scalable framework for streaming credit card fraud detection with Spark. *Information Fusion, 41*, 182–194.

[5] Carcillo, F., Le Borgne, Y.-A., Caelen, O., & Bontempi, G. (2021). Streaming graph-based fraud detection. *Machine Learning, 110*, 123–145.

[6] Dal Pozzolo, A., Caelen, O., Johnson, R. A., & Bontempi, G. (2015). Calibrating probability with undersampling for unbalanced classification. In *2015 IEEE Symposium Series on Computational Intelligence* (pp. 159–166).

[7] Dal Pozzolo, A., Boracchi, G., Caelen, O., Alippi, C., & Bontempi, G. (2018). Credit card fraud detection: A realistic modeling and a novel learning strategy. *IEEE Transactions on Neural Networks and Learning Systems, 29*(8), 3784–3797.

[8] El-Hajj, M., & S. M. (2023). Machine learning as a tool for assessment and management of fraud risk in banking transactions. *Journal of Risk and Financial Management, 18*(3),

Article 130.

[9] Jurgovsky, J., Granitzer, M., Ziegler, K., Calabretto, S., Portier, P.-E., He-Guelton, L., & Caelen, O. (2018). Sequence classification for credit-card fraud detection. *Expert Systems with Applications, 100*, 234–245.

[10] Khandani, A. E., Kim, A. J., & Lo, A. W. (2010). Consumer credit-risk models via machine-learning algorithms. *Journal of Banking & Finance, 34*(11), 2767–2787.

[11] 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.

[12] Mytnyk, B., Tkachyk, O., Shakhovska, N., Fedushko, S., & Syerov, Y. (2023). Application of artificial intelligence for fraudulent banking operations recognition. *Big Data and Cognitive Computing, 7*(2), Article 93.

[13] 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.

[14] Pourhabibi, T., Ong, K.-L., Kam, B. H., & Boo, Y. L. (2020). Fraud detection: A systematic literature review of graph-based anomaly detection approaches. *Decision Support Systems, 133*, Article 113303.

[15] Srokosz, M., Bobyk, A., Ksiezopolski, B., & Wydra, M. (2023). Machine-learning-based scoring system for antifraud CISIRTs in banking environment. *Electronics, 12*(1), Article 251.

[16] Taha, A., & Malebary, S. J. (2020). An intelligent approach to credit card fraud detection using an optimized light gradient boosting machine. *IEEE Access, 8*, 25579–25587.

[17] Whitrow, C., Hand, D. J., Adams, N. M., Juszczak, P., & Weston, D. J. (2009). Transaction aggregation as a strategy for credit card fraud detection. *Data Mining and Knowledge Discovery, 18*(1), 30–55.

[18] Xie, Y., et al. (2023). Financial transaction fraud detector based on imbalance learning and graph neural network. *Applied Soft Computing, 149*, Article 110984.

[19] Zhang, Y., et al. (2023). Enhancing financial fraud detection with hierarchical graph attention networks: A study on integrating local and extensive structural information. *Finance Research Letters, 58*, Article 104458.

[20] Zheng, W., Yan, L., Gou, C., & Wang, F.-Y. (2021). Federated meta-learning for fraud detection in financial institutions. *IEEE Transactions on Computational Social Systems, 8*(6), 1517–1528.